# Asymmetry and Long Memory in Volatility Modelling*

**Manabu Asai**
Faculty of Economics
Soka University, Japan

**Michael McAleer**
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Institute of Economic Research
Kyoto University, Japan

**Marcelo C. Medeiros**
Department of Economics
Pontifical Catholic University of Rio de Janeiro, Brazil

**EI 2010-60**

Revised: October 2010

* The authors are most grateful to a Co-Editor, Associate Editor and two referees for very helpful comments and suggestions, and Marcel Scharth for efficient research assistance. For financial support, the first author acknowledges the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science, and Australian Academy of Science, the second author is most grateful to the Australian Research Council, National Science Council, Taiwan, and Japan Society for the Promotion of Science, and the third author wishes to acknowledge CNPq, Brazil.

## Abstract

A wide variety of conditional and stochastic variance models has been used to estimate latent volatility (or risk). In this paper, we propose a new long memory asymmetric volatility model which captures more flexible asymmetric patterns as compared with existing models. We extend the new specification to realized volatility by taking account of measurement errors, and use the Efficient Importance Sampling technique to estimate the model. As an empirical example, we apply the new model to the realized volatility of Standard and Poor's 500 Composite Index to show that the new specification of asymmetry significantly improves the goodness of fit, and that the out-of-sample forecasts and Value-at-Risk (VaR) thresholds are satisfactory. Overall, the results of the out-of-sample forecasts show the adequacy of the new asymmetric and long memory volatility model for the period including the global financial crisis.

**Keywords:** Asymmetric volatility, long memory, realized volatility, measurement errors, efficient importance sampling.

# 1 Introduction

The accurate specification and modelling of risk are integral to optimal portfolio selection and risk management using high frequency and ultra high frequency data. In this context, a wide variety of conditional and stochastic variance models has been used to estimate latent volatility (or risk) using high frequency data, while the availability of tick data has led to alternative models of realized volatility to estimate integrated volatility in analysing ultra high frequency data (see McAleer (2005) for a comprehensive review of univariate and multivariate, and symmetric and asymmetric, conditional and stochastic volatility models, and Asai, McAleer and Yu (2006) for a detailed review of alternative specifications and estimation algorithms for multivariate stochastic volatility models).

In the framework of diffusion processes, the daily variance of stock return is expressed as an integral of the intraday variance, which is called the integrated variance. If the microstructure noise is ignored, we may estimate the integrated variance by the sum of squared returns of ultra high frequency data. Such an estimator is called the realized variance, which corresponds to an estimate of the integrated variance, namely the true daily variance. In this paper, we refer to the square root of the integrated variance and of the realized variance as the Integrated Volatility (IV) and Realized Volatility (RV), respectively. For a recent extensive review of the RV literature, see McAleer and Medeiros (2008), and Bandi and Renò (2008), Todorov (2009) and Shephard and Sheppard (2010), among others, for more recent developments regarding the modelling and estimation of stochastic volatility using high frequency data.

Recent empirical results from the RV literature show two typical features in volatility, namely the asymmetric effect on volatility caused by previous returns, and the long-range dependence in volatility. The former issue has been investigated by Bollerslev and Zhou (2006), Bollerslev, Litovinova and Tauchen (2006), Bollerslev, Sizova and Tauchen (2010), Chen and Ghysels (2008), Martens, van Dijk and de Pooter (2009), and Patton and Sheppard (2010), among others. With respect to the latter point, the autoregressive fractionally integrated model has been used by Andersen, Bollerslev, Diebold and Labys (2001), Koopman, Jungbacker and Hol (2005) and Pong, Shackelton, Taylor and Xu (2004), among others, while other studies have used the heterogeneous autoregressive model of Corsi (2009) to approximate the hyperbolic decay rates associated with long memory models.

The purpose of the paper is to propose a new specification of the asymmetric and long memory volatility model, which allows flexible patterns in order to capture empirical regularities. Based on the general specification, we examine alternative stochastic volatility models that have recently been developed and estimated. Some of the corresponding SV models are in Harvey and Shephard (1996), Danielsson (1994), and Asai and McAleer (2005, 2010), with similar attempts having been considered by Bollerslev, Sizova and Tauchen (2010), Martens, van Dijk and de Pooter (2009), and Corsi and Renò (2010). Bollerslev, Sizova and Tauchen (2010) develop an equilibrium model with a continuous time long memory process, while our paper takes a discrete-time approach. Compared with Martens, van Dijk and de Pooter (2009) and Corsi and Renò (2010), our model incorporates a more general specification of the asymmetric effect and exact long memory process.

Upon estimating RV by using ultra high frequency data, one of the major problems that has arisen is that of microstructure noise. Several authors have proposed alternative methods for removing the microstructure noise (see, for example, Bandi and Russell (2006), Barndorff-Nielsen, Hansen, Lunde and Shephard (2008), Zhang, Mykland and Aït-Sahalia (2005), and Hansen, Large and Lunde (2008)). Some methods have provided bias-corrected and consistent estimators of the integrated variance, while other methods have not. Recently, Asai, McAleer and Medeiros (2009) have shown that, even when a bias-corrected and consistent estimator is used, non-negligible measurement errors remain in estimating and forecasting IV.

Barndorff-Nielsen and Shephard (2002) considered the decomposition of RV as the sum of IV and measurement error, which they call the RV error. In other words, RV is considered to be a proxy for IV. With respect to the third of our aims, we propose a new asymmetric model for RV by extending the general asymmetric volatility model, with an additional term to capture RV errors. It should be noted that introducing a correction for measurement error in the RV process renders the true volatility process unobservable. In order to estimate the proposed model, we employ the efficient importance sampling (EIS) ML method proposed by Liesenfeld and Richard (2003, 2006). The EIS evaluates the log-likelihood function of the model, including the latent process, by using simulations, such as the Monte Carlo Likelihood (MCL) technique of Durbin and Koopman (1997). Compared with the MCL method, the EIS method is applicable to various kinds of latent models (see the discussion in Liesenfeld and Richard (2003)).

The remainder of the paper is organized as follows. Section 2 develops a general long-memory asymmetric volatility model, and examines five kinds of asymmetric SV models. By using the structure of asymmetric effects, Section 3 proposes a new model for RV based on correcting for RV errors. Section 4 discusses the EIS-ML method, while Section 5 presents the empirical results for the RV model using Standard and Poor's 500 Composite Index, and evaluates the new specification of asymmetry with respect to goodness of fit, out-of-sample forecasts, and Value-at-Risk (VaR) thresholds. Section 6 gives some concluding remarks.

## 2 Structure of Asymmetric Volatility Models with Long Memory

In this section, we propose a new asymmetric volatility model, and compare it with stochastic volatility (SV) models that have recently been developed and estimated.

The return process is given by

$$r_t = m_t + V_t z_t, \quad z_t \sim \text{i.i.d.}(0,1),$$

where $m_t$ and $V_t$ are the time-varying mean and volatility processes, respectively, and $z_t$ is the standardized disturbance. We assume that the log-volatility follows an ARFIMA($p$,$d$,$q$) process,

$$\ln V_{t+1} = \alpha + (1-L)^{-d}\,\Phi^{-1}(L)\Theta(L)\xi_t, \qquad (1)$$

where $L$ is the lag-operator, $\Phi(L)$ and $\Theta(L)$ are the lag polynomials for the AR and MA coefficients, and $(1-L)^d$ is the fractional difference operator. As suggested by Nelson (1990, 1991) for conditional volatility models, the innovation term in the volatility equation plays an important role in considering asymmetry and leverage effects.

We suggest a generalized error, such that

$$\xi_t = \xi_t^* - E\left(\xi_t^*\right) + \eta_t, \quad \eta_t \sim N\left(0, \sigma_\eta^2\right),$$
$$\xi_t^* = \gamma_1 z_t + \gamma_2 \left|z_t\right| + \gamma_3 z_t I\left(0 \le z_t < \delta\right) + \gamma_3 \delta I\left(\delta \le z_t\right), \tag{2}$$

where $\gamma_1$, $\gamma_2$ and $\gamma_3$ are parameters, and $I\left(0 \le z < \delta\right)$ is the indicator function, which takes the value of one if $0 \le z < \delta$, and zero otherwise. The first two terms in $\xi_t^*$ play similar roles as in the EGARCH model. As shown in Harvey and Shephard (1996), the negative sign of the coefficient of $z_t$ produces the dynamic relationship between current return and future volatility, which is called the 'leverage' effect. Generally, a sufficient condition for univariate SV models to have a leverage effect is that $\xi_t$ is negatively correlated with $z_t$. For our new model, a negative sign for $\gamma_1$ is expected. Hence, $\gamma_1 z_t$ controls the leverage effect in the new model. On the other hand, $\gamma_2 \left|z_t\right|$ governs the size effect. When $\gamma_1 = \gamma_3 = 0$, the term makes the log-volatility increase according to the size of the standardized error.

