Realized Stochastic Volatility with General Asymmetry and Long Memory[*]

Manabu Asai

Faculty of Economics
Soka University, Japan

Chia-Lin Chang

Department of Applied Economics
Department of Finance
National Chung Hsing University, Taiwan

Michael McAleer

Department of Quantitative Finance
National Tsing Hua University, Taiwan
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam, The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid, Spain
and
Institute of Advanced Sciences
Yokohama National University, Japan

Revised: April 2017

Abstract

The paper develops a novel realized stochastic volatility model of asset returns and realized volatility that incorporates general asymmetry and long memory (hereafter the RSV-GALM model). The contribution of the paper ties in with Robert Basmann's seminal work in terms of the estimation of highly non-linear model specifications ("Causality tests and observationally equivalent representations of econometric models", *Journal of Econometrics*, 1988), especially for specifying causal effects from returns to future volatility. This paper discusses asymptotic results of a Whittle likelihood estimator for the RSV-GALM model and a test for general asymmetry, and analyses the finite sample properties. The paper also develops an approach to obtain volatility estimates and out-of-sample forecasts. Using high frequency data for three US financial assets, the new model is estimated and evaluated. The paper compares the forecasting performance of the new model with a realized conditional volatility model.

1 Introduction

Over the past two decades, realized measures of volatility have received unprecedented attention in the academic literature on modeling and forecasting of stock market returns volatility. In the traditional literature on generalized autoregressive conditional heteroskedasticity (GARCH) models, Engle and Gallo (2006) and Shephard and Sheppard (2010), among others, incorporated realized measures for modeling and forecasting volatility. In addition, Hansen, Huang, and Shek (2012) suggested the realized GARCH framework, which provides a structure for the joint modeling of returns and realized measures of volatility (see also Hansen and Huang (2016)).

Turning to the literature on stochastic volatility (SV), realized measures of volatility are used as realized values of unobservable volatilities. However, even though we can obtain a consistent estimator of true volatility, there are non-negligible differences that are referred to as the 'realized volatility error' (see Barndorff-Nielsen and Shephard (2002)). For removing the estimation bias caused by the realized volatility error in estimating stochastic volatility (SV) models, Barndorff-Nielsen and Shephard (2002), Bollerslev and Zhou (2002), Takahashi, Omori and Watanabe (2009), and Asai, McAleer and Medeiros (2012a,b) showed that it is useful to use an ad hoc approach that accommodates an error term with constant variance. As in the realized GARCH model, Takahashi, Omori and Watanabe (2009) suggested a specification based on daily returns and a realized volatility measure simultaneously, which we will call the 'realized SV' (RSV) model.

As in the GARCH literature, it is important to accommodate asymmetric effects, long memory, and heavy-tails in SV models. For asymmetric effects, it is popular to consider the negative correlation between returns and future volatility in the specification, which is the leverage effect (see Hull and White (1987), Wiggins (1987), Chesney and Scott (1989), Harvey and Shephard (1996), and Yu (2005), among others). As an extension, Asai and McAleer (2011) suggested a general asymmetric model by adding the size effect based on the specification of the exponential

1

GARCH (EGARCH) model of Nelson (1991). Rather than the negative correlation, So, Li, and Lam developed a threshold stochastic volatility model, in the light of the threshold GARCH model of Glosten , Jagannathan, and Runkle (1992). These SV models uses the information given in returns.

For SV models of realized measures of volatility, there is an extensive literature on asymmetric effects that have been investigated by Bollerslev and Zhou (2006), Bollerslev, Litovinova, and Tauchen (2006), Martens, van Dijk, and de Pooter (2009), Chen and Ghysels (2010), Bollerslev, Sizova, and Tauchen (2012), Asai, McAleer and Medeiros (2012a), and Patton and Sheppard (2015), among others. For the specification of asymmetric effects, the empirical results of Chen and Ghysels (2010) led to semi-parametric estimates of the news impact curve (NIC). Their results imply that: (i) a negative shock to future volatility is larger than a positive shock of the same magnitude; (ii) a negative shock and a large positive shock increase future volatility; and (iii) a small positive shock decreases future volatility. Based on the results of Chen and Ghysels (2010), Asai, McAleer and Medeiros (2012a) developed a flexible parametric specification of asymmetric effects for SV models of realized measures of volatility.

With respect to long-range dependence in volatility, Breidt, Crato and de Lima (1998) developed the long memory SV (LMSV) model, in which log-volatility follows the ARFIMA process. Using the information of realized measure, Andersen et al. (2001, 2003), Pong et al. (2004), Koopman, Jungbacker, and Hol (2005), and Asai, McAleer and Medeiros (2012a) consider the autoregressive fractionally integrated model, while other studies have used the heterogeneous autoregressive (HAR) model of Corsi (2009) to approximate the hyperbolic decay rates.

Although SV models are known to be more appropriate for describing the tail thickness of financial returns than ARCH-type models, extreme movements in returns occur more frequently in the observed data than the model implies. One way to cope with this problem in discrete time

is to assume a heavy-tailed distribution, such as Student's t distribution or a generalized error distribution (GED). Following Harvey, Ruiz, and Shephard (1994), Sandmann and Koopman (1998), Liesenfeld and Jung (2000), and Asai (2008, 2009) estimated SV models with the (standardized) t distribution. Liesenfeld and Jung (2000) and Asai (2009) compared both distributions, and found that the t distribution fits generally better than GED.

In the literature of RSV models, Takahashi, Omori and Watanabe (2009) considered the basic SV model without asymmetric effects, long memory, or fat-tails. Koopman and Scharth (2013) accommodated leverage effects and the sum of AR(1) processes to capture longer persistence. Recently, Shirota, Hizu, and Omori (2014) suggested the RSV model with leverage, long memory, and heavy-tails. For estimating their RSV model, Shirota, Hizu, and Omori (2014) used the Bayesian Markov chain Monte Carlo (MCMC) technique. There is room for improving the specification of asymmetric effects, and to develop a frequentest approach for estimation and forecasting volatility.

The purpose of the paper is to extend the work of Shirota, Hizu, and Omori (2014) by accommodating a general asymmetric function for volatility. This paper is closely related to Basmann (1988) for considering causality for future volatility (see also Chang and McAleer (2017)). Applying discussion of Basmann (1988), we develop a reduced form for the RSV model with respect to the general asymmetric function. The reduced form enables use of the Whittle likelihood estimator with asymptotic results, and to conduct a test of general asymmetry. With the first step estimate, we can extract the asymmetric function with an artificial regression. We also extend the approach of Harvey (1998) for estimation and forecasting latent volatility.

The remainder of the paper is organized as follows. Section 2 develops the new RSV-GALM model using a general asymmetric function, and derives a reduced form with respect to the asymmetric effects. Section 3 suggests a two step estimator, consisting of the Whittle likelihood estima-

tion of the reduced form and the artificial regression for the asymmetric function. We show that the first step estimator has the asymptotic normal distribution, and consider a test for general asymmetry which has an asymptotic chi-squared distribution. Section 3 investigates the finite sample performance of the estimator and test statistic, and develop an approach to estimate and forecast volatility by extending the method of Harvey (1998). Section 4 provides an empirical example for three stocks traded on the New York Stock Exchange. Section 5 gives some concluding remarks.

2   Realized Stochastic Volatility Model with Asymmetry and Long Memory

2.1   Model and News Impact Function

Let y_t and x_t be returns and the logarithm of a realized measure of volatility, respectively. Consider the following structure:

$$y_t = \varepsilon_t \exp\left(\frac{1}{2}h_t\right), \quad \varepsilon_t = \frac{z_t}{\sqrt{\kappa_t/(\nu-2)}}, \quad z_t \sim \text{iid}N(0,1) \quad \kappa_t \sim \text{iid}\chi^2(\nu), \tag{1}$$

$$x_t = h_t + e_t, \quad e_t \sim \text{iid}N(0,\sigma_e^2) \tag{2}$$

$$h_{t+1} = \mu + \alpha(L)\xi_t, \quad \alpha(L) = \sum_{j=0}^{\infty} \alpha_j L^j = (1-L)^{-d}[\phi(L)]^{-1}\theta(L), \tag{3}$$

$$\xi_t = \tau(\varepsilon_t) + \eta_t, \quad E[\tau(\varepsilon_t)] = 0, \quad \eta_t \sim \text{iid}N(0,\sigma_\eta^2), \tag{4}$$

where h_t is the unobservable log-volatility, $\tau(\varepsilon)$ is an asymmetric function (to be defined later), L is the lag operator, $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$, and d is the parameter of long memory with $d \in [0, \frac{1}{2})$. Assume that the roots of $\phi(z) = 0$ and $\theta(z) = 0$ are distinct, and they are located outside the unit circle, in order to satisfy the stationary and invertibility conditions, respectively.

By the specification, the unobservable log-volatility, h_t, follows the autoregressive fractionally integrated moving average (ARFIMA) process of order (p, d, q). For observed series, y_t follows

4

the stochastic volatility model with a heavy-tailed distribution, as ε_t follows the standardized t distribution with the degrees-of-freedom parameter, ν ($\nu > 2$). Furthermore, x_t is specified as log-volatility plus a noise process. The asymmetric function, $\tau(\varepsilon)$, will be specified to capture a general effect from a standardized return to one-step-ahead log-volatility. For this paper, the model given in (1)–(4) will be called the 'realized stochastic volatility model with general asymmetry and long memory' (RSV-GALM).

As given in Hansen, Huang and Shek (2012), the asymmetric function is constructed by the Hermite polynomials:

$$\tau(\varepsilon) = \sum_{j=1}^{r} \tau_j a_j(\varepsilon), \quad E[a_j(\varepsilon)] = 0 \text{ for all } j, \tag{5}$$

where

$$a_1(\varepsilon) = \varepsilon, \quad a_2(\varepsilon) = \varepsilon^2 - 1, \quad a_3(\varepsilon) = \varepsilon^3 - 3\varepsilon,$$

$$a_4(\varepsilon) = (\varepsilon^4 - 6\varepsilon^2 + 3) - E(\varepsilon^4 - 6\varepsilon^2 + 3), \ldots$$

The condition $2r < \nu$ is required for the existence of the variance of ξ_t. The RSV-GALM model reduces to the RSV model of Shirota, Hizu, and Omori (2013) by setting $\tau_j = 0$ ($j \geq 2$). We will refer to the latter model as 'RSV-ALM' in order to distinguish it from the general model.