Turning to the last two terms in $\xi_t^*$, they contribute to capturing asymmetric effects with greater flexibility. Figure 1 shows the relationship between $\xi$ and $z$, and implies that negative shocks and large positive shocks increase future volatility via $\xi$, but small positive shocks decrease volatility. Such a phenomenon has recently been observed in Chen and Ghysels (2008) with a semi-parametric method for realized volatility. Recently, Patton and Sheppard (2010) also attempt to explain it by considering the sign of jumps on the realized volatility measure.

We consider five special cases as follows:

Model (i) Equations (3) and (4), with restrictions $\gamma_1 = \gamma_2 = \gamma_3 = 0$.

Model (ii) Equations (5) and (6), with restrictions $\gamma_1 < 0$ and $\gamma_2 = \gamma_3 = 0$.

Model (iii) Equations (7) and (8), with $\gamma_3 = 0$ and $z_t$ replaced by $r_t - m_t$.

Model (iv) Equations (9) and (10), with $\gamma_3 = 0$ and $\gamma_2 |z_t|$ replaced by $\gamma_2 |r_t - m_t|$.

Model (v) Equations (11) and (12), with $\gamma_3 = 0$.

In order to understand these concepts, it is convenient to consider a simple AR(1) model of log-volatility. Setting $d = 0$, $\Phi(L) = 1 - \phi L$, and $\Theta(L) = 1$ in (13), we have

$$\ln V_{t+1} = (1-\phi)\alpha + \phi \ln V_t + \xi_t . \tag{14}$$

Taking $V_t$ as the latent process for the stochastic volatility, we may find the following correspondence. Model (i) is the basic SV model of Taylor (1982), which is symmetric as positive and negative shocks to returns have identical effects on future volatility. Model (ii) corresponds to the SV model suggested in Harvey and Shephard (1996), and re-examined by Yu (2005) (see also Asai and McAleer (2009) for a correction of Yu's (2005) news impact function). Model (iii) was proposed in Danielsson (1994), and was estimated in Asai and McAleer (2005). Model (iv) was suggested in Asai and McAleer (2005) to capture both leverage and asymmetric effects. Model (v) adapts the EGARCH model of Nelson (1991) to the SV literature, and was suggested and estimated in Asai and McAleer (2010). In contrast to Model (iii), Model (v) uses the standardized returns in forecasting future volatility, and can capture various types of asymmetric and leverage effects.

As compared with existing models, the new model in (15) and (16) allows log-volatility to follow the ARFIMA process, and incorporates more flexible asymmetric effects.

## 3 Model Specification for Realized Volatility

Let $p(t+\tau)$ be the logarithmic price of a given asset at time $\tau$ $(0 \le \tau \le 1)$ on day $t$ $(t = 1, 2, \ldots)$. We assume that $p(t+\tau)$ follows a continuous time diffusion process,

$$dp(t+\tau) = \mu(t+\tau)d\tau + \sigma(t+\tau)dW(t+\tau), \tag{17}$$

where $\mu(t+\tau)$ is the drift component, $\sigma(t+\tau)$ is the instantaneous volatility (or standard deviation), and $W(t+\tau)$ is a standard Brownian motion. Let $r_t$ be the daily return, defined as $r_t = p(t) - p(t-1)$. Conditionally on

$$\Im_t \equiv \Im\left\{\mu(t+\tau-1), \sigma(t+\tau-1)\right\}_{\tau=0}^{\tau=1},$$

which is the σ-algebra (information set) generated by the sample paths of $\mu(t+\tau-1)$ and $\sigma(t+\tau-1)$ $(0 \le \tau \le 1)$, we have

$$r_t \Big| \Im_t \sim N\left(\int_0^1 \mu(t+\tau-1)d\tau, \int_0^1 \sigma^2(t+\tau-1)d\tau\right).$$

The term $\int_0^1 \sigma^2(t+\tau-1)d\tau$ is known as the *integrated variance*, which is a measure of the day-*t* ex post volatility. The integrated variance is typically the object of interest as a measure of the true daily volatility.

With respect to the model of the instantaneous volatility, there are several specifications, which are called "continuous-time Stochastic Volatility (SV)" models (see Ghysels, Harvey and Renault (1996), for example). Hull and White (1987) allow the squared volatility to follow a diffusion process:

$$d\sigma^2 = \alpha\sigma^2 d\tau + \omega\sigma^2 dB\,, \tag{18}$$

where $B$ is a second Brownian motion, and $\alpha$ and $\omega$ are parameters. Here, we have omitted $(t+\tau)$ in order to simplify the notation. Hull and White (1987) assume a negative correlation between $W$ and $B$, thereby incorporating leverage effects. The model in (19) is closely related to the GARCH diffusion, which is derived as the diffusion limit of a sequence of GARCH(1,1) models (see Nelson (1990)).

Wiggins (1987) assumes that the log-volatility follows a Gaussian Ornstein-Uhlenbeck (OU) process:

$$d\log\sigma^2 = \alpha\left(\mu - \log\sigma^2\right)\sigma^2 d\tau + \omega dB\,. \tag{20}$$

In the specification, we may introduce leverage effects by assuming a negative correlation between $W$ and $B$. The asymmetric SV model of Harvey and Shephard (1996) is considered to be an Euler-Maruyama approximation of the continuous-time model (21), with negative correlation. Three major extensions of such diffusion-based SV models incorporate jumps to volatility process (Eraker, Johannes and Polson (2003)), model volatility as a function of a number of factors (Chernov et al. (2003)), and allow the log-volatility to follow a long memory process (Comte and Renault (1998)).

If the underlying process of the instantaneous volatility is a continuous-time SV model, the resulting integrated variance is still a stochastic process. At this stage, it may be useful to distinguish the differences and similarities among the conditional variance, stochastic variance, and integrated variance. As shown in Nelson (1990), it is possible to consider the diffusion limits of typical conditional variance models, such as the GARCH model and the exponential GARCH model of Nelson (1991). Hence, conditional variance models are considered to be approximations of continuous-time SV models. Alternative approximations are the (discrete-time) SV models of Taylor (1982) and Harvey and Shephard (1996), which are obtained by the Euler-Maruyama discretization of the continuous-time SV models. Compared with the class of GARCH models, discrete-time SV models give better approximations in the sense that the latter can be derived straightforwardly from continuous-time SV models. Therefore, the conditional and (discrete-time) stochastic variance can be considered as approximations of the integrated variance obtained by continuous-time SV models.

In the literature, there have been numerous extensions of GARCH models, while extensions of SV models are still being developed. There are many cases where it is not straightforward to consider a continuous-time SV model which corresponds to such an extension. For this reason, in the previous section we considered asymmetric long-memory models of the integrated volatility directly.

Although the integrated variance is unobservable, it is possible to estimate it using high frequency data. Such estimates are called "Realized Volatility (RV)". Zhang, Mykland and Aït-Sahalia (2005) and Barndorff-Nielsen, Hansen, Lunde and Shephard (2008) have proposed consistent estimator of the integrated variance, under the existence of microstructure noise (for extensive reviews of the RV literature, see Bandi and Russell (2006) and McAleer and Medeiros (2008)). As observed in Barndorff-Nielsen and Shephard (2002), we can always decompose RV as the sum of IV and a measurement error, which they call the 'RV error'. According to their analysis, even if we have a consistent estimator of IV, the RV contains a measurement error, which is not negligible.

At this stage we should consider the possible confusion regarding 'conditional' volatility. The RV is an estimator of IV, which is the ex-post daily variance of the price process conditional on the sigma algebra, defined after equation (22). However, this is quite different from the conditional volatility in the ARCH class, as the latter is conditional on the sigma algebra defined by past observed information, such as the return series (see the detailed discussion in Andersen, Bollerslev, Diebold and Labys (2001) and Andersen, Bollerslev and Diebold (2010)). Therefore, the conditional volatility based on the extensions of the ARCH models contains less information as compared with the IV.

Now, we specify the new asymmetric model for realized volatility (RV), noting the correspondence that $m_t = \int_0^1 \mu(t+\tau-1)d\tau$ , $V_t^2 = \int_0^1 \sigma^2(t+\tau-1)d\tau$ and $z_t \sim N(0,1)$ . Assume that the RV is a consistent estimator of integrated volatility (IV). Barndorff-Nielsen and Shephard (2002) refer to the measurement error, defined by the difference between RV and IV, as the RV error. Barndorff-Nielsen and Shephard (2002), Bollerslev and Zhou (2002) and Asai, McAleer and Medeiros (2009) showed it is useful to employ an ad-hoc approach which accommodates an error term with constant variance.

Let $y_t$ be the daily log RV, in which RV is a consistent estimate of IV. The new asymmetric model for RV to be analysed in the paper is given by

$$
\begin{aligned}
& y_t = \ln V_t + U_t, \quad E\left(U_t\right) = 0, \quad V\left(U_t\right) = \sigma_u^2, \\
& \ln V_{t+1} = \alpha + \left(1-L\right)^{-d} \Phi^{-1}\left(L\right)\Theta\left(L\right)\xi_t \\
& \xi_t = \xi_t^* - E\left(\xi_t^*\right) + \eta_t, \quad \eta_t \sim N\left(0, \sigma_\eta^2\right), \\
& \xi_t^* = \gamma_1 z_t + \gamma_2 \left|z_t\right| + \gamma_3 z_t I\left(0 \le z_t < \delta\right) - \gamma_3 \delta I\left(\delta \le z_t\right), \\
& z_t = r_t / V_t,
\end{aligned}
\tag{23}
$$

where $z_t$ is the standardized return and follows the standard normal distribution. This specification enables $U_t$ to capture the measurement errors in RV. We will refer to this model as the "RV-ARFIMA(*p*,*d*,*q*)-AS $\left(\gamma_1, \gamma_2, \gamma_3\right)$ -noise" model. The model allows various types of symmetric and/or asymmetric effects, long-memory property, and takes account of the realized volatility errors. If the measurement errors are neglected, we will have a special case with $\sigma_u = 0$. It should be noted that we consider the mean subtracted return, $r_t$, instead of return.

## 4 EIS-ML Estimation

The likelihood function for the asymmetric model in equation (24) includes high-dimensional integration, which is difficult to calculate numerically. We employ the Efficient Importance Sampling (EIS) method developed by Liesenfeld and Richard (2003, 2006) for evaluating the log-likelihood.

The pilot method for the EIS is the Accelerated Gaussian Importance Sampling (AGIS) approach, as developed in Danielsson and Richard (1993). The AGIS approach is designed to estimate dynamic latent variable models, where the latent variable follows a linear Gaussian process. While the AGIS technique has limited applicability, the EIS is applicable to models with more flexible classes of distributions and specifications for the latent variables. As in the case of AGIS, EIS is a Monte Carlo technique for the evaluation of high-dimensional integrals. The EIS relies on a sequence of simple low-dimensional least squares regressions to obtain a very accurate global approximation of the integrand. This approximation leads to a Monte Carlo sampler, which produces highly accurate Monte Carlo estimates of the likelihood.

### 4.1 Likelihood Evaluation via EIS

Let $y_t$ be an observable variable and $h_t = \ln V_t$ be a latent variable. We denote the joint density of $Y_T = \{y_t\}_{t=1}^{T}$ and $H_T = \{h_t\}_{t=1}^{T}$ as $f(Y_T, H_T; \theta)$, indexed by the unknown parameter vector $\theta$. In dynamic latent variable models, the joint density is typically formulated as:

$$f(Y_T, H_T; \theta) = \prod_{t=1}^{T} f(y_t, h_t | Y_{t-1}, H_{t-1}, \theta) = \prod_{t=1}^{T} g(y_t | h_t, Y_{t-1}, \theta)\, p(h_t | H_{t-1}, Y_{t-1}, \theta),$$

where $g(\cdot)$ denotes the conditional density of $y_t$ given $(h_t, y_{t-1})$, and $p(\cdot)$ the conditional density of $h_t$ given $(H_{t-1}, Y_{t-1})$. For ease of notation, it is assumed that the initial conditions are known constants, but EIS can easily accommodate alternative (stochastic) assumptions. It should be noted that we excluded the density of return series. This approach is not efficient, but the loss in efficiency is minor, by construction.

The likelihood function is given by the *T*-dimensional integral:

$$L(\theta; Y_T) = \int f(Y_T, H_T; \theta)\, dH_T \ ,$$

and a natural MC estimate of $L(\theta; Y_T)$ is given by

$$\hat{L}(\theta; Y_T) = \frac{1}{N} \sum_{i=1}^{N} \left[ \prod_{t=1}^{T} g\left(y_t \middle| \tilde{h}_t^{(i)}, Y_{t-1}, \theta\right) \right], \tag{25}$$

where $\left\{\tilde{h}_t^{(i)}(\theta)\right\}_{t=1}^{T}$ denotes a trajectory drawn from the sequence of $T$ densities. Each $\tilde{h}_t^{(i)}(\theta)$ is drawn from the conditional density $p\left(h_t \middle| \tilde{H}_{t-1}^{(i)}(\theta), Y_{t-1}, \theta\right)$.

In order to understand the EIS, we first note that EIS searches for a sequence of samplers that exploits the sample information on $h_t$ conveyed by $y_t$. Let $\left\{m\left(h_t \middle| H_{t-1}, x_t\right)\right\}_{t=1}^{T}$ denote a sequence of auxiliary samplers, indexed by the auxiliary parameters $X_n = \{x_t\}_{t=1}^{T}$. Regardless of the values of the auxiliary parameters, the likelihood function, $L(\theta; Y_T)$, is rewritten as

$$L(\theta; Y_T) = \int \prod_{t=1}^{T} \left[ \frac{f\left(y_t, h_t \middle| Y_{t-1}, H_{t-1}, \theta\right)}{m\left(h_t \middle| H_{t-1}, x_t\right)} \right] \prod_{t=1}^{T} m\left(h_t \middle| H_{t-1}, x_t\right) dH_T,$$

and the corresponding importance sampling MC estimate of the likelihood is given by

$$\tilde{L}(\theta; Y_T, X_T) = \frac{1}{N} \sum_{i=1}^{N} \left[ \prod_{t=1}^{T} \left\{ \frac{f\left(y_t, \tilde{h}_t^{(i)}(x_t) \middle| Y_{t-1}, \tilde{H}_{t-1}^{(i)}(x_{t-1}), \theta\right)}{m\left(\tilde{h}_t^{(i)}(x_t) \middle| \tilde{H}_{t-1}^{(i)}(x_{t-1}), x_t\right)} \right\} \right], \qquad (26)$$

where $\left\{\tilde{h}_t^{(i)}(x_t)\right\}_{t=1}^{T}$ denotes a trajectory drawn from the sequence of auxiliary importance samplers, *m*.

The EIS chooses a sequence of *m* densities by selecting values of the auxiliary parameters, $X_T$, which provide a good match between the product in the numerator and that in the denominator in equation (26) to minimize the MC sampling variance of $\tilde{L}(\theta; Y_T, X_T)$. In order to implement EIS, it requires constructing a positive functional approximation, $k(H_t; x_t)$, for the density $f\left(y_t, h_t \middle| Y_{t-1}, H_{t-1}, \theta\right)$, with the requirement that it be analytically integrable with respect to $h_t$. In Bayesian terminology, $k\left(H_t;x_t\right)$ plays a role of a density kernel for $m\left(h_t \middle| H_{t-1},x_t\right)$, which is then given by

$$m\left(h_t \middle| H_{t-1},x_t\right)=\frac{k\left(H_t;x_t\right)}{\chi\left(H_{t-1},x_t\right)}, \tag{27}$$

where $\chi\left(H_{t-1},x_t\right)=\int k\left(H_t;x_t\right)dh_t$. Then, the EIS requires solving a back-recursive sequence of low-dimensional least squares problems of the form:

$$\begin{aligned}\hat{x}_t=\arg\min_{x_t}\sum_{i=1}^{N}\Bigg[\ln\left\{f\left(y_t,\tilde{h}_t^{(i)}(\theta)\middle|Y_{t-1},\tilde{H}_{t-1}^{(i)}(\theta),\theta\right)\xi\left(\tilde{H}_t^{(i)}(\theta),x_{t+1}(\theta)\right)\right\}\\ -c_t-\ln k\left(\tilde{H}_t^{(i)}(\theta);x_t\right)\Bigg]^2,\end{aligned} \tag{28}$$

for $t:T\rightarrow 1$, with $\xi\left(H_T,x_{T+1}\right)\equiv 1$. As in equation (25), $\left\{\tilde{h}_t^{(i)}(\theta)\right\}_{t=1}^{T}$ denotes a trajectory drawn from the *p* densities, and the $c_t$ are unknown constants to be estimated jointly with $x_t$. If the density kernel $k\left(H_t;x_t\right)$ is chosen within the exponential family of distributions, the EIS least squares problems become linear in $x_t$ under the canonical representation of exponential kernels.

The EIS estimate of the likelihood function for a given value of $\theta$ is obtained by substituting $\left\{\hat{x}_t(\theta)\right\}_{t=1}^{T}$ for $\left\{x_t\right\}_{t=1}^{T}$ in equation (26). In order to obtain maximally efficient importance samplers, a small number of iterations of the EIS algorithm is required, where the natural samplers *p* are replaced by the previous stage importance samplers. For such iterations to converge to fixed values of the auxiliary parameters, $\hat{x}_t$, which are expected to produce optimal importance samplers, it is necessary to apply the technique of Common Random Numbers (CRNs).

### 4.2 Implementation Issues

As we consider the nonlinear ARFIMA($p,d,q$) process, it is not straightforward to incorporate it in the likelihood function. Hence, we suggest using an AR($J$) approximation of the $\mathrm{AR}(\infty)$ representation of the ARFIMA part, which is similar to the MA($J$) approximation of the FIEGARCH model by Bollerslev and Mikkelsen (1996), in the sense that the coefficient of the $J$-th lagged term is almost zero and is negligible for large $J$, such as $J$ = 1000.