The asymmetric function (5) allows a flexible pattern for the relationship between returns and future volatility. For the case $r = 2$, we obtain $\tau(\varepsilon) = \tau_1\varepsilon + \tau_2(\varepsilon^2 - 1)$, and equation (4) with $\tau_1 < 0$ and $\tau_2 > 0$ enables the RSV-GALM model to accommodate asymmetric effects based on the empirical findings in Chen and Ghysels (2010). Their findings indicate that a negative shock in returns tends to produce a larger future volatility than a positive shock in returns, and that a small positive shock often decreases volatility. Figure 2 illustrates an empirical example of the asymmetric effect from returns, y, to its one-step-ahead volatility, $\bar{\sigma}_{t+1}^2(y)$, using the news impact function (NIF) of log-volatility (see Yu (2005), Asai and McAleer (2009)), where the log of $\bar{\sigma}_{t+1}^2(y)$

is defined by:

$$\log \bar{\sigma}_{t+1}^2(y) = E[h_{t+1}|y_t = y, h_t = \mu] = \mu + E[\xi_t|y_t = y, h_t = \mu] = \mu + \sum_{j=1}^{r} \tau_j a(e^{-\mu/2}y). \quad (6)$$

Figure 2 captures the empirical results of Chen and Ghysels (2010).

2.2 Asymmetric Function and Observationally Equivalent Representation

As noted in Scott (1987) and Harvey, Ruiz, and Shephard (1994), it is useful to take a logarithm of the squared (or absolute) returns. Define $\boldsymbol{v}_t = (x_t^\dagger, y_t^\dagger)'$, where $x_t^\dagger = x_t - \mu$, $y_t^\dagger = \log y_t^2 - \mu_{\log y^2}$, $\mu_{\log y^2} = E(\log y_t^2)$. By the transformation, we obtain $y_t^\dagger = (h_t - \mu) + \zeta_t$, where $\zeta_t = \log \varepsilon_t^2 - E(\log \varepsilon_t^2)$. As in Harvey, Ruiz, and Shephard (1994), it is straightforward to show $E(\log \varepsilon_t^2) = \psi(1/2) - \psi(\nu/2) + \log(\nu - 2)$ and $V(\zeta_t) = \psi'\left(\frac{1}{2}\right) + \psi'\left(\frac{\nu}{2}\right)$, where $\psi(z)$ is the digamma function defined by $\psi(z) = \frac{d \log \Gamma(z)}{dz}$ (see equation (26.3.46) in Abramovits and Stegun (1970)).

The above transformation affects the representation of the asymmetric function. The RSV-GALM model (1)-(5) has the representation:

$$\boldsymbol{v}_t = \left(\begin{array}{c} e_t + \sum_{j=0}^{\infty} \alpha_j \xi_{t-j-1} \\ \zeta_t + \sum_{j=0}^{\infty} \alpha_j \xi_{t-j-1} \end{array} \right) = \sum_{j=0}^{\infty} G_j \boldsymbol{u}_{t-j}, \quad (7)$$

where $\boldsymbol{u}_t = (e_t, \zeta_t, \xi_t)'$, and

$$G_0 = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right), \quad G_j = \left(\begin{array}{ccc} 0 & 0 & \alpha_{j-1} \\ 0 & 0 & \alpha_{j-1} \end{array} \right) \quad (j \geq 1).$$

Moreover, the mean and covariance matrix of $\boldsymbol{u}_t$ are given, respectively, by:

$$E(\boldsymbol{u}_t) = \boldsymbol{0}, \quad V(\boldsymbol{u}_t) = \Sigma_u = \left(\begin{array}{ccc} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & \sigma_{\zeta\xi} \\ 0 & \sigma_{\zeta\xi} & \sigma_\xi^2 \end{array} \right),$$

where

$$\sigma_\zeta^2 = \psi'\left(\frac{1}{2}\right) + \psi'\left(\frac{\nu}{2}\right), \quad \sigma_\xi^2 = \sum_{j=1}^{r} \tau_j^2 V(a_j(\varepsilon_t)) + \sigma_\eta^2, \quad \sigma_{\zeta\xi} = \sum_{i=1}^{\lfloor r/2 \rfloor} \tau_{2i} E[a_{2i}(\varepsilon_t) \log \varepsilon_t^2],$$

and $\lfloor x \rfloor$ is the floor function, which gives the greatest integer less than or equal to x. The values of σ_ξ^2 and $\sigma_{\zeta\xi}$ are finite if $r < \nu/2$ (see Appendix A.1 for the derivation of $E[\varepsilon_t^r \log \varepsilon_t^2]$).

As discussed in Harvey and Shephard (1996), the information regarding the sign of ε_t is lost by the transformation of y_t. Under the assumption that ε_t has the standardized t distribution, $\sigma_{\zeta\xi}$ excludes the effects of the Hermite polynomials of the odd numbers. Unlike the simple asymmetric SV model of Harvey and Shephard (1996), the general asymmetric function (5) with $r \geq 2$ produces non-zero $\sigma_{\zeta\xi}$.

In order to consider the causality from y_t to $\exp(h_{t+1})$, we return to the important concept of the 'observationally equivalent representation' of Basmann (1988) for causality analysis. Basmann (1988) uses a structural vector autoregression (SVAR) model, which has its reduced form derived by multiplying a matrix for normalization. Basmann (1988) shows that a causal relationship will change, depending on the choice of the matrix, which is an example of the problems associated with observationally equivalent representations. For this problem, estimation of the SVAR model requires imposing restrictions, as discussed in Waggoner and Zha (2003).

For the RSV-GALM model, the observationally equivalent representation regarding the asymmetric effect is given by (7). The parameters $(\sigma_e^2, \sigma_\xi^2, \sigma_{\zeta\xi})$ are determined by ν, σ_η^2, and $(\tau_1, \ldots, \tau_r)$. Among them, σ_e^2 is uniquely determined by ν, and vice-versa. On the contrary, it is not possible to identify $(\tau_1, \ldots, \tau_r)$ and σ_η^2 from the information of σ_ξ^2 and $\sigma_{\zeta\xi}$ with ν. For example, if $r = 2$, we obtain:

$$
\sigma_\xi^2 = \tau_1^2 + 2\tau_2^2 \times \frac{\nu - 1}{\nu - 4} + \sigma_\eta^2
$$

$$
\sigma_{\zeta\xi} = \tau_2 \left[\psi\left(\frac{\nu}{2}\right) - \psi\left(\frac{\nu - 2}{2}\right) + \frac{1}{2}\left\{ \psi\left(\frac{3}{2}\right) - \log\left(\frac{1}{2}\right) \right\} - \psi\left(\frac{1}{2}\right) \right].
$$

(8)

Given the structure, ν and τ_2 are identified by $(\sigma_\zeta^2, \sigma_{\zeta\xi})$, but it is not possible to separate $\tau_1^2 + \sigma_\eta^2$ from the information of $(\sigma_e^2, \sigma_\xi^2, \sigma_{\zeta\xi})$. In this sense, we may call the representation (7) as the 'reduced form' of the asymmetric function. In general, $\sigma_{\zeta\xi} = 0$ implies that the asymmetric function is specified as $\tau(\epsilon) = 0$ or $\tau(\epsilon) = \tau_1\varepsilon$. In the next section, we will develop the test for general asymmetry, $\sigma_{\zeta\xi} \neq 0$.

The selection of the order, r, of the asymmetric function (5) requires knowledge of ν, which is unknown *a priori*. Hence, an important feature of the reduced form (7) is that we can handle the model without knowledge of r. Rather than the estimation of a restricted model, the paper considers extracting the asymmetric function after estimating the reduced form (7), and uses the estimated asymmetric function for forecasting volatility. These issues will be discussed in the next section.

3 Estimation and Forecasting

3.1 Whittle Likelihood Estimation

We use the limiting theory of Hosoya (1997) to show the asymptotic property of the Whittle likelihood estimator for the reduced form (7) of the RSV-GALM model. As Hosoya (1997) considers a vector long memory process, which allows differences in the dimensions of the observed series and disturbances, we can directly check the conditions given in Hosoya (1997).

Let $\boldsymbol{\delta} = (d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma_e^2, \sigma_\zeta^2, \sigma_{\zeta\xi}, \sigma_\xi^2)'$ be the vector of parameters. By the specification, the process $\{\boldsymbol{v}_t\}$ in (7) is a second-order stationary process and has a spectral density matrix defined by $\boldsymbol{f}(\omega) = \frac{1}{2\pi}\boldsymbol{k}(\omega;\delta)\Sigma_u\boldsymbol{k}(\omega;\delta)^*$, where $\boldsymbol{k}(\omega;\delta) = \sum_{j=0}^{\infty} G_j e^{i\omega j}$, which yields:

$$\boldsymbol{f}(\omega) = \frac{1}{2\pi} \begin{pmatrix} K_{11}(\omega) & K_{12}(\omega) \\ K_{12}(\omega)^* & K_{22}(\omega) \end{pmatrix}, \tag{9}$$

with

$$K_{11}(\omega) = \sigma_e^2 + \sigma_\xi^2|\alpha(e^{i\omega})|^2,$$

$$K_{12}(\omega) = \sigma_{\zeta\xi}\alpha(e^{i\omega})e^{i\omega} + \sigma_\xi^2|\alpha(e^{i\omega})|^2,$$

$$K_{22}(\omega) = \sigma_\zeta^2 + \sigma_{\zeta\xi}\left(\alpha(e^{i\omega})e^{i\omega} + \alpha(e^{-i\omega})e^{-i\omega}\right) + \sigma_\xi^2|\alpha(e^{-i\omega})|^2.$$

The (1,1)-element of $\boldsymbol{f}(\omega)$ is the spectral density of x_t, which can be interpreted as the conventional signal plus noise process. The (2,2)-element of $\boldsymbol{f}(\omega)$ is the spectral density of $\log y_t^2$, and it is the

same as that derived in Zaffaroni (2009). The off-diagonal elements are the cross-spectral densities of x_t and $\log y_t^2$.