Based on the above truncation, we have the distributions of $y_t$ and $h_t$. The RV-ARFIMA($p,d,q$)-$\mathrm{AS}(\gamma_1, \gamma_2, \gamma_3)$-noise model in equation (29) assumes that RVs, $y_t$, given the latent log-volatility, $h_t$, follow the normal distribution:

$$g(y_t | h_t, \theta) \propto \exp\left\{ -\frac{1}{2\sigma^2}(y_t - h_t)^2 \right\}.$$

Conditional on $(H_{t-1}, r_{t-1})$, the log-volatility, $h_t\ (= \ln V_t)$, follows the normal distribution:

$$p(h_t | H_{t-1}, r_{t-1}, \theta) \propto \exp\left\{ -\frac{1}{2\sigma_t^2}(h_t - \alpha - l_t)^2 \right\},$$

where

$$l_t = \begin{cases} 0 & \text{for} \quad t = 1 \\ \sum_{i=1}^{t-1} \lambda_i (h_{t-i} - \alpha) + \xi_{t-1}^* - E(\xi_{t-1}^*) & \text{for} \quad t = 2, \ldots, J \\ \sum_{i=1}^{J} \lambda_i (h_{t-i} - \alpha) + \xi_{t-1}^* - E(\xi_{t-1}^*) & \text{for} \quad t = J+1, \ldots, T \end{cases}$$

and

$$\sigma_t^2 = \begin{cases} \Lambda'_{J,J} C_J \Lambda_{J,J} + \sigma_\xi^2 & \text{for} \quad t = 1 \\ \Lambda'_{t,J} C_{J-t+1} \Lambda_{t,J} + \sigma_\eta^2 & \text{for} \quad t = 2, \dots, J \\ \sigma_\eta^2 & \text{for} \quad t = J+1, \dots, T \end{cases}$$

where $\xi_t^* = \gamma_1 r_t e^{-h_t} + \gamma_2 |r_t| e^{-h_t} + \gamma_3 r_t e^{-h_t} I\left(0 \le r_t e^{-h_t} < \delta\right) + \gamma_3 \delta I\left(\delta \le r_t e^{-h_t}\right)$, $\sigma_\xi^2$ is the variance of $\xi_t$ determined by $\left(\gamma_1, \gamma_2, \gamma_3, \delta, \sigma_\eta^2\right)$, $\Lambda_{t,J} = \left(\lambda_t, \lambda_{t+1}, \dots, \lambda_J\right)'$, and $C_j$ is the unconditional covariance matrix of $\left(h_1, \dots, h_j\right)'$. Note that it is assumed that $p\left(h_1\right)$ follows the normal distribution with mean zero and the unconditional variance of $h_t$. Regarding the initial distributions for $t = 2, \dots, J$, we used the decomposition:

$$p\left(h_1, \dots, h_J\right) = p\left(h_1\right) \prod_{t=2}^{J} p\left(h_t \middle| h_1, \dots, h_{t-1}, r_1, \dots, r_{t-1}\right).$$

which produces the combination of the conditional and unconditional mean and variance given above.

We chose *m* as the parametric extension of the natural samplers, *p*. Hence, the parameterization for *k* is given by

$$k\left(H_t; x_t, r_{t-1}\right) = p\left(h_t \middle| H_{t-1}, r_{t-1}, \theta\right) \zeta\left(h_t, x_t\right),$$

where the auxiliary function $\zeta\left(h_t, x_t\right)$ is itself a Gaussian density kernel. Under this parameterization, the natural sampler, *p*, cancels out in the least squares problem in equation (28), to the effect that $\ln \zeta\left(h_t, x_t\right)$ serves to approximate $\ln g\left(y_t \middle| h_t, Y_{t-1}, r_{t-1}, \theta\right)$ $+ \ln \chi\left(H_t, x_{t+1}, r_t\right)$. In particular, the appropriate auxiliary function for the asymmetric model is given by $\ln \zeta\left(h_t, x_t\right) = \exp\left(x_{1t} h_t + x_{2t} h_t\right)$, with $x_t = \left(x_{1t}, x_{2t}\right)$, and the density kernels of the importance samplers have the form

$$k\left(H_t;x_t,r_{t-1}\right)\propto \exp\left[-\frac{1}{2}\left\{\left(\frac{\lambda_t}{\sigma_t}\right)^2-2\left(\frac{\alpha+l_t}{\sigma_t^2}+x_{1t}\right)h_t+\left(\frac{1}{\sigma_t^2}-2x_{2t}\right)h_t^2\right\}\right].$$

Accordingly, the conditional mean and variance of $h_t$ on *m* are given by

$$\mu_{m,t}=\sigma_{m,t}^2\left(\frac{\alpha+l_t}{\sigma_t^2}+x_{1t}\right),\quad \sigma_{m,t}^2=\frac{\sigma_t^2}{1-2\sigma_t^2x_{2t}},\tag{30}$$

respectively. Integrating $k\left(H_t;x_t,r_{t-1}\right)$ with respect to $h_t$, and omitting irrelevant multiplicative factors, leads to the following expression for the integrating constant:

$$\chi\left(H_{t-1},x_t,r_{t-1}\right)\propto \exp\left\{\frac{\mu_{m,t}^2}{2\sigma_{m,t}^2}-\frac{\left(\alpha+l_t\right)^2}{2\sigma_t^2}\right\}.\tag{31}$$

Based on these functional forms, the computation of an EIS estimate of the likelihood for the asymmetric model requires the following steps:

**Step (0)**: Use the natural samplers, *p*, to draw *N* trajectories of the latent variable, $\left\{\tilde{h}_t^{(i)}\left(\theta\right)\right\}_{t=1}^T$.

**Step (*t*)**: $\left(t:T\rightarrow 1\right)$: Use these random draws to solve the back-recursive sequence of least squares problems, as defined in equation (28). The step *t* least squares problem is characterized by the following linear auxiliary regression:

$$\begin{aligned}&-\frac{1}{2\sigma^2}\left(y_t-h_t\right)^2+\ln\chi\left(\tilde{h}_t^{(i)}\left(\theta\right),\hat{x}_{t+1}\left(\theta\right)\right)\\&\quad=\text{constant}+x_{1t}\tilde{h}_t^{(i)}\left(\theta\right)+x_{2t}\left\{\tilde{h}_t^{(i)}\left(\theta\right)\right\}^2+u_t^{(i)},\quad i:1\rightarrow N,\end{aligned}$$

where $u_t^{(i)}$ denotes the regression error term. The initial condition for the integrating constant (in equation (31)) is given by $\chi\left(h_T, x_{T+1}, r_T\right) \equiv 1$.

**Step (T + 1)**: The EIS samplers, $\left\{m\left(h_t \middle| H_{t-1}, \hat{x}_t(\theta)\right)\right\}_{t=1}^{T}$, which are characterized by the conditional mean and variance given in equation (30), are used to draw $N$ trajectories $\left\{\tilde{h}_t^{(i)}\left(\hat{a}_t(\theta)\right)\right\}_{t=1}^{T}$, from which the EIS estimate of the likelihood is calculated according to equation (26).

We set $N = 50$, as Liesenfeld and Richard (2003) reported that 50 is sufficient for univariate and nonlinear latent variable models, such as SV. After 7-10 iterations, $\tilde{L}\left(\theta; Y_T, X_T, R_T\right)$ converged for each $\theta$. The next section gives the EIS-ML estimates for the asymmetric model of RV.

For the case of neglecting measurement errors (that is, $\sigma_u = 0$), $h_t$ is observable, so it is possible to perform maximum likelihood estimation without simulations. By comparing the log-likelihood with the EIS log-likelihood above, we have the conventional likelihood ratio test statistics, which follows the $\chi^2(1)$ distribution under the null hypothesis that $\sigma_u = 0$.

### 4.3 Monte Carlo Experiments

In this subsection we present the results of a Monte Carlo study to investigate the small sample performance of the estimation procedure presented in subsection 4.1. We generate $R$ simulated time series for RV-ARFIMA(1,$d$,0)-AS$\left(\gamma_1, \gamma_2, \gamma_3\right)$-noise model in equation (32) and for some given 'true' parameter vector $\theta$. Subsequently, we treat $\theta$ as unknown and estimate it for each series using the EIS maximum likelihood method described in subsections 4.1 and 4.2. We compute the sample mean, standard deviation and root mean squared error (RMSE) and compare it with the 'true' parameter value.

The ‘true’ parameter values for generating Monte Carlo samples are given in the first column of Table 1, which is obtained by our empirical analysis in Section 5. The results given in Table 1 are for the typical sample size $T = 2500$ with the number of iterations set to $R = 300$. Table 1 shows that the most of the values of the standard deviation are close to those of the RMSE, indicating that the bias in finite samples is negligible.

# 5 Empirical Results

## 5.1 Data and Preliminary Results

The empirical analysis focuses on the RV of Standard and Poor’s 500 Composite Index. In order to estimate the daily realized volatility, we use the two time scales estimator (TTSE) of Zhang, Mykland and Aït-Sahalia (2005) with five-minute grids, which is a consistent estimator of the daily realized volatility. The sample period is Jan/3/1996 to March/29/2007, giving $T = 2796$ observations of RV.

As a preliminary analysis, we consider the new Fractional Integrated EGARCH-*t* models given in Section 3 as

$$\begin{aligned} & r_t = \sigma_t z_t, \quad z_t \sim St(\nu), \\ & \ln \sigma_{t+1}^2 = \alpha + (1-L)^{-d} \, \Phi^{-1}(L)\Theta(L)\xi_t, \\ & \xi_t = \xi_t^* - E(\xi_t^*), \\ & \xi_t^* = \gamma_1 z_t + \gamma_2 |z_t| + \gamma_3 z_t I(0 \le z_t < \delta) - \gamma_3 \delta I(\delta \le z_t), \end{aligned} \tag{33}$$

where $St(\nu)$ denotes the standardized *t* distribution, with degrees of freedom given by *v*. Note that this model implicitly specifies that $\sigma_\eta = 0$, so that $\sigma_t$ is determined by the past information. We denote this as the FIEGARCH(*p*,*d*,*q*)-*t*-AS$(\gamma_1, \gamma_2, \gamma_3)$ model and, for the case *d*=0, as the EGARCH(*p*,*q*)-*t*-AS$(\gamma_1, \gamma_2, \gamma_3)$ model.

We estimated two kinds of models, namely EGARCH(1,1)-*t*-AS $(\gamma_1,\gamma_2,\gamma_3)$ and FIEGARCH(1,*d*,1)-*t*-AS$(\gamma_1,\gamma_2,0)$. Table 2 shows the ML estimates of these models, with initial values of 1000. For the former model, all the estimated parameters, except for $\alpha$ and $\gamma_3$, are significant at the five percent level. The estimate of $\phi$ is close to 0.99, showing high persistence in volatility. The estimate of $\gamma_1$ is negative, while that of $\gamma_2$ is positive. The estimate of $1/\nu$ is 0.08, indicating that the estimate of $\nu$ is close to 13. The results are typical for the EGARCH-*t* specification. For the long memory model, all the estimated parameters, except for $\gamma_1$ and $1/\nu$, are significant. This specification shows the lack of importance of asymmetric effects and heavy-tailed conditional distributions. The AIC and BIC favour the FIEGARCH(1,*d*,1)-*t*-AS$(\gamma_1,\gamma_2,0)$ model. Similar results are also found in the literature with the FIEGARCH-*t* specification.

### 5.2 Estimates for RV Models

In the following, we will show that the empirical results for RV models are substantially different from those associated with EGARCH models. It should be noted that it is inadequate to compare the log-likelihood of EGARCH models with that of RV models as the former is based on $r_t$ while the latter is based on the RV, $y_t$. Furthermore, the fat tails of the conditional distribution of $r_t$ are irrelevant for the estimation of the RV model.

Table 3 shows the EIS-ML results of the RV-AR(1)-AS $(\gamma_1,\gamma_2,\gamma_3)$ -noise model. Regarding asymmetry, we consider four specifications, namely AS$(0,0,0)$, AS$(\gamma_1,0,0)$, AS$(\gamma_1,\gamma_2,0)$, and AS$(\gamma_1,\gamma_2,\gamma_3)$. All the estimated parameters are significant at the 5% level. As the AS$(\gamma_1,\gamma_2,\gamma_3)$ model has the smallest AIC and BIC, we report the empirical results only for this specification.

The estimate of $\sigma_u$ is close to 0.4, showing that the RV errors are not negligible. The estimate of $\phi$ is 0.986, while that of $\sigma_\eta$ is 0.11, which are typical of SV models. The estimate of $\gamma_1$ is negative, while that of $\gamma_2$ is positive. Unlike the estimates of the EGARCH model, the estimate of $\gamma_3$ is negative and significant. Figure 2 gives the news impact from $z_t$ to $\ln V_{t+1}$, showing that negative shocks and large positive shocks increase future volatility, but small positive shocks decrease volatility.

Table 4 presents the EIS-ML results for the RV-ARFIMA(1,*d*,0)-AS$\left(\gamma_1,\gamma_2,\gamma_3\right)$-noise model. As before, we consider four kinds of asymmetric effects. The AIC and BIC selected the AS$\left(\gamma_1,\gamma_2,\gamma_3\right)$ model, so we will concentrate the empirical analysis on this model. All the estimated parameters are significant at the five percent level. The estimate of $\sigma_u$ is close to 0.4, indicating that the RV errors are not negligible. The estimate of *d* is 0.47, showing that the log-volatility has long memory and is a stationary process. The estimate of $\phi$ is positive and close to 0.4, which is against the typical value of -0.1 in the RV literature. The difference can be explained by the existence of RV errors, $U_t = y_t - \ln V_t$. As shown in the Monte Carlo experiments of Asai, McAleer and Medeiros (2009), even minor RV errors can cause bias in the estimates if the RV error is neglected in estimation. The signs of $\gamma_1$, $\gamma_2$ and $\gamma_3$ are the same as in the case of Table 3. Figure 3 shows the news impact from $z_t$ to $\ln V_{t+1}$.

From Tables 2 and 3, we find that the RV-ARFIMA(1,*d*,0)-AS$\left(\gamma_1,\gamma_2,\gamma_3\right)$-noise model has the smallest AIC, while BIC chooses the RV-AR(1)-AS$\left(\gamma_1,\gamma_2,\gamma_3\right)$-noise model. These tables indicate that having the additional term, $\gamma_3$, significantly improves the goodness of fit of the model.

### 5.3 Forecasting Analysis

Regarding the RV-ARFIMA(1,*d*,0)-AS $\left(\gamma_1,\gamma_2,\gamma_3\right)$ -noise model, we examine the performance of the out-of-sample forecasts using the following four approaches: (i) test for equal forecast accuracy; (ii) test model specification; (iii) test the forecasts of the VaR thresholds; (iv) model selection. The benchmark model is the Leverage Heterogeneous Autoregressive (LHAR) model, suggested in Corsi and Renò (2010). The LHAR model is based on the Heterogeneous Autoregressive (HAR) model of Corsi (2009), which approximates a long memory process, with an extension regarding the leverage effect. Hence, the LHAR model accommodates both long range dependence and the leverage effect. The LHAR model is given by

$$y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 (y_t)_{t-5} + \beta_4 (y_t)_{t-20} + \beta_5 r_{t-1} I[r_{t-1} < 0] \\ + \beta_6 (r_t)_{t-5} I[(r_t)_{t-5} < 0] + \beta_7 (r_t)_{t-20} I[(r_t)_{t-20} < 0] + \text{error},$$

where $(y_t)_{t-h}$ denotes the *h*-horizon normalized realized volatility, defined by

$$(y_t)_{t-h} = \frac{y_{t-1} + y_{t-2} + \cdots + y_{t-h}}{h},$$

and $(r_t)_{t-h}$ is defined by the same manner. $I[r < 0]$ is the indicator function which take one if *r* is negative, and zero otherwise. A similar model is suggested by Martens, van Dijk and de Pooter (2009). Note that it is possible to include the positive part of heterogeneous returns, but they are usually insignificant.

Fixing the sample size at 2,500, we re-estimated the model and computed one-step-ahead forecasts of log-volatility for the last 150 days.

First, we report the result for the Harvey, Leybourne, and Newbold (1997) modification of the Diebold and Mariano (1995) test of equal predictive accuracy. The new asymmetric and long-memory volatility model is compared against the LHAR model. The test statistic follows the standard normal distribution asymptotically under the null hypothesis of equal accuracy. Table 5 shows the test results, indicating the difference between the two forecasts.

Second, we test the model specification, based on the Mincer-Zarnowitz regression, namely

$$x_t = a + b\hat{x}_{t|t-1} + e_t, \quad t = 1, 2, \ldots, 150$$

where $x_t$ can be the observed RV or log-RV on day *t*, and $\hat{x}_{t|t-1}$ is the one-step-ahead forecast of $x_t$ on day *t*. If the model is correctly specified, then $a = 0$ and $b = 1$. Table 6 show the estimates of the coefficients and the heteroskedasticity-consistent *F* test statistics for the joint null hypothesis, regarding the LHAR model and the RV-ARFIMA(1,*d*,0)-AS$\left(\gamma_1, \gamma_2, \gamma_3\right)$-noise model, respectively. With respect to the LHAR model, the *F* tests in both cases rejected the null hypothesis that the model is correctly specified. However, for the new asymmetric and long-memory model, the *F* test did not reject the null hypothesis. As the new model is based on log-RV, the estimates for log-RV are very close to the values expected under the null hypothesis.