Let $\boldsymbol{I}_n(\boldsymbol{v}, \omega)$ is the periodogram matrix defined by:

$$\boldsymbol{I}_n(\boldsymbol{v}, \omega) = \boldsymbol{w}_n(\omega)\boldsymbol{w}_n(\omega)^*, \quad -\pi < \omega \leq \pi,$$

where $\boldsymbol{w}_n(\omega)$ is the finite Fourier transform defined by:

$$\boldsymbol{w}_n(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{t=1}^{n} \boldsymbol{v}_t e^{it\omega}.$$

For the purpose of deriving the quasi-likelihood function, we will treat the process $\boldsymbol{v}(t)$ as Gaussian. Choose the frequencies ω_j, $j = 1, \ldots, n$, equi-spaced in the region $(-\pi, \pi]$ in such a way that $\boldsymbol{f}(\omega)$ is continuous at $\omega = \omega_j$ Then the finite Fourier transform $\boldsymbol{w}_n(\omega_j)$, $j = 1, \ldots, n$, will have a complex-valued multivariate normal distribution which, for large n, is approximately independent, each with probability density function given by:

$$\pi^{-1} \{\det \boldsymbol{f}(\omega_j; \boldsymbol{\delta})\}^{-1/2} \exp\left[-\frac{1}{2} \text{tr}\left\{\boldsymbol{f}^{-1}(\omega_j; \boldsymbol{\delta})\boldsymbol{w}_n(\omega_j)\boldsymbol{w}_n(\omega_j)^*\right\}\right], \quad j = 1, \ldots, n.$$

As $\boldsymbol{w}_n(\omega_j)$, $j = 1, \ldots, n$, constitutes a sufficient statistic for $\boldsymbol{\delta}$, an approximate log-likelihood function of $\boldsymbol{\delta}$ based on $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is, excluding the constant term, given by:

$$\bar{L}_n(\boldsymbol{\delta}) = -\frac{1}{2} \sum_{j=1}^{n} \left[\log \det \boldsymbol{f}(\omega_j; \boldsymbol{\delta}) + \text{tr}\left\{\boldsymbol{f}^{-1}(\omega_j; \boldsymbol{\delta})\boldsymbol{I}_n(\boldsymbol{v}, \omega_j)\right\}\right]. \tag{10}$$

In integral form, the equation (10) has the expression:

$$-\frac{n}{4\pi}\left[\int_{-\pi}^{\pi} \log \det \boldsymbol{f}(\omega; \boldsymbol{\delta})d\omega + \int_{-\pi}^{\pi} \text{tr}\left\{\boldsymbol{f}^{-1}(\omega; \boldsymbol{\delta})\boldsymbol{I}_n(\boldsymbol{v}, \omega)\right\} d\omega\right]. \tag{11}$$

The function $\bar{L}_n(\boldsymbol{\delta})$ is called the quasi-log-likelihood function. The approximation was originally proposed by Whittle (1952) for scalar-valued stationary processes (see also Dunsmuir and Hannan (1976), Taniguchi and Kakizawa (2000)). Define the QML estimator, $\hat{\boldsymbol{\delta}}_n$, which is obtained by minimizing $-\bar{L}_n(\boldsymbol{\delta})$. In practice, we use the discrete quasi-log-likelihood (10) with frequency

$\omega_j = 2\pi j/n$ $(j = 1, \dots, \lfloor n/2 \rfloor)$, for the symmetry of the Fourier transform, as in the usual empirical analysis.

Following Hosoya (1997), define the quantity:

$$R_j(\boldsymbol{\delta}) = H_j(\boldsymbol{\delta}) + \int_{-\pi}^{\pi} \text{tr}\{\boldsymbol{h}_j(\omega, \boldsymbol{\delta})\boldsymbol{f}(\omega)\}d\omega,$$

where

$$H_j(\boldsymbol{\delta}) = \frac{\partial}{\partial \delta_j} \int_{-\pi}^{\pi} \log \det \boldsymbol{f}(\omega; \boldsymbol{\delta})d\omega,$$

$$\boldsymbol{h}_j(\omega; \boldsymbol{\delta}) = \frac{\partial}{\partial \delta_j} \boldsymbol{f}^{-1}(\omega; \boldsymbol{\delta}).$$

Noting that

$$\det \boldsymbol{f}(\omega; \boldsymbol{\delta}) = \frac{1}{2\pi} \left\{ \sigma_e^2 \sigma_\zeta^2 + 2\sigma_e^2 \sigma_{\zeta\xi} \sum_{j=0}^{\infty} \alpha_j \cos(\omega(j+1)) + (\sigma_e^2 \sigma_\xi^2 + \sigma_\zeta^2 \sigma_\xi^2 - \sigma_{\zeta\xi}^2) |\alpha(e^{i\omega})|^2 \right\},$$

$R_j(\boldsymbol{\delta})$ is measurable with respect to $\boldsymbol{\delta}$ almost everywhere ω. Denote W as the matrix of derivatives, $W_{jl} = \partial R_j/\partial \delta_l$, evaluated at $\boldsymbol{\delta} = \boldsymbol{\delta}_0$.

By checking the conditions of Hosoya (1997), we can obtain the asymptotic results of the QML estimator. Let the QML estimator, $\hat{\boldsymbol{\delta}}_n \in \Theta$, a compact parameter space and the true parameter vector, $\boldsymbol{\delta}_0$, be in the interior of the parameter space, and add a technical assumption that $K_{jl}(\omega) = 0$ for $\omega = 0$ in equation (9). Example 3.1 in Hosoya (1996) uses the assumptions, and the latter assumption has no effect on maximizing the discrete quasi-log-likelihood (10) with frequency $\omega_j = 2\pi j/n$ $(j = 1, \dots, [n/2])$. Applying the approach to prove Theorem 2 in Chan and Tsai (2008) and Theorem 1 in Tsai, Rachinger, and Lin (2015), we can verify Assumptions A, C, and D of Hosoya (1997) to show the consistency and asymptotic normality of the QML estimator, in order to obtain:

$$\sqrt{n}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}_0) \xrightarrow{d} N(0, W^{-1}U(W^*)^{-1}), \tag{12}$$

10

where U is the matrix with (j, l)th element represented as:

$$U_{jl} = 4\pi \int_{-\pi}^{\pi} \text{tr} \left[\boldsymbol{h}_j(\omega; \delta_0) \boldsymbol{f}(\omega) h_l(\omega; \delta_0) \boldsymbol{f}(\omega) \right] d\omega \bigg|_{\delta=\delta_0}$$

$$+ \sum_{a,b,c,d=2}^{3} C_{abcd} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\omega_1) \boldsymbol{h}_j(\omega_1; \delta_0) k(\omega_1) \big|_{\delta=\delta_0} d\omega_1 \right]_{ab} \qquad (13)$$

$$\times \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\omega_2) \boldsymbol{h}_l(\omega_2; \delta_0) k(\omega_2) \big|_{\delta=\delta_0} d\omega_2 \right]_{cd},$$

and

$$\int_{-\pi}^{\pi} k^*(\omega) \boldsymbol{h}_j(\omega; \delta_0) k(\omega) \big|_{\delta=\delta_0} d\omega = 0 \quad \text{for } \theta_j \in (d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q), \qquad (14)$$

with C_{abcd} is the fourth cumulant of the element of $\boldsymbol{u}_t$ obtained by $E(u_a u_b u_c u_d) - E(u_a u_b) E(u_c u_d) -$

$E(u_a u_c) E(u_b u_d) - E(u_a u_d) E(u_b u_c)$. Those cumulants exist if $r < \nu/4$ (see Appendix A.2). As

the first element of $\boldsymbol{u}_t$, namely, e_t, follows the independent normal distribution, $C_{abcd} = 0$ if at

least one of (a, b, c, d) is one.

Although U_{jl} defined by Theorem 2.2 in Hosoya (1997) is based on the fourth-order spectral

density, it can be simplified as in (13) under Assumption F of Hosoya (1997) (see also equation

(5.3.22) of Taniguchi and Kakizawa (2000) and Theorem 2 of Zaffaroni (2009)). As Assumption

F states that the joint fourth cumulant of the elements of u_{a,t_1}, u_{b,t_2}, u_{c,t_3}, u_{d,t_4} is equal to C_{abcd}

if $t_1 = t_2 = t_3 = t_4$, and is equal to zero otherwise, it can be verified by the structure of the

RSV-GALM model. Furthermore, Σ_u in equation (9) is irrelevant to $(d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$,

and so it satisfies the so-called 'innovation-free specification' of Assumption E of Hosoya (1997),

yielding (14) (see also Assumption 5.3.7 of Taniguchi and Kakizawa (2000)).

Before estimation of the general asymmetric function, we consider a test for general asymmetry

of the null hypothesis:

$$H_0 : \sigma_{\zeta\xi} = 0, \quad \text{vs.} \quad H_1 : \sigma_{\zeta\xi} \neq 0.$$

11

The likelihood ratio (LR) statistic is defined by:

$$LR = 2\left\{\bar{L}_n(\hat{\boldsymbol{\delta}}_n) - \bar{L}_n(\hat{\boldsymbol{\delta}}_n^0)\right\},$$

where $\hat{\boldsymbol{\delta}}_n^0$ is the constrained QML estimator. Under mild conditions in Theorem 2.4 of Hosoya (1997), the LR statistic follows the asymptotic χ_1^2 distribution (see also Theorem 5.3.8 of Taniguchi and Kakizawa (2000)).

3.2 Forecasting with the Asymmetric Function

Using the QML estimates above, this paper suggests a method of extracting the asymmetric function (5) and use it to forecast volatility. For this purpose, we consider the minimum mean square linear estimator (MMSLE) of h_t, by extending the work of Harvey (1998) for a long memory SV model. Second, we obtain the estimate of ξ_t to construct an auxiliary regression for (5). Finally, we suggest an approach for forecasting volatility with the estimated asymmetric function.

Define $\boldsymbol{x}^\dagger = (x_1^\dagger, \dots, x_n^\dagger)'$, $\boldsymbol{y}^\dagger = (y_1^\dagger, \dots, y_n^\dagger)'$, $\boldsymbol{h} = (h_1, \dots, h_n)'$, $\boldsymbol{e} = (e_1, \dots, e_n)'$, and $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_n)'$ on (7). Using $\boldsymbol{h}$ explicitly, the vector form of (7) is given by:

$$\boldsymbol{x}^\dagger = \boldsymbol{h} - \mu\mathbf{1}_n + \boldsymbol{e}, \quad \boldsymbol{y}^\dagger = \boldsymbol{h} - \mu\mathbf{1}_n + \boldsymbol{\zeta},$$

where $\mathbf{1}_n$ is an $n \times 1$ vector of ones. Then, we obtain the MMSLE of $\boldsymbol{h}$ as:

$$\tilde{\boldsymbol{h}} = \mu\mathbf{1}_n + b_1^{-1}\left(I_n - \Sigma_B^{-1}\right)\left(\sigma_e^{-2}\boldsymbol{x}^\dagger + \sigma_\zeta^{-2}\boldsymbol{y}^\dagger + b_2 B_2'\Sigma_*^{-1}\boldsymbol{y}_0^\dagger\right), \tag{15}$$

where $\boldsymbol{y}_0^\dagger = (0, y_1^\dagger, \dots, y_{n-1}^\dagger)'$, $b_1 = \sigma_e^{-2} + \sigma_\zeta^{-2}$, $b_2 = \sigma_{\zeta\xi}/\sigma_\zeta^2$, $\Sigma_B = I_n + b_1 B_2^{-1}\Sigma_*(B_2')^{-1}$, and B_2 and Σ_* are defined in Appendix A.3, as is the derivation of the MMSLE. If $\sigma_{\zeta\xi} = 0$, the MMSLE is given by:

$$\tilde{\boldsymbol{h}} = \mu\mathbf{1}_n + b_1^{-1}\left(I_n - \Sigma_b^{-1}\right)\left[\sigma_e^{-2}\boldsymbol{x}^\dagger + \sigma_\zeta^{-2}\boldsymbol{y}^\dagger\right],$$

where $\Sigma_b = I_n + b_1\Sigma_h$ and $V(\boldsymbol{h}) = \Sigma_h$. In this case, the MMLSE is simplified as $\boldsymbol{y}_0^\dagger$ is unnecessary. If $\sigma_{\zeta\xi} = 0$ and $\sigma_e^2 \to \infty$, then $b_1 \to \sigma_\zeta^{-2}$ and the MMSLE reduces to that obtained in Harvey

(1998). As μ and $\mu_{\log y^2}$ are unknown, we need to estimate them using the sample mean. Harvey (1998) recommends using the volatility estimate:

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_{\tilde{y}}^2 \exp(\tilde{h}_t),$$

where $\tilde{\sigma}_{\tilde{y}}^2 = \frac{1}{n} \sum_{t=1}^n \tilde{y}_t^2$, and $\tilde{y}_t = y_t \exp(-\tilde{h}_t/2)$ are the heteroskedasticity-corrected observations.