Third, we calculated the VaR thresholds, accommodating the filtered historical simulation (FHS) approach, which is an effective method for predicting VaR thresholds (see Kuester et al. (2006) for some recent studies regarding the FHS approach). In short, the FHS approach estimates the empirical distribution of the standardized returns, then obtains the 100*p* percentiles to compute the 100*p* percent VaR thresholds. In our analysis, each time we estimated the model with 2,500 observations, we computed the 100*p* percentiles of the empirical distribution based on the last 500 observations, discarding the first 2,000 observations. Combined with the one-day-ahead forecasts of log-volatility, we computed the 100*p* percent VaR thresholds.

In order to assess the estimated VaR thresholds, the unconditional coverage and independence tests developed by Christoffersen (1998) are widely used. A drawback of the Christoffersen (1998) test for independence is that it tests against a particular alternative of a first-order dependence. The duration-based approach in Christoffersen and Pelletier (2004) allows for testing against more general forms of dependence but still requires a specific alternative. Recently, Candelon et al. (2010) have developed a more robust procedure which does not need a specific distributional assumption for the durations under the alternative. Consider the “hit sequence” of VaR violations, which takes a value of one if the loss is greater than the VaR threshold, and takes the value zero if the VaR is not violated. If we could predict the VaR violations, then that information may help to construct a better model. Hence, the hit sequence of violations should be unpredictable, and should follow an independent Bernoulli distribution with parameter *p*, indicating that the duration of the hit sequence should follow a geometric distribution The GMM duration-based test developed by Candelon et al. (2010) works with the J-statistic based on the moments defined by the orthonormal polynomials associated with the geometric distribution. The conditional coverage test and independence test based on *q* orthnormal polynomials have asymptotic $\chi_q^2$ and $\chi_{q-1}^2$ distributions under their respective null distributions. The unconditional coverage test is given as a special case of the conditional coverage test with $q = 1$.

Table 6 shows the percentage of VaR violations and test results for the LHAR model and new asymmetric and long-memory volatility model, respectively. For both the LHAR model and RV-ARFIMA(1,*d*,0)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise model, the tests did not reject the null hypothesis for the 5% and 1% VaR thresholds, indicating that the estimated VaR thresholds are satisfactory. We also conducted the unconditional coverage and independence tests developed in Christoffersen (1998), and the results are unchanged.

Finally, we select the forecasts using the following MZ equation:

$$x_t = a + b_1 \hat{x}_{t|t-1}^{AS} + b_2 \hat{x}_{t|t-1}^{LHAR} + e_t, \quad t = 1, 2, \ldots, 150$$

where $\hat{x}_{t|t-1}^{i} \left( i = AS, LHAR \right)$ is the one-step-ahead forecast of $x_t$ on day *t*, based on the RV-ARFIMA(1,*d*,0)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise model (*AS*) and the LHAR model. We select the forecasts by the conventional *t* test. As before, we consider two dependent variables, namely volatility and log-volatility. Table 8 gives the results. In both cases, the coefficients of $\hat{x}_{t|t-1}^{LHAR}$ are insignificant, indicating that the data prefer the forecasts of the RV-ARFIMA(1,*d*,0)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise model. Overall, the results of the out-of-sample forecasts favour our new asymmetric and long memory volatility model.

### 5.4 Global Financial Crisis

In addition to the previous analysis, we examine the adequacy of the new RV-ARFIMA(1,*d*,0)-AS $(\gamma_1, \gamma_2, \gamma_3)$ -noise model for the period including the global financial crisis, starting from the bankruptcy of Lehman Brothers, that is, Sep/15/2008.

For the analysis, we chose IBM as the individual stock for the period Jan/03/2000 to April/27/2009, giving $T$ = 2334 observations for RV. We obtained one-step-ahead forecasts as before for the last 150 observations corresponding to the period starting from the bankruptcy of Lehman Brothers. We use the LHAR model as a benchmark.

Table 9 gives the estimates for the MZ equations. The new model does not reject the null hypothesis, $a = 0$ and $b = 1$, showing the adequacy of the new model, while the LHAR model does reject the null hypothesis. The results with two forecasts show the significance of the forecast of the new model and insignificance for the LHAR model.

We also conducted the HLN tests for volatility and log-volatility, producing values of the test statistics of 4.97 and 9.14, respectively. The results indicate the superiority of the RV-ARFIMA(1,$d$,0)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise model, rejecting equal predictive accuracy. For the forecasting period, the stock price of IBM is so volatile that there is no day in which a negative return exceeds the boundary of -1.98 times RV. Hence, we cannot conduct tests of the VaR thresholds. Thus, we report that the number of violations for the 1% threshold is zero for the new model, while it is 4 times (0.027%) for the LHAR model. Overall, the results of the out-of-sample forecasts show the adequacy of the new asymmetric and long memory volatility model for the period including the global financial crisis.

## 6 Concluding Remarks

We proposed a new asymmetric and long-memory volatility model. Regarding the leverage effect, the new model sensitively captures the effects of both large and small, and positive and negative, shocks. Based on the new specification, this paper examined alternative univariate volatility models that have recently been developed and estimated.

We extended the specification of asymmetric and long memory volatility in order to model RV by taking account of the RV errors. This is a general model which includes not only various kinds of asymmetric effects, but also short and long memory specifications. We applied the EIS-ML method to estimate the model of RV, and reported the results for a Monte Carlo experiment.

The empirical results for the RV of Standard and Poor's 500 Composite Index showed the existence of RV errors. The estimates of the short and long memory models supported the new specification of asymmetric effect, which satisfies the following three conditions: (i) negative shocks to returns increase future volatility; (ii) large positive shocks to returns increase future volatility, but a negative shock has a larger effect on volatility than does a positive shock of equal magnitude; and (iii) small positive shocks to returns decrease future volatility. Overall, the new specification of asymmetry significantly improved the goodness of fit, and the out-of-sample forecasts and VaR thresholds were satisfactory.

## References

Andersen, T.G., Bollerslev, T., and F.X. Diebold (2010), "Parametric and Nonparametric Volatility Measurement", in: L.P. Hansen and Y. A¨ıt-Sahalia (eds.), *Handbook of Financial Econometrics*, Amsterdam: North-Holland, pp. 67-138.

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

Asai, M. and M. McAleer (2005), "Dynamic Asymmetric Leverage in Stochastic Volatility Models", *Econometric Reviews*, **24**, 317-332.

Asai, M. and M. McAleer (2010), "Alternative Asymmetric Stochastic Volatility Models", to appear in *Econometric Reviews*.

Asai, M. and M. McAleer (2009), "Multivariate Stochastic Volatility, Leverage and News Impact Surfaces", *Econometrics Journal*, **12**, 292-309.

Asai, M., M. McAleer and M. Medeiros (2009), "Modelling and Forecasting Daily Volatility with Noisy Realized Volatility Measures", Unpublished paper, Soka University.

Asai, M., M. McAleer and J. Yu (2006), "Multivariate Stochastic Volatility: A Review", *Econometric Reviews*, **25**, 145-175.

Bandi, F.M. and R. Renò (2008), "Nonparametric Stochastic Volatility", Unpublished paper, Graduate School of Business, University of Chicago.

Bandi, F.M. and J.R. Russell (2006), "Market Microstructure Noise, Integrated Variance Estimators, and the Accuracy of Asymptotic Approximations", unpublished paper, Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O.E., and N. Shephard (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models", *Journal of the Royal Statistical Society, Series B*, **64**, 253□280.

Barndorff-Nielsen, O.E., P.H. Hansen, A. Lunde and N. Shephard (2008), "Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise", *Econometrica*, **76**, 1481-1536.

Bollerslev, T. and H.O. Mikkeslen (1996), "Modeling and Pricing Long Memory in Stock Market Volatility", *Journal of Econometrics*, **73**, 151-184.

Bollerslev, T. and H. Zhou (2002), "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility", *Journal of Econometrics*, **109**, 33-65.

Bollerslev, T. and H. Zhou (2006), "Volatility Puzzles: A Simple Framework for Gauging Return-Volatility Regressions", *Journal of Econometrics*, **131**, 123-150.

Bollerslev, T., J. Litovinova and G. Tauchen (2006), "Leverage and Volatility Feedback Effects in High-Frequency Data", *Journal of Financial Econometrics*, **4**, 353-384.

Bollerslev, T., N. Sizova and G. Tauchen (2010), "Volatility in Equilibrium: Asymmetries and Dynamic Dependencies", Unpublished Paper, Department of Economics, Duke University.

Candelon, B., G. Colletaz, C. Hurlin, and S. Tokpavi (2010), "Backtesting Value-at-Risk: A GMM Duration-Based Approach", to appear in *Journal of Financial Econometrics*.

Chen, X. and E. Ghysels (2008), "News - Good or Bad - and its Impact on Volatility Predictions over Multiple Horizons", Unpublished paper, University of North Carolina at Chapel Hill.

Chernov, M., A.R. Gallant, E. Ghysels and G. Tauchen (2003), "Alternative Models for Stock Price Dynamics", *Journal of Econometrics*, **116**, 225-257.

Christoffersen, P. (1998), "Evaluating Interval Forecasts," *International Economic Review*, **39**, 841–62.

Christoffersen, P. and D. Pelletier (2004), "Backtesting Value-at-Risk: A Duration-Based Approach", *Journal of Financial Econometrics*, **2**, 84-108.

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

Corsi, F. (2009), “A Simple Approximate Long-Memory Model of Realized Volatility,” *Journal of Financial Econometrics*, **7**, 174-196.

Corsi, F. and R. Renò (2010), “HAR Volatility Modelling with Heterogeneous Leverage and Jumps”, Unpublished Paper, Università di Siena.

Danielsson, J. (1994), “Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood”, *Journal of Econometrics*, **64**, 375-400.

Danielsson, J., and J.-F. Richard (1993), “Quadratic Acceleration for Simulated Maximum Likelihood Evaluation”, *Journal of Applied Econometrics*, **8**, 153-173.

Diebold, F. X., and R. S. Mariano (1995), “Comparing Predictive Accuracy,” *Journal of Business and Economic Statistics*, **13**, 253-263.

Durbin, J., and S.J. Koopman (1997), “Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models”, *Biometrika*, **84**, 669-684.

Eraker, B., M. Johannes, and N.G. Polson (2003), “The Impact of Jumps in Returns and Volatility”, *Journal of Finance*, **53**, 1269-1300.

Ghysels, E., A.C. Harvey and E. Renault (1996), “Stochastic Volatility”, in C.R. Rao and G.S. Maddala (eds.), *Statistical Methods in Finance*, Amsterdam: North-Holland, pp. 119-191.

Hansen, P.R., J. Large and A. Lunde (2008), “Moving Average-based Estimators of Integrated Variance”, *Econometric Reviews*, **27**, 79-111.

Harvey, D., S. Leybourne, and P. Newbold (1997), “Testing the Equality of Mean Squared Errors,” *International Journal of Forecasting*, **13**, 281-291.

Harvey, A.C. and N. Shephard (1996), “Estimation of an Asymmetric Stochastic

Volatility Model for Asset Returns", *Journal of Business and Economic Statistics*, **14**, 429-434.

Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatility", *Journal of Finance*, **42**, 281-300.

Koopman, S.J., B. Jungbacker and E. Hol (2005), "Forecasting Daily Variablity of the S&P 100 Stock Index Using Historical Realized and Implied Volatility Measurements", *Journal of Empirical Finance*, **12**, 445-475.

Kuester, K., S. Mittnik and M.S. Paolella (2006), "Value-at-Risk Prediction: A Comparison of Alternative Strategies", *Journal of Financial Econometrics*, **4**, 53-89.

Liesenfeld, R., and J.-F. Richard (2003), "Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics", *Journal of Empirical Finance*, **10**, 505-531.

Liesenfeld, R., and J.-F. Richard (2006), "Classical and Bayesian Analysis of Univariate and Multivariate Stochastic Volatility Models", *Econometric Reviews*, **25**, 335-361.

Martens, M., D. van Dijk, and M. de Pooter (2009), "Forecasting S&P 500 Volatility: Long Memory, Level Sifts, Leverage Effects, Day-of-the-Week Seasonality, and Macroeconomic Announcements", *International Journal of Forecasting*, **25**, 282-303.

McAleer, M. (2005), "Automated Inference and Learning in Modeling Financial Volatility", *Econometric Theory*, **21**, 232-261.

McAleer, M. and M. Medeiros (2008), "Realized Volatility: A Review", *Econometric Reviews*, **27**, 10-45.

Nelson, D.B. (1990), "ARCH Models as Diffusion Approximations", *Journal of Econometrics*, **45**, 7–38.

Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, **59**, 347-370.

Patton, A. and K. Sheppard (2010), “Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility”, Unpublished Paper, Duke University.

Pong S., M.B. Shackelton, S.J. Taylor and X. Xu (2004), “Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA models, *Journal of Banking and Finance*, **28**, 2541-2563.

Shephard, N. and K. Sheppard (2010), “Realising the Future: Forecasting with High-Frequency Based Volatility (HEAVY) Models”, *Journal of Applied Econometrics*, **23**, 197-231.

Taylor, S.J. (1982), “Financial Returns Modelled by the Product of Two Stochastic Processes - A Study of Daily Sugar Prices 1961-79”, in O. D. Anderson (Ed.), *Time Series Analysis: Theory and Practice, 1*, Amsterdam: North-Holland, pp. 203–226.

Todorov, V. (2009), “Estimation of Continuous-Time Stochastic Volatility Models with Jumps Using High-Frequency Data”, *Journal of Econometrics*, **148**, 131-148.

Wiggins, J.B. (1987), “Option Values Under Stochastic Volatility: Theory and Empirical Estimates”, *Journal of Financial Economics*, **19**, 351-372.

Yu, J. (2005), “On Leverage in a Stochastic Volatility Model”, *Journal of Econometrics*, **127**, 165-178.

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**, 1394 – 1411.

**Table 1: Monte Carlo Results for EIS-ML Estimator for RV-AR(1)-AS$\left(\gamma_1,\gamma_2,\gamma_3\right)$-noise Model**

| Parameters | True | Mean | Standard deviation | RMSE |
|---|---|---|---|---|
| $d$ | 0.4727 | 0.4588 | (0.0366) | [0.0392] |
| $\phi$ | 0.4373 | 0.4453 | (0.0655) | [0.0660] |
| $\sigma_\eta$ | 0.1739 | 0.1792 | (0.0156) | [0.0165] |
| $\alpha$ | 0.0046 | -0.0018 | (0.0019) | [0.0067] |
| $\gamma_1$ | -0.0274 | -0.0160 | (0.0156) | [0.0193] |
| $\gamma_2$ | 0.0511 | 0.1179 | (0.0155) | [0.0686] |
| $\gamma_3$ | -0.2428 | -0.2954 | (0.0541) | [0.0754] |
| $\delta$ | 0.8841 | 0.8413 | (0.1262) | [0.1332] |
| $\sigma_u$ | 0.3858 | 0.3845 | (0.0089) | [0.0090] |

**Table 2: ML Estimates of the New EGARCH Class**

| Parameters | New EGARCH-*t* | | FIEGARCH(1,*d*,0)-*t* | |
|---|---|---|---|---|
| $d$ | | | 0.4067 | (0.0256) |
| $\phi$ | 0.9877 | (0.0038) | -0.2651 | (0.0450) |
| $\alpha$ | 0.1397 | (0.4560) | 0.3101 | (0.0777) |
| $\gamma_1$ | -0.0936 | (0.0138) | 0.0115 | (0.0441) |
| $\gamma_2$ | 0.0670 | (0.0204) | 2.4407 | (0.0776) |
| $\gamma_3$ | -0.0783 | (0.0967) | | |
| $\delta$ | 0.7259 | (0.4408) | | |
| $1/\nu$ | 0.0784 | (0.0191) | 0.0004 | (0.0410) |
| Log-Like | -2419.93 | | -1721.05 | |
| AIC | 4853.86 | | 3454.09 | |
| BIC | 4892.31 | | 3489.71 | |

**Note**: Standard errors are in parentheses. The first 1,000 observations are used for the initial values for the FIEGARCH-*t* model.

**Table 3: EIS Estimates of RV-AR(1)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise**

| Parameters | AS$(0,0,0)$ | AS$(\gamma_1,0,0)$ | AS$(\gamma_1,\gamma_2,0)$ | AS$(\gamma_1,\gamma_2,\gamma_3)$ |
|---|---|---|---|---|
| $\phi$ | 0.9747 | 0.9728 | 0.9870 | 0.9856 |
| | (0.0051) | (0.0040) | (0.0044) | (0.0044) |
| $\sigma_\eta$ | 0.1478 | 0.1111 | 0.1110 | 0.1103 |
| | (0.0091) | (0.0075) | (0.0074) | (0.00720) |
| $\alpha$ | -0.3148 | -0.1795 | -0.8439 | 1.2246 |
| | (0.1091) | (0.0788) | (0.3227) | (0.5127) |
| $\gamma_1$ | | -0.0681 | -0.0649 | -0.0418 |
| | | (0.0046) | (0.0043) | (0.0062) |
| $\gamma_2$ | | | 0.0424 | 0.0561 |
| | | | (0.0074) | (0.0079) |
| $\gamma_3$ | | | | -0.1934 |
| | | | | (0.0471) |
| $\delta$ | | | | 0.4902 |
| | | | | (0.0605) |
| $\sigma_u$ | 0.4054 | 0.4092 | 0.4125 | 0.4116 |
| | (0.0073) | (0.0067) | (0.0067) | (0.0067) |
| Log-Like | -1921.94 | -1821.52 | -1806.24 | -1793.51 |
| AIC | 3851.88 | 3653.04 | 3624.48 | 3603.03 |
| BIC | 3875.63 | 3682.72 | 3660.10 | 3650.52 |