Before extracting the asymmetric function, consider the $\text{AR}(\infty)$ representation of the latent log-volatility as:

$$h_{t+1} = \mu + \sum_{j=1}^{\infty} \beta_j (h_{t+1-j} - \mu) + \xi_t, \tag{16}$$

where $\sum_{j=1}^{\infty} \beta_j L^j = 1 - [\alpha(L)]^{-1}$. As in Hosking (1981), we can approximate:

$$h_{t+1} = \mu + \sum_{j=1}^{P} \beta_j (h_{t+1-j} - \mu) + \xi_t,$$

with some large P. Denote the pre-sample data with size J as $\boldsymbol{x}_p^{\dagger} = (x_{-J+1}^{\dagger}, \dots, x_0^{\dagger})'$ and $\boldsymbol{y}_p^{\dagger} = (y_{-J+1}^{\dagger}, \dots, y_0^{\dagger})'$, in order to set corresponding pre-sample values $\boldsymbol{h}_p = (h_{-J+1}, \dots, h_0)'$ using equation (15). Then, we obtain the estimates of ξ_t as:

$$\tilde{\xi}_t = (\tilde{h}_{t+1} - \mu) - \sum_{j=1}^{J+t} \beta_j (\tilde{h}_{t+1-j} - \mu) \quad (t = 1, \dots, n-1),$$

in order to construct an auxiliary regression:

$$\tilde{\xi}_t = \sum_{j=1}^{r} \tau_j a_j (\tilde{\sigma}_t^{-1} y_t) + \text{ error.} \tag{17}$$

Note that the maximum value of r must satisfy the condition that $r_{\max} \leq \lfloor \nu/4 \rfloor$ for the fourth cumulants used in the covariance matrix in equation (12). We can estimate the parameters via the conventional ordinary least squares method.

After estimating the asymmetric function, we can obtain l-step-ahead forecast of log-volatility by:

$$\hat{h}_{n+l|n} = \mu + \sum_{j=1}^{n+J+l-1} \beta_j (\hat{h}_{n+l-j|n} - \mu) + \mathbf{1}_{\{l=1\}} \sum_{i=1}^{r} \tau_i a_i (\tilde{\sigma}_n^{-1} y_n), \tag{18}$$

13

where $\hat{h}_{n+l|n} = \tilde{h}_{n+l}$ for $l \leq 0$, and $\mathbf{1}_{\{l=1\}}$ is the indicator function, which takes one when $l = 1$ and zero otherwise.

3.3 Monte Carlo Experiments

Simulation experiments were conducted in order to assess the performance of the two step estimator for the RSV-GALM model, and the quasi-likelihood ratio test of the general asymmetry.

We consider the RSV-GALM $(1, d, 0)$ with $r = 2$, that is, the second-order specification for the asymmetric function (5). For the two step estimator, the parameter values are specified as:

$$(\mu, \sigma_e, \nu, d, \phi, \tau_1, \tau_2, \sigma_\eta) = \begin{cases} (-0.123, 0.25, 11, 0.4, 0.63, -0.05, 0.05, 0.120) & \text{for DGP1} \\ (-0.250, 0.70, 16, 0.2, 0.86, -0.02, 0.02, 0.216) & \text{for DGP2} \end{cases}$$

in the following way. First, the long memory parameter and autoregressive parameter (d, ϕ) were selected such that the first-order autocorrelation function of h_t is 0.95. Second, the parameters $(\nu, \tau_1, \tau_2, \sigma_\eta)$ were chosen so that $\sqrt{V(h_t)} = 0.5$ for DGP1, and 0.7 for DGP2. Third, μ was selected for the variance of y_t to take one, that is, $\exp(\mu + 0.5V(h_t)) = 1$. Finally, σ_e was chosen so that $\sigma_e/\sqrt{V(h_t)} = 0.5$ for DGP1, and 1 for DGP2. For each parameter vector, we generated a sample of size $T = 2048$, and estimated the RSV-GALM model with $r = 2$ using the two step estimator. The number of replications is 2000.

Table 1 shows the sample means, standard deviations, and root mean squared errors of the two step estimator. Table 1(a) presets the results of the QML estimators in the first step. Table 1(a) gives the corresponding true values of $(\rho_{\zeta\xi}, \sigma_\xi)$, where $\rho_{\zeta\xi} = \sigma_{\zeta\xi} / \sqrt{\sigma_\zeta^2 \sigma_\xi^2}$, via equation (8). Compared with the bias of (d, ϕ, σ_e), those of $(\nu, \rho_{\zeta\xi})$ are relatively large, due to the fourth cumulants in (13) caused by the non-normal disturbance. Those bias are expected to disappear, as the sample size increases.

Table 1(b) shows the results of the estimates of the remaining parameters. While μ was

14

estimated by the sample mean of x_t, $(\tau_1, \tau_2, \sigma_\eta)$ were estimated by the OLS method for the auxiliary regression (17). The bias in μ is negligible. There is bias in the estimates of τ_1 and τ_2 toward zero, while the bias in σ_η is affected by those of $|\tau_1|$ and $|\tau_2|$. These biases are expected to disappear as the values of $|\tau_1|$ and $|\tau_2|$ increase, and/or as $n \to \infty$.

Next, we examine the finite sample performance of the quasi-likelihood ratio test for general asymmetry with null hypothesis $\sigma_{\zeta\xi} = 0$. For this purpose, we modify the parameter settings of (τ_2, σ_η) in DGP1 and DGP2, in order to specify $\rho_{\zeta\xi} = 0$, 0.1, 0.3, keeping the remaining conditions the same. Table 2 gives the new parameters, and corresponding rejection frequencies at the 5 percent significance level. Table 2 indicates that the rejection frequency increases as the value of $\rho_{\zeta\xi}$ increases. The rejection frequencies under the alternative hypothesis are higher as $\sigma_e/\sqrt{V(h_t)}$ is smaller. On the other hand, the rejection frequency under the null hypothesis is closer to 0.05 when $\sigma_e/\sqrt{V(h_t)}$ is larger.

4 Empirical Analysis

4.1 Data and Preliminary Analysis

We estimate the RMESV-ALM model using daily returns and realized volatilities for three major stocks traded on the New York Stock Exchange, namely: Alcoa Inc. (AA), Bank of America (BAC), and International Business Machines (IBM). For each return computed for 1-min intervals of the trading day at t between 9:30a.m. and 4:00 p.m., we calculated the daily volatility using the preaveraged and truncated realized volatility (PTRV) estimator of Koike (2016) (see also Asai and McAleer (2017)). The estimator gives a consistent estimator of the integrated covariance matrix, and is robust to microstructure noise and jumps.

We also calculate the corresponding returns for the three assets. Following the methodology of Aït-Sahalia and Jacod (2012), we obtained the 'robust' returns by removing the jump and noise components. Although the realized kernel estimator of Barndorff-Nielsen et al. (2008) is consistent

and robust to microstructure noise and jumps, we use the PTRV estimator for disentangling these components on volatility and returns simultaneously. We denote robust returns and log of PTRV as y_t and x_t, respectively. The sample period is from July 2, 1998 to October 26, 2012, giving 3548 observations. Figure 1 shows the time series plot of the PTRV estimator. We divide the sample into three parts, with sizes 1000, 2048, and 500. We use the second $n = 2048$ observations for estimating the parameters in the first step, while the first $J = 1000$ observations are used as a pre-sample for estimating $\{h_t\}$ and the parameters in the second step. We preserve the last $F = 500$ observations for forecasting. The second period is from July 1, 2002 to October 13, 2010, covering the period of the global financial crisis, as shown in Figure 1.

Table 3 presents the descriptive statistics of the returns and log-volatilities. The empirical distribution of the returns is highly leptokurtic, and is skewed to the left. Compared with the returns series, the distribution of log-volatility is closer to the normal distribution, but it is skewed to the right, and the kurtosis exceeds three. It should be noted that it is straightforward to verify that ξ_t in the RSV-GALM model with $r \geq 2$ can describe the heavy-tails and skewness to the right.

For a benchmark model, we consider the realized EGARCH model of Hansen and Huang (2016). By accommodating heavy-tails, the model is given by equation (1) and:

$$x_t = \varphi_1 + \varphi_2 h_t + \lambda_1 \varepsilon_t + \lambda_2(\varepsilon_t^2 - 1) + e_t, \quad e_t \sim \text{iid} N(0, \sigma_e^2),$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \tau_1 \varepsilon_t + \tau_2(\varepsilon_t^2 - 1) + \tau_0 e_t$$

Although the disturbance in h_{t+1} equation is predetermined, unlike the RSV models, the disturbance of x_t is expected to cover the difference. The structure of the realized EGARCH model implies that there is an instantaneous asymmetric relationship between y_t and x_t, which is not considered in the RSV models. On the other hand, the realized EGARCH model misses the long memory in this specification. As examined in Hansen and Huang (2016), the specification gives

a better fit than the realized GARCH model of Hansen, Huang, and Shek (2012). The density function of the standardized t distribution is given by:

$$f(\varepsilon) = [\pi(\nu - 2)]^{-1/2} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{\varepsilon^2}{\nu - 2}\right)^{-(\nu+1)/2}.$$

As in the conventional ARCH family, we can estimate the model via maximum likelihood (ML) estimation.

It should be noted that the realized EGARCH model uses the level and squares of the standardized shock, ε_t, whereas the standard EGARCH model of Nelson (1991) uses the levels and absolute values of the standardized shocks.

Table 4 shows the ML estimates of the realized EGARCH model. All the estimated parameters are significant at the five percent level. The estimates of ν indicate that the standardized t distribution is better than the normal distribution. For the specification of h_{t+1}, we observe that ϕ is close to 0.98 as in the conventional EGARCH specification. While the estimates of τ_1 are negative, those of τ_2 are positive. As in the results of Hansen and Huang (2016), τ_0 is positive. For the model of x_t, φ_2 is close to one and φ_1 is close to minus one. Although $(\varphi_1, \varphi_2) = (0, 1)$ is expected by theory, there are significant differences. The estimates of σ_e lie in the range (0.37,0.43). For the simultaneous asymmetric effects between y_t and x_t, the estimates of λ_1 are negative, and those of λ_2 are positive. These results support the empirical analysis of Hansen and Huang (2016).

4.2　Empirical Results

We estimate the RSV-GALM model, and compare its one-step-ahead forecasts with those of the realized EGARCH model.

Table 5(a) presents the QML estimates for the first step. For estimating the asymptotic covariance matrix of the QML estimator, we use the simulated fourth cumulants given in Table 6,

which will be explained below. The estimates of d lie in the range (0.47,0.49), and are significant at the five percent level, showing that long memory in volatility improves the in-sample performance. For the general asymmetry, we test the null hypothesis, $\rho_{\zeta\xi} = 0$. Both the t test and QLR test reject the null hypothesis for AA and IBM, while the null hypothesis is not rejected for BAC. Compared with the estimates of the realized EGARCH model, the estimates of σ_e are smaller, implying that the systematic variation of x_t explained by the RSV-GALM model is larger than for the realized EGARCH model.

Table 5(b) shows the results of the second step estimator. In the second step, we use the maximum number of r as $r_{\max} = \min(5, \lfloor 0.25\nu \rfloor)$, and select the optimal order of the general asymmetric function by choosing the maximum number of j, of which τ_j is significant at the five percent level. While $r_{\max}$ is 3 for AA, and 5 for BAC and IBM, the optimal r is 2 for AA and IBM, and 1 for BAC. The result for BAC corresponds to the above QLR test. Figure 2 shows the news impact function for these three datasets.

Using the estimates in Table 5(b), we generated a random sample of (ε_t, η_t) with size of 1000000, in order to obtain the fourth cumulants of (ζ_t, ξ_t), and estimated $W^{-1}U(W^*)^{-1}$ in equation (12) via numerical derivation and integration. The standard errors are given in the parentheses in Table 5(a).

For the RSV-GALM ($r \geq 2$) and RSV-ALM ($r = 1$) models, we examine the performance of the out-of-sample forecasts using the root mean squared forecast error (RMSFE) and the Diebold and Mariano (1995) test for equal forecast accuracy. The benchmark model is the realized EGARCH model. We use x_{n+l} ($l = 1, \ldots, F$) as the proxy of the true log-volatility. Fixing the sample size at 2048 as the rolling window, we re-estimated the model and computed one step ahead forecasts of log-volatility for the last 500 days. RMSFE is defined as:

$$\sqrt{\frac{1}{F} \sum_{l=1}^{F} \left(\hat{x}_{n+l|n+l-1} - x_{n+l} \right)^2},$$

18

where $\hat{x}_{n+l|n+l-1} = \hat{h}_{n+l|n+l-1}$ for the RSV model, and $\hat{x}_{n+l|n+l-1} = \hat{\varphi}_1 + \varphi_2 \hat{h}_{n+l|n+l-1}$ for the realized EGARCH model As above, we select the optimal r for each time to obtain the one step ahead forecast.

Table 7 shows that the RSV-GALM model produces the smallest RMSFE for all three datasets. The results of the Diebold and Mariano (1995) test indicate that the forecasts of the RSV-GALM model are significantly different from those of the remaining two models.

For each time predicting x_{n+l}, we select optimal r. Table 8 presents the frequencies of the selected r in forecasting the RSV-GALM model. The frequencies for selecting $r = 2$ is highest for all three datasets, while the frequencies where a higher r is selected are not negligible. For all three assets, $r = 0$ is not selected. Figure 3 shows the changes in the optimal r. The first half of the trajectory for AA is relatively calm and $r = 2$ is sufficient, while the latter period fluctuates from one to four. For BAC, there is a tendency that $r = 1$ or 2 are selected at the beginning, but with fluctuations for most of the forecasting period. There is a calm period in the middle for IBM.

Table 8 also gives the correlation coefficients between $\exp(0.5x_t)$ and r, showing a weak negative correlation. The result implies that there is a weak tendency to select small r when volatility is large. A more extensive analysis for choosing an optimal value of r requires further research.

5 Conclusion

In the framework of realized stochastic volatility, this paper developed general asymmetry, long memory, and fat-tails, in order to construct the RSV-GALM model. The general symmetric function allows the model to describe flexible asymmetric patterns, corresponding to recent empirical findings. This paper derives the reduced form of the RSV-GALM model, and suggested Whittle likelihood estimation of the model, and developed a test for general asymmetry.

This paper also gives the asymptotic results of the estimator. As the second step, this paper developed an artificial regression for estimating the the general asymmetric function. The paper gives an approach to obtain estimates and forecasts of the RSV-GALM model. The empirical results based on the three financial assets from the US market indicate that the new RSV-GALM model outperforms existing models in forecasting volatility.

The new RSV-GALM model opens the possibility for many interesting research directions. First, we may include a simultaneous asymmetric effect between returns and a realized measure of volatility, as in Hansen and Huang (2016). Second, we may consider multiple realized measures of volatility for improving forecasts, as in Engle and Gallo (2006). Third, we may develop a time-varying structure of the asymmetric function, based on the empirical results. These issues require considerable further research.

References

Abramovits, M. and N. Stegun (1970), *Handbook of Mathematical Functions*, Dover Publications, N.Y.

Aït-Sahalia, Y., and J. Jacod (2012), "Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data", *Journal of Economic Literature*, **50**, 1007–1050.

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.

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

Asai, M. (2008), "Autoregressive Stochastic Volatility Models with Heavy-Tailed Distributions: A Comparison with Multifactor Volatility Models", *Journal of Empirical Finance*, **15**, 332–341.

Asai, M. (2009), "Bayesian Analysis of Stochastic Volatility Models with Mixture-of-Normal Distributions", *Mathematics and Computers in Simulation*, **79**, 2579–2596.

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

Asai, M. and M. McAleer (2011), "Alternative Asymmetric Stochastic Volatility Models", *Econometric Reviews*, **30**, 548–564.

Asai, M. and M. McAleer (2017), "The Impact of Jumps and Leverage in Forecasting Co-Volatility", to appear in *Econometric Reviews*.

Asai, M., M. McAleer and M. Medeiros (2012a), "Estimation and Forecasting with Noisy Realized Volatility", *Computational Statistics & Data Analysis*, **56**, 217–230.

Asai, M., M. McAleer and M. Medeiros (2012b), "Asymmetry and Long Memory in Volatility Modeling", *Journal of Financial Econometrics*, **10**, 495–512.

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.

Basmann, R. (1988), "Causality Tests and Observationally Equivalent Representations of Econometric Models", *Journal of Econometrics*, **39**, 69–104.

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

Bollerslev, T., N. Sizova, and G. Tauchen (2012), "Volatility in Equilibrium: Asymmetries and Dynamic Dependencies", *Review of Finance*, **16**, 31–80.

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.

Breidt, F.J., N. Crato and P. de Lima (1998), "The Detection and Estimation of Long Memory", *Journal of Econometrics*, **83**, 325–348.

Chang, C.-L. and M. McAleer (2017), "A Simple Test for Causality in Volatility", *Econometrics*, **5**:15, 1-5.

Chen, X. and E. Ghysels (2010), "News - Good or Bad - and Its Impact on Volatility Predictions over Multiple Horizons", *Review of Financial Studies*, **24**, 46–81.

Chesney, M., and L.O. Scott (1989), "Pricing European Currency Options: A Comparison of The Modified Black-Scholes Model and A Random Variance model", *Journal of Financial and Quantitative Analysis*, **24**, 267–84.

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

Diebold, F. and R. Mariano (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, **13**, 253–263.

Dunsmuir, W. (1979), "A Central Limit Theorem for Parameter Estimation in Stationary Vector Time Series and Its Application to Models for A Signal Observed with Noise", *Annals of Statistics*, **7**, 490–506.

Dunsmuir, W. and E.J. Hannan (1976), "Vector Linear Time Series Models", *Advances in Applied Probability*, **8**, 339–364.

Engle, R. F., and G. Gallo (2006), "A Multiple Indicators Model for Volatility Using Intra-Daily Data", *Journal of Econometrics*, **131**, 3–27.

Gradshteyn, I.S. and I.M. Ryzhik (1980), *Table of Integrals, Series, and Products*, Academic Press, San Diego.

Glosten, L., R. Jagannathan, and D. Runkle (1992), "On the Relation between the Expected Value and Volatility of Nominal Excess Returns on Stocks", *Journal of Finance*, **46**, 1779–1801.

Hansen, P.R. and Z. Huang (2016), "Exponential GARCH Modeling with Realized Measures of Volatility", *Journal of Business & Economic Statistics*, **34**, 269–287.

Hansen, P.R., Z. Huang, and H.H. Shek (2012), "Realized GARCH: A Complete Model of Returns and Realized Measures of Volatility", *Journal of Applied Econometrics*, **27**, 877–906.

Harvey, A. (1998), "Long Memory in Stochastic Volatility", In: Knight, J. and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, Oxford: Butterworth-Haineman, 307–320.

Harvey, A.C., E. Ruiz, and N. Shephard (1994), "Multivariate Stochastic Variance Models", *Review of Economic Studies*, **61**, 247–264.

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.

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

Hosoya, Y. (1997), "A Limit Theory for Long-Range Dependence and Statistical Inference on Related Models", *Annals of Statistics*, **25**, 105–137.

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

Koike, Y. (2016), "Estimation of Integrated Covariances in the Simultaneous Presence of Non-synchronicity, Microstructure Noise and Jumps", *Econometric Theory*, **32**, 533-611.

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.

Koopman, S.J. and M. Scharth (2013), "The Analysis of Stochastic Volatility in the Presence of Daily Realized Measures", *Journal of Financial Econometrics*, **11**, 76–115.

Liesenfeld, R. and R.C. Jung (2000), "Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions", *Journal of Applied Econometrics*, **15**, 137–160.

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.

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

Patton, A. and K. Sheppard (2015), "Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility", *Review of Economics and Statistics*, **97**, 683–697

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.

Sandmann, G., and S. J. Koopman (1998), "Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood", *Journal of Econometrics*, **87**, 271–301.

Scott, L.O. (1987), "Option Pricing When the Variance Changes Randomly: Theory, Estimation and an Application", *Journal of Financial and Quantitative Analysis*, **22**, 419–438.

Shephard, N, and K. Sheppard (2010), "Realising the Future: Forecasting with High Frequency Based Volatility (HEAVY) Models", *Journal of Applied Econometrics*, **25**, 197–231.

Shirota, S., T. Hizu, and Y. Omori (2014), "Realized Stochastic volatility with Leverage and Long Memory", *Computational Statistics & Data Analysis*, **76**, 618–641.

Takahashi, M., Y. Omori, and T. Watanabe (2009), "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously", *Computational Statistics & Data Analysis*, **53**, 2404–2426.