**Note**: Standard errors are in parentheses.

**Table 4: EIS Estimates of RV-ARFIMA(1,*d*,0)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise**

| Parameters | AS$(0,0,0)$ | AS$(\gamma_1,0,0)$ | AS$(\gamma_1,\gamma_2,0)$ | AS$(\gamma_1,\gamma_2,\gamma_3)$ |
|---|---|---|---|---|
| $d$ | 0.4955 | 0.4987 | 0.4748 | 0.4727 |
| | (0.0039) | (0.00089) | (0.0090) | (0.0076) |
| $\phi$ | 0.3261 | 0.3676 | 0.4166 | 0.4373 |
| | (0.0603) | (0.0438) | (0.0538) | (0.0291) |
| $\sigma_\eta$ | 0.2416 | 0.1750 | 0.1852 | 0.1739 |
| | (0.0225) | (0.0147) | (0.0157) | (0.0080) |
| $\alpha$ | -0.5832 | 0.0021 | 0.0051 | 0.0046 |
| | (0.2394) | (0.0026) | (0.0027) | (0.0020) |
| $\gamma_1$ | | -0.0865 | -0.0827 | -0.0275 |
| | | (0.0061) | (0.0063) | (0.0075) |
| $\gamma_2$ | | | 0.0226 | 0.0511 |
| | | | (0.0077) | (0.0076) |
| $\gamma_3$ | | | | -0.2428 |
| | | | | (0.0287) |
| $\delta$ | | | | 0.8841 |
| | | | | (0.0196) |
| $\sigma_u$ | 0.3648 | 0.3844 | 0.3827 | 0.3858 |
| | (0.01208) | (0.0081) | (0.0085) | (0.0067) |
| Log-Like | -1908.27 | -1819.30 | -1811.37 | -1792.16 |
| AIC | 3826.54 | 3650.60 | 3636.75 | 3602.31 |
| BIC | 3856.22 | 3686.22 | 3678.30 | 3655.74 |

**Note**: Standard errors are in parentheses.

**Table 5: HLN Tests for Equal Forecast Accuracy**

| HLN | Test Stat. | P-value |
|---|---|---|
| Volatility | 2.5688 | 0.0102 |
| Log-Volatility | 4.8427 | 0.0000 |

**Note**: HLN is the test for equal forecast accuracy of Harvey, Leybourne, and Newbold (1997), where the new asymmetric volatility model is compared with LHAR. The test statistic follows the standard normal distribution asymptotically under the null hypothesis of equal accuracy.

**Table 6: Tests for Model Specification by MZ Equation**

$$x_t = a + b\hat{x}_{t|t-1} + e_t$$

| Model | LHAR | | RV-ARFIMA(1,*d*,0)-AS$(\gamma_1, \gamma_2, \gamma_3)$-noise | |
|---|---|---|---|---|
| Dependent variable | Volatility | Log-Volatility | Volatility | Log-Volatility |
| Constant | 0.1899<br>(0.0492) | -0.4098<br>(0.0740) | -0.2352<br>(0.1381) | -0.0045<br>(0.1305) |
| Forecast | 0.6387<br>(0.0798) | 0.6218<br>(0.0753) | 1.758<br>(0.4208) | 0.9838<br>(0.1109) |
| *F* test | 9.6675<br>[0.0078] | 26.292<br>[0.0000] | 4.1500<br>[0.1256] | 0.2790<br>[0.8698] |

**Note**: Heteroskedasticity-consistent standard errors are in parentheses, and *p*-values are in brackets. '*F* test' denotes the value of the heteroskedasticity-robust *F* test for the null hypothesis $H_0 : a = 0, b = 1$.

**Table 7: Backtesting VaR Thresholds**

| Model | LHAR | | | | RV-ARFIMA(1,*d*,0)-AS$(\gamma_1,\gamma_2,\gamma_3)$-noise | | | |
|---|---|---|---|---|---|---|---|---|
| VaR | % Violation | UC | ID | CC | % Violation | UC | ID | CC |
| 5% | 0.0533 | 0.0541<br>[0.8160] | 0.6013<br>[0.9630] | 0.6009<br>[0.9880] | 0.0467 | 0.0281<br>[0.8669] | 1.0801<br>[0.8793] | 1.1930<br>[0.9456] |
| 1% | 0.0133 | 0.6146<br>[0.4331] | 0.7976<br>[0.8728] | 1.2318<br>[0.9418] | 0.0267 | 0.9961<br>[0.3183] | 0.4477<br>[0.8408] | 1.4197<br>[0.9222] |

**Note**: '% Violation' is the percentage of days when returns are less than the VaR threshold. UC, IND CC are the GMM duration-base tests for unconditional coverage, independence and conditional coverage, developed by Candelon et al. (2010). The number of orthonormal polynomials is set to 5. *P*-values are in brackets.

**Table 8: Model Selection by MZ Equation**

$$x_t = a + b_1\hat{x}_{t|t-1}^{AS} + b_2\hat{x}_{t|t-1}^{LHAR} + e_t, \quad t = 1, 2, \ldots, 150$$

| Dependent Variable | Const | $\hat{x}_{t\|t-1}^{AS}$ | $\hat{x}_{t\|t-1}^{LHAR}$ |
|---|---|---|---|
| Volatility | -0.2848<br>(0.1831) | 2.1334<br>(0.8320) | -0.2123<br>(0.2896) |
| Log-Volatility | -0.0190<br>(0.1346) | 0.8628<br>(0.1931) | 0.1109<br>(0.1226) |

**Note**: Heteroskedasticity-consistent standard errors are in parentheses.

**Table 9: Model Selection by the MZ equation for IBM data**

$$x_t = a + b_1 \hat{x}^{AS}_{t|t-1} + b_2 \hat{x}^{LHAR}_{t|t-1} + e_t, \quad t = 1, 2, \ldots, 150$$

| Dependent Variable | Const | $\hat{x}^{AS}_{t\|t-1}$ | $\hat{x}^{LHAR}_{t\|t-1}$ |
|---|---|---|---|
| Volatility | 1.1367 | 1.1264 | |
| | (1.3667) | (0.2285) | |
| Volatility | -1.4107 | | 3.8264 |
| | (1.7943) | | (0.7576) |
| Volatility | -1.2534 | 0.8924 | 1.3978 |
| | (1.6501) | (0.2271) | (0.7264) |
| Log-Volatility | 0.1854 | 0.9636 | |
| | (0.1049) | (0.0603) | |
| Log-Volatility | 1.1691 | | 0.8098 |
| | (0.0801) | | (0.0903) |
| Log-Volatility | 0.1856 | 0.9554 | 0.0153 |
| | (0.1050) | (0.0781) | (0.0870) |

**Note**: Heteroskedasticity-consistent standard errors are in parentheses.

Figure 1: News Impact from $Z_t$ to $\xi_t$

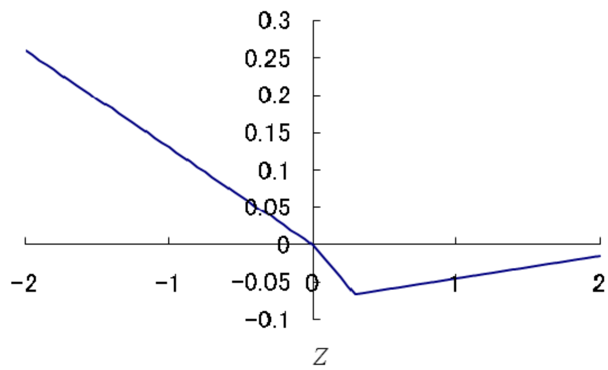


Figure 2: News Impact from $Z_t$ to $\ln V_{t+1}$;
RV-AR(1)-AS($\gamma_1,\gamma_2,\gamma_3$)-Noise Model

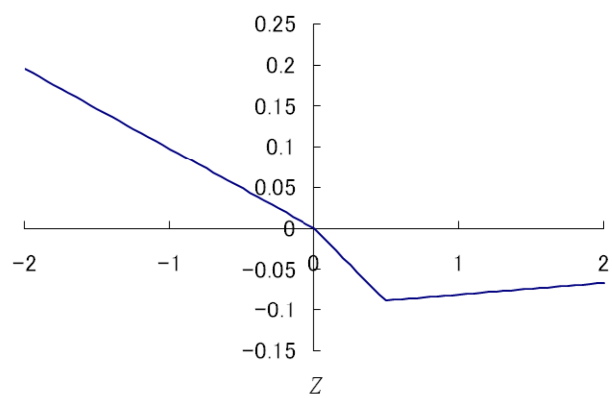


Figure 3: News Impact from $Z_t$ to $\ln V_{t+1}$;
RV-ARFIMA(1,$d$,0)-AS($\gamma_1,\gamma_2,\gamma_3$)-Noise Model