Taniguchi, M. (1982), "On Estimation of the Integrals of the Fourth Order Cumulant Spectral Density", *Biometrika*, **69**, 117–122.

Taniguchi, M. and Y. Kakizawa (2000), *Asymptotic Theory of Statistical Inference for Time Series*, New York: Springer-Verlag.

Waggoner, D.F. and T. Zha (2003), "Likelihood Preserving Normalization in Multiple Equation Models", *Journal of Econometrics*, **114**, 329–347.

Wiggins, J.B. (1987), "Option Values under Stochastic Volatility: Theory and Empirical Estimates", *Journal of Financial Economics*, **19**, 351–72.

Whittle, P. (1952), "Some Results in Time Series Analysis", *Skandivanisk Aktuarietidskrift*, **35**, 48–60.

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

Zaffaroni, P. (2009), "Whittle Estimation of EGARCH and Other Exponential Volatility Models", *Journal of Econometrics*, **151**, 190–200.

Appendix

A.1 Derivation of $E(\varepsilon^r \log \varepsilon^2)$

In the Appendix, we omit the subscript t for random variables, when it is redundant. Noting that $\varepsilon = z\sqrt{(\nu - 2)/\kappa}$, we obtain

$$E(\log \varepsilon^2) = E[\log z^2] - E[\log(\kappa/\nu)] + \log[(\nu - 2)/\nu] = \psi(1/2) - \psi(\nu/2) + \log(\nu - 2)$$

by equation (26.4.36) in Abramovits and Stegun (1970).

If r is an odd number, it is straightforward to show that $E(\varepsilon^r \log \varepsilon^2) = 0$. For an even number $r = 2R$, we obtain:

$$E(\varepsilon^r \log \varepsilon^2) = (\nu - 2)^R E[\kappa^{-R}] E[z^{2R} \log z^2] - (\nu - 2)^R E(z^{2R}) E[\kappa^{-R} \log \kappa]$$
$$+ \left(\frac{\nu - 2}{\nu}\right)^R \log(\nu - 2) \times E[T^{2R}],$$

where $T = z/\sqrt{\kappa/\nu}$ follows the Student t distribution with degrees-of-freedom parameter, ν. It is straightforward to show that:

$$E(z^{2R}) = \prod_{j=1}^{R}(2j - 1), \quad E(T^{2R}) = \nu^R \prod_{j=1}^{R}\frac{2j - 1}{\nu - 2j}, \quad E(\kappa^{-R}) = \prod_{j=1}^{R}\frac{1}{\nu - 2j}.$$

By using equation (26.4.36) in Abramovits and Stegun (1970), we obtain:

$$E[\kappa^{-R} \log \kappa] = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^\infty \log(\kappa)\kappa^{(\nu-2R)/2-1}e^{-\kappa/2}d\kappa$$
$$= \frac{\Gamma((\nu - 2R)/2)}{2^R\Gamma(\nu/2)}E[\log \chi^2_{\nu-2R}] = \frac{\Gamma((\nu - r)/2)}{2^{r/2}\Gamma(\nu/2)}\left\{\psi\left(\frac{\nu - r}{2}\right) - \log\left(\frac{1}{2}\right)\right\}$$
$$E[z^{2R} \log z^2] = \frac{2^{(2R-1)/2}\Gamma((2R + 1)/2)}{\Gamma(1/2)}E[\log \chi^2_{2R+1}]$$
$$= \frac{2^{(r-1)/2}\Gamma((r + 1)/2)}{\sqrt{\pi}}\left\{\psi\left(\frac{r + 1}{2}\right) - \log\left(\frac{1}{2}\right)\right\}.$$

As a result, if r is even, it follows that:

$$E(\varepsilon^r \log \varepsilon^2) = \left(\prod_{j=1}^{r/2} \frac{\nu - 2}{\nu - 2j} \right) \times \frac{2^{(r-1)/2} \Gamma((r+1)/2)}{\sqrt{\pi}} \left\{ \psi \left(\frac{r+1}{2} \right) - \log \left(\frac{1}{2} \right) \right\}$$
$$- \left[\prod_{j=1}^{r/2} (2j-1) \right] \times \left(\frac{\nu - 2}{2} \right)^{r/2} \frac{\Gamma((\nu - r)/2)}{\Gamma(\nu/2)} \left\{ \psi \left(\frac{\nu - r}{2} \right) - \log \left(\frac{1}{2} \right) \right\}$$
$$+ \left(\prod_{j=1}^{r/2} \frac{\nu(2j-1)}{\nu - 2j} \right) \times \left(\frac{\nu - 2}{\nu} \right)^{r/2} \log(\nu - 2).$$

It is straightforward to show that $E(\varepsilon^r \log \varepsilon^2) \to E(z^r \log z^2)$ as $\nu \to \infty$.

For the existence of σ_ξ^2 and $\sigma_{\zeta\xi}$ in equation (7), it is sufficient to guarantee $E(\varepsilon^{2r}) < \infty$ and $|E(\varepsilon^r \log \varepsilon^2)| < \infty$ by $r < \nu/2$.

A.2 Fourth Moments of u_t

Appendix A.2 examines the conditions for the existence of fourth moments of u_t. As discussed in the main text, the normal random variables, e_t and η_t, have all finite moments, and so it is sufficient to check the conditions for $E(\zeta^4)$, $E(\zeta^3 \xi)$, $E(\zeta^2 \xi^2)$, $E(\zeta \xi^3)$, and $E(\xi^4)$.

For the fourth moment of ζ, we decompose ζ as $\zeta = z^* - \kappa^*$, where $z^* = \log z^2 - E(\log z^2)$ and $\kappa^* = \log \kappa - E(\log \kappa)$. By definition, it is straightforward to verify $E(\zeta^4) = E(z^{*4}) + 6E(z^{*2})E(\kappa^{*2}) + E(\kappa^{*4})$. From equation (26.4.36) in Abramovits and Stegun (1970), which gives the kurtosis of z^* and κ^*, we obtain:

$$E(\zeta^4) = \psi^{(3)} \left(\frac{1}{2} \right) + \psi^{(3)} \left(\frac{\nu}{2} \right) + 3 \left[\psi' \left(\frac{1}{2} \right) + \psi' \left(\frac{\nu}{2} \right) \right]^2,$$

where $\psi^{(3)}(z) = d^3 \psi(z)/dz^3$ is the penta gamma function. As $\nu > 2$ by equation (1), $E(\zeta^4)$ is finite.

If $\nu > 4r$, we can show that $E(\varepsilon^{4r}) < \infty$, by the property of the t distribution, which guarantees $E[\{\tau(\varepsilon)\}^4] < \infty$ and $E(\xi^4)$.

Noting that $E(\zeta \xi^3) = E(\xi^3 \log \varepsilon^2) - E(\xi^3)E(\log \varepsilon^2)$, we can show $E(\xi^3 \log \varepsilon^2)$, $E(\log \varepsilon^2)$ and

25

$E(\xi^3)$ are finite using the results in Appendix A.1 and the property of the t distribution, if $3r < \nu$.

Similarly, we can show that $E(\zeta^2\xi^2)$ and $E(\xi\zeta^3)$ are finite, if $4r < \nu$.

Overall, the conditions for the existence of cumulants of $\boldsymbol{u}_t$ is $4r < \nu$. Given the parameter values, we can estimate them by the sample fourth moments, using a simulated random variables of $\boldsymbol{u}_t$.

A.3 Derivation of MMSLE of $\boldsymbol{h}$

As in Harvey (1998), we assume the Gaussian distribution for $\boldsymbol{u}_t$ to derive the MMSLE of $\boldsymbol{h}$ intuitively. Under the Gaussian assumption, the conditional distribution of $\xi_t|\{\zeta_t = y_t^\dagger - h_t^\dagger\}$ is given by $N(b_2(y_t^\dagger - h_t^\dagger), b_3)$, where b_2 is defined by (15) and $b_3 = \sigma_\xi^2 - \sigma_{\zeta\xi}^2/\sigma_\zeta^2$.

Consider the decomposition of the autoregressive terms in (16) as:

$$h_t^\dagger = \sum_{j=1}^{t-1} \beta_j h_{t-j}^\dagger + \xi_{t-1} + s_t,$$

where

$$s_t = \sum_{j=t}^\infty \beta_j h_{t-j}^\dagger = \sum_{j=0}^\infty \beta_{t+j} h_{-j}^\dagger = \sum_{j=0}^\infty \left(\sum_{i=0}^j \beta_{t+i}\alpha_{j-i} \right) \xi_{-2-j}.$$

By definition, the autocovariance function of s_t is given by:

$$E(s_t s_{t'}) = \sigma_\xi^2 \sum_{j=0}^\infty \left(\sum_{i=0}^j \beta_{t+i}\alpha_{j-i} \right) \left(\sum_{i=0}^j \beta_{t'+i}\alpha_{j-i} \right),$$

which depends on (t, t'). In vector form, given $\boldsymbol{y}_0^\dagger$, we obtain:

$$\boldsymbol{h}^\dagger = L_\beta \boldsymbol{h}_0^\dagger + b_2(\boldsymbol{y}_0^\dagger - \boldsymbol{h}_0^\dagger) + \boldsymbol{\xi}^*, \quad \boldsymbol{\xi}^* \sim N(\boldsymbol{0}, \Sigma_*), \quad \Sigma_* = b_3 I_n + B_3 + \Sigma_s,$$

where Σ_s is the covariance matrix of $\boldsymbol{s} = (s_1, \dots, s_n)'$, $\boldsymbol{h}^\dagger = \boldsymbol{h} - \mu\boldsymbol{1}_n$, $\boldsymbol{h}_0^\dagger = S\boldsymbol{h}^\dagger$, and

$$
S = \left(\begin{array}{cc|c} 0 & \cdots & 0 & 0 \\ \hline & & 0 \\ & I_{n-1} & \vdots \\ & & 0 \end{array} \right), \quad
L_\beta = \left(\begin{array}{cccc} \beta_1 & 0 & \cdots & 0 \\ \beta_2 & \beta_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \beta_n & \cdots & \beta_2 & \beta_1 \end{array} \right), \quad
B_3 = \left(\begin{array}{cccc} \sigma_\xi^2 - b_3 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{array} \right).
$$

We can use the above equations to obtain the conditional distribution of $\boldsymbol{h}|\boldsymbol{y}_0^\dagger$.

The density function of $\boldsymbol{h}^\dagger$ conditional on $(\boldsymbol{x}^\dagger, \boldsymbol{y}^\dagger)$ is proportional to:

$$f(\boldsymbol{h}^\dagger|\boldsymbol{x}^\dagger, \boldsymbol{y}^\dagger) \propto f(\boldsymbol{x}^\dagger|\boldsymbol{h}^\dagger)f(\boldsymbol{y}^\dagger, \boldsymbol{h}^\dagger)$$

$$\propto f(\boldsymbol{x}^\dagger|\boldsymbol{h}^\dagger)f(y_1^\dagger|h_1^\dagger)f(h_1^\dagger)\prod_{t=2}^n f(y_t^\dagger|h_t^\dagger)f(h_t^\dagger|y_{t-1}^\dagger, h_{t-1}^\dagger, \dots, h_1^\dagger)$$

$$\propto f(\boldsymbol{x}^\dagger|\boldsymbol{h}^\dagger)f(\boldsymbol{y}^\dagger|\boldsymbol{h}^\dagger)f(\boldsymbol{h}^\dagger|\boldsymbol{y}_0^\dagger),$$

where

$$f(\boldsymbol{x}^\dagger|\boldsymbol{h}^\dagger) \propto \exp\left(-\frac{1}{2\sigma_e^2}(\boldsymbol{x}^\dagger - \boldsymbol{h}^\dagger)'(\boldsymbol{x}^\dagger - \boldsymbol{h}^\dagger)\right)$$

$$f(\boldsymbol{y}^\dagger|\boldsymbol{h}^\dagger) \propto \exp\left(-\frac{1}{2\sigma_\zeta^2}(\boldsymbol{y}^\dagger - \boldsymbol{h}^\dagger)'(\boldsymbol{y}^\dagger - \boldsymbol{h}^\dagger)\right)$$

$$f(\boldsymbol{h}^\dagger|\boldsymbol{y}_0^\dagger) \propto \exp\left(-\frac{1}{2}\left(B_2\boldsymbol{h}^\dagger - b_2\boldsymbol{y}_0^\dagger\right)'\Sigma_*^{-1}\left(B_2\boldsymbol{h}^\dagger - b_2\boldsymbol{y}_0^\dagger\right)\right),$$

with $B_1 = L_\beta - b_2 I_m$, $B_2 = I_n - B_1 S$. The mean of $\boldsymbol{h}^\dagger$, conditional on $(\boldsymbol{x}^\dagger, \boldsymbol{y}^\dagger)$, is given by:

$$\boldsymbol{\mu}_{h|\cdot} = \left(b_1 I_n + B_2'\Sigma_*^{-1}B_2\right)^{-1}\left(\sigma_e^{-2}\boldsymbol{x}^\dagger + \sigma_\zeta^{-2}\boldsymbol{y}^\dagger + b_2 B_2'\Sigma_*^{-1}\boldsymbol{y}_0^\dagger\right)$$

$$= b_1^{-1}\left(I_n - \Sigma_B^{-1}\right)\left(\sigma_e^{-2}\boldsymbol{x}^\dagger + \sigma_\zeta^{-2}\boldsymbol{y}^\dagger + b_2 B_2'\Sigma_*^{-1}\boldsymbol{y}_0^\dagger\right),$$

where $\Sigma_B = I_n + b_1 B_2^{-1}\Sigma_*(B_2')^{-1}$. Then the MMLSE of $\boldsymbol{h}$ is given by $\tilde{\boldsymbol{h}} = \mu\boldsymbol{1}_n + \boldsymbol{\mu}_{h|\cdot}$.

If $\sigma_{\zeta\xi} = 0$, we obtain:

$$f(\boldsymbol{h}^\dagger) \propto \exp\left(-\frac{1}{2}(\boldsymbol{h}^\dagger)'\Sigma_h^{-1}\boldsymbol{h}^\dagger\right),$$

and the MMSLE of $\boldsymbol{h}$ is given by:

$$\tilde{\boldsymbol{h}} = \mu\boldsymbol{1}_n + b_1^{-1}(I_n - \Sigma_b^{-1})\left(\sigma_e^{-2}\boldsymbol{x}^\dagger + \sigma_\zeta^{-2}\boldsymbol{y}^\dagger\right),$$

where $\Sigma_b = I_n + b_1\Sigma_h$.

Figure 1: Daily Realized Measure of Volatility
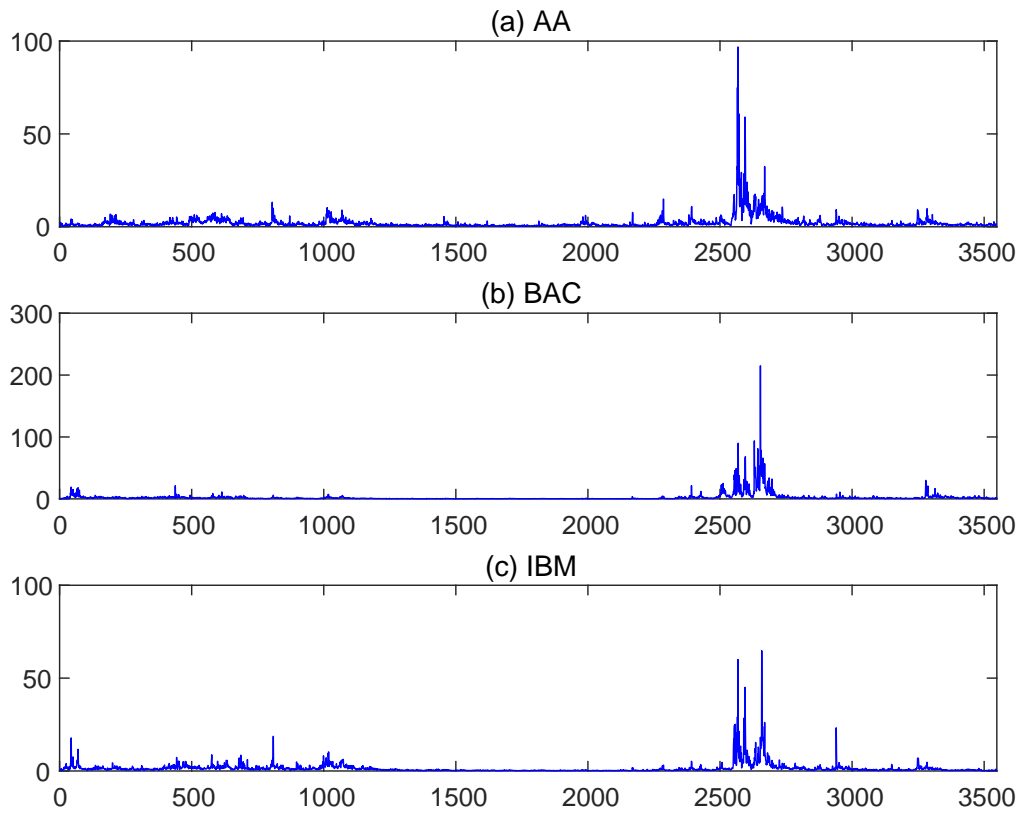
Note: We use the preaveraged and truncated realized volatility estimator of Koike (2016).

Figure 2: News Impact Functions of the RSV-GALM Model

(a) AA

(b) BAC

(c) IBM
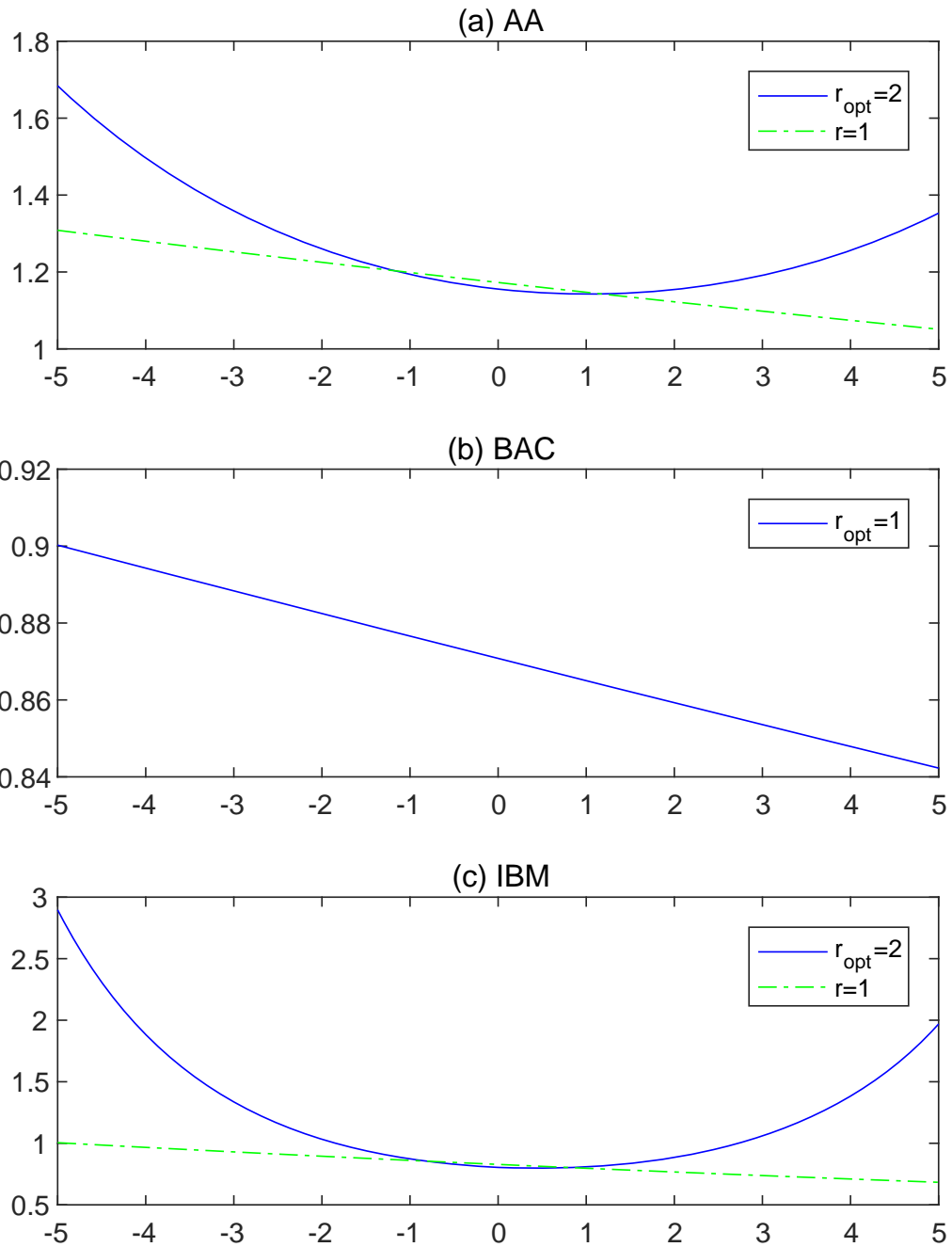
Note: The figure shows the news impact function, $\bar{\sigma}_{t+1}(y)$, defined by equation (6) for the optimal general asymmetric function.

Figure 3: Changes in Optimal r for the General Asymmetric Function

Table 1: Finite Sample Performance of Estimators in Each Step

(a) First Step QML Estimator

Param.	DGP 1				DGP 2			
	True	Mean	Std.Dev.	RMSE	True	Mean	Std.Dev.	RMSE
d	0.400	0.3629	0.1399	0.1447	0.200	0.3134	0.1557	0.1926
ϕ	0.630	0.6523	0.1243	0.1263	0.860	0.7873	0.0929	0.1180
σ_e	0.250	0.2498	0.0075	0.0075	0.700	0.7002	0.0174	0.0174
$1/\nu$	0.091	0.0581	0.0702	0.0775	0.063	0.0460	0.0652	0.0672
$\rho_{\zeta\xi}$	0.264	0.3135	0.0388	0.0626	0.072	0.0867	0.0681	0.0697
σ_ξ	0.155	0.1568	0.0136	0.0137	0.219	0.2150	0.0366	0.0369

(b) Second Step Estimator

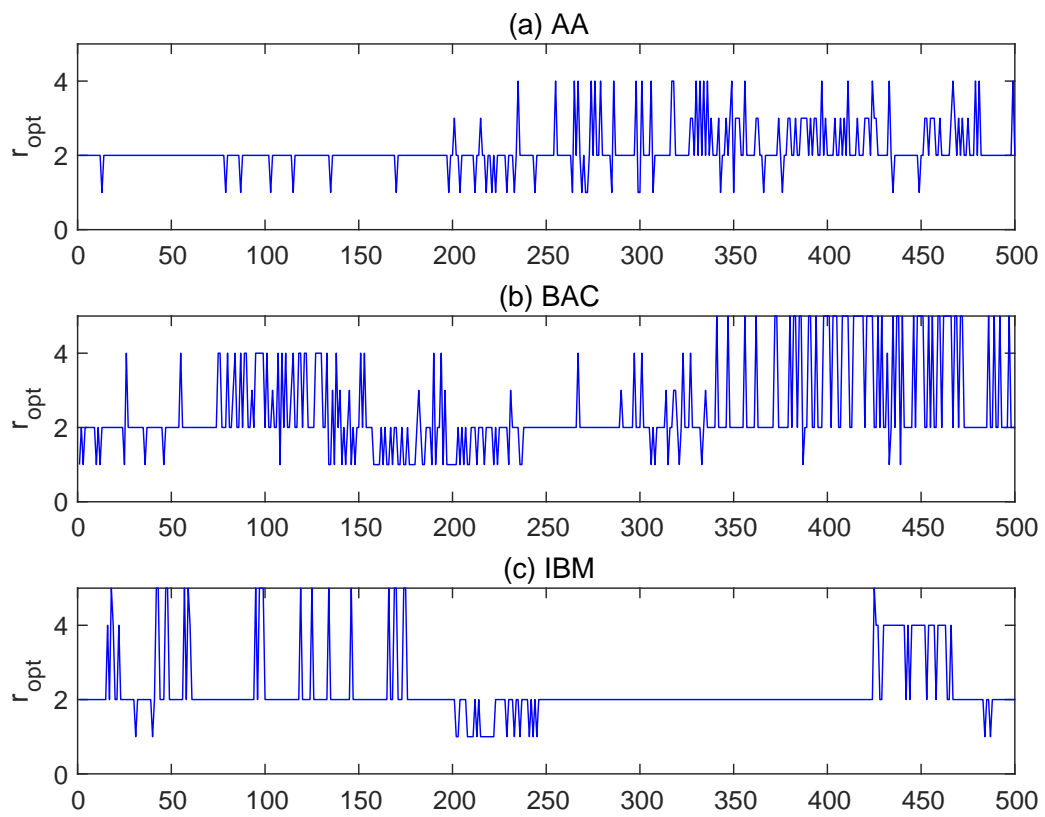| Param.	DGP 1				DGP 2			
	True	Mean	Std.Dev.	RMSE	True	Mean	Std.Dev.	RMSE
μ	-0.123	-0.1293	0.2143	0.2144	-0.250	-0.2503	0.1332	0.1332
τ_1	-0.050	-0.0164	0.0029	0.0337	-0.020	-0.0034	0.0022	0.0168
τ_2	0.050	0.0286	0.0035	0.0217	0.020	0.0060	0.0053	0.0149
σ_η	0.120	0.0820	0.0101	0.0393	0.216	0.0907	0.0235	0.1275

Note: While μ is estimated by the sample mean of x_t, the remaining parameters are estimated by OLS in equation (17).

Table 2: Finite Sample Performance of the Quasi-Likelihood Ratio Test for General Asymmetry

DGP	(τ_2, σ_η)	$\rho_{\zeta\xi}$	RF
DGP1m1	(0, 0.148)	0	0.0300
DGP1m2	(0.019, 0.143)	0.1	0.6250
DGP1m3	(0.058, 0.111)	0.3	0.9995
DGP2m1	(0, 0.218)	0	0.0445
DGP2m2	(0.028, 0.215)	0.1	0.4055
DGP2m3	(0.083, 0.174)	0.3	0.9970

Note: 'RF' denote the rejection frequency.

Table 3: Descriptive Statistics of Returns and Log-Volatilities

Data	Average	Std. Dev.	Skewness	Kurtosis
Returns				
AA	-0.1395	2.9083	-8.0238	193.85
BAC	-0.0927	2.7907	-0.3602	16.700
IBM	-0.0312	1.7676	-0.0499	8.9003
Log-Volatility				
AA	0.3300	0.7974	0.9027	4.4002
BAC	-0.0571	1.2431	0.4414	3.2899
IBM	-0.2538	0.9994	0.4812	3.3037

Table 4: ML Estimates for Realized EGARCH Models

Param	AA	BAC	IBM
φ_1	-1.1535	-1.0474	-0.9089
	(0.0652)	(0.0465)	(0.0365)
φ_2	1.0843	0.9646	1.0622
	(0.0395)	(0.0259)	(0.0252)
λ_1	-0.0361	-0.0465	-0.0233
	(0.0087)	(0.0086)	(0.0084)
λ_2	0.0759	0.0315	0.0849
	(0.0063)	(0.0020)	(0.0062)
σ_e	0.3821	0.4228	0.3775
	(0.0052)	(0.0053)	(0.0045)
ν	14.9692	7.8367	14.1215
	(5.3168)	(1.3587)	(3.6341)
μ	1.2558	0.7668	0.5701
	(0.0744)	(0.0575)	(0.0362)
ϕ	0.9742	0.9807	0.9779
	(0.0045)	(0.0036)	(0.0036)
τ_1	-0.0511	-0.0783	-0.0616
	(0.0077)	(0.0101)	(0.0074)
τ_2	0.0476	0.0209	0.0683
	(0.0057)	(0.0018)	(0.0049)
τ_0	0.3408	0.5482	0.3315
	(0.0193)	(0.0235)	(0.0171)

Note: Standard errors are in parentheses.

Table 5: Estimates for RSV-GALM Models

(a) First Step QML Estimates

Param	AA	BAC	IBM
σ_e	0.0053	0.2208	0.0072
	(0.0001)	(0.0084)	(0.0001)
ν	12.0033	18.5878	26.5050
	(0.0047)	(0.0020)	(0.0012)
$\rho_{\zeta\xi}$	0.0891	0.0762	0.1313
	(0.0036)	(0.0264)	(0.0031)
σ_ξ	0.4055	0.3437	0.4121
	(0.0117)	(0.0083)	(0.0146)
d	0.4894	0.4864	0.4730
	(0.0115)	(0.0044)	(0.0107)
ϕ	-0.0156	0.3576	0.0420
	(0.0020)	(0.0069)	(0.0026)
QLR	10.0083	1.2650	18.0633
	[0.0016]	[0.2607]	[0.0000]

Note: Standard errors are in parentheses. QLR denotes the quasi-likelihood ratio test statistic for general asymmetry. P-values are given in brackets.

(b) Second Step Estimates

Param	AA	BAC	IBM
μ	0.3186	-0.3767	-0.4611
	(0.0199)	(0.0316)	(0.0245)
τ_1	-0.0522	-0.0127	-0.0643
	(0.0090)	(0.0020)	(0.0091)
τ_2	0.0323	-0.0019	0.0610
	(0.0066)	(0.0015)	(0.0070)
τ_3	-0.0066	0.0009	-0.0027
	(0.0039)	(0.0007)	(0.0046)
τ_4		0.2428×10^{-5}	0.0004
		(0.0003)	(0.0024)
τ_5		-0.6277×10^{-6}	0.0009
		(0.0001)	(0.0012)
σ_η	0.4010	0.0899	0.3988
	(0.0063)	(0.0014)	(0.0063)
$r_{\max}$	3	5	5
r_{opt}	2	1	2

Note: Standard errors are in parentheses. Results under the optimal r are omitted as the differences are negligible.

Table 6: Simulated Moments and Cumulants

(a) Second and Fourth Moments

Data	$E(\zeta_t^2)$	$E(\zeta_t\xi_t)$	$E(\xi_t^2)$	$E(\zeta_t^4)$	$E(\zeta_t^3\xi_t)$	$E(\zeta_t^2\xi_t^2)$	$E(\zeta_t^2\xi_t^2)$	$E(\xi_t^4)$
AA	5.1045	0.0715	0.1712	174.16	1.1579	0.9148	0.0509	0.0934
BAC	5.0366	0.0567	0.1223	174.30	0.8663	0.6245	0.0222	0.0450
IBM	5.0070	0.1260	0.1827	173.01	1.8845	0.9687	0.0837	0.1046

Note: Moments are estimated via simulated data with sample size of 1000000, based on the estimated parameters in the first and second steps.

(b) Fourth Cumulants

Data	C_{2222}	C_{2223}	C_{2233}	C_{2333}	C_{3333}
AA	95.988	0.0629	−1.7075	0.0142	0.0054
BAC	98.200	0.0102	−1.2230	0.0014	0.0002
IBM	97.795	−0.0083	−1.7753	0.0147	0.0045

Note: C_{abcd} denotes the fourth cumulants of $(u_{a,t}, u_{b,t}, u_{c,t}, u_{d,t})$ with $\boldsymbol{u}_t = (e_t, \zeta_t, \xi_t)'$.

Table 7: RMSFE for Out-of-Sample Forecasts

Data	RSV-GALM	RSV-ALM	RealEGARCH
AA	0.2359	0.2475∗	0.3537∗
BAC	0.3433	0.3548∗	0.4228∗
IBM	0.3506	0.3624∗	0.3876∗

Note: '∗' denotes the test statistic for equal forecast accuracy to RSV-GALM is rejected at the five percent level.

Table 8: Frequency of Selecting r for Forecasting via RSV-GALM Models

Data	0	1	2	3	4	5	Corr.
AA	0.000	0.060	0.796	0.090	0.054	0.000	−0.1797
BAC	0.000	0.122	0.644	0.030	0.078	0.126	−0.0074
IBM	0.000	0.050	0.834	0.000	0.074	0.042	−0.2090

Note: 'Corr' indicates the correlation coefficient between x_t and the selected r for the forecasting period.