Econometric Theory, **40**, 2024, 790–826.
doi:10.1017/S0266466622000585

EXPONENTIAL REALIZED GARCH-ITÔ VOLATILITY MODELS

DONGGYU KIM
College of Business, Korea Advanced Institute of Science and Technology (KAIST)

This paper introduces a novel Itô diffusion process to model high-frequency financial data that can accommodate low-frequency volatility dynamics by embedding the discrete-time nonlinear exponential generalized autoregressive conditional heteroskedasticity (GARCH) structure with log-integrated volatility in a continuous instantaneous volatility process. The key feature of the proposed model is that, unlike existing GARCH-Itô models, the instantaneous volatility process has a nonlinear structure, which ensures that the log-integrated volatilities have the realized GARCH structure. We call this the exponential realized GARCH-Itô model. Given the autoregressive structure of the log-integrated volatility, we propose a quasi-likelihood estimation procedure for parameter estimation and establish its asymptotic properties. We conduct a simulation study to check the finite-sample performance of the proposed model and an empirical study with 50 assets among the S&P 500 compositions. Numerical studies show the advantages of the proposed model.

1. INTRODUCTION

In financial practice, volatility plays a pivotal role. Low-frequency and high-frequency financial data are widely used to analyze volatility dynamics. For example, generalized autoregressive conditional heteroskedasticity (GARCH) models are introduced to catch low-frequency volatility dynamics, such as volatility clustering, by employing the squared low-frequency log return as the innovation (Engle, 1982; Bollerslev, 1986). However, when volatility changes rapidly, it is often difficult to catch the change using only low-frequency log returns as the innovations (Andersen et al., 2003). On the other hand, high-frequency financial data are available to construct the so-called realized volatility for estimating daily integrated volatility. Examples include two-time scale realized volatility (Zhang, Mykland, and Aït-Sahalia, 2005), multiscale realized volatility (MSRV; Zhang, 2006), kernel realized volatility (Barndorff-Nielsen et al., 2008), quasi-maximum likelihood estimator (QMLE; Aït-Sahalia, Fan, and Xiu, 2010;

The author thanks the Editor (Professor Peter C.B. Phillips), the Co-Editor (Professor D. Kristensen), and anonymous two referees for their careful reading of this paper and valuable comments. The research of the author was supported in part by the National Research Foundation of Korea (NRF) (2021R1C1C1003216). Address correspondence to Donggyu Kim, College of Business, Korea Advanced Institute of Science and Technology (KAIST), Seoul, South Korea; e-mail: donggyukim@kaist.ac.kr.

Xiu, 2010), pre-averaging realized volatility (PRV; Jacod et al., 2009), and robust PRV (Fan and Kim, 2018; Shin, Kim, and Fan, 2021). These realized volatility estimators contain high-frequency information about market volatility, and many studies show that incorporating high-frequency information helps account for low-frequency market dynamics (Corsi, 2009; Shephard and Sheppard, 2010; Hansen, Huang, and Shek, 2012; Kim and Wang, 2016). Several conditional volatility models have been developed to combine high-frequency and low-frequency data and enhance volatility estimation and prediction by employing realized volatility as the volatility proxy. Examples include the realized volatility-based modeling approaches (Andersen and Bollerslev, 1997a, 1997b, 1998a, 1998b; Andersen et al., 2003), the heterogeneous autoregressive (HAR) models (Corsi, 2009), the realized GARCH models (Hansen et al., 2012), the high-frequency-based volatility (HEAVY) models (Shephard and Sheppard, 2010), and the unified GARCH/SV-Itô models (Kim and Wang, 2016; Kim and Fan, 2019; Song et al., 2021). These models have been developed based on the linear autoregressive structure of realized volatilities. However, we often observe that nonlinear autoregressive structures, such as exponential functions, better capture the volatility dynamics (Nelson, 1991; Kawakatsu, 2006; Hansen and Huang, 2016). This may be because log volatilities often have a stronger linear autoregressive relationship. In fact, when variables are close to normal distributions, linear models work well. To check the normality of realized volatilities, we draw normal quantile–quantile (QQ) plots of realized volatilities and log-realized volatilities for Apple Inc. (AAPL) stock. Figure 1 shows that the log transformation makes the realized volatilities closer to a normal distribution. Most assets show the same phenomena. Thus, we can conjecture that log volatilities better explain volatility dynamics. To harness this feature, Hansen and Huang (2016) employed the exponential GARCH structure with the log-realized volatility as the innovation, and their empirical study indicated that the nonlinear GARCH structure helps account for market dynamics. Although, as discussed above, empirical studies have supported that incorporating high-frequency data with a nonlinear autoregressive structure better captures the market dynamics, the mathematical gap between the empirical low-frequency discrete-time nonlinear volatility models, such as the exponential realized GARCH, and high-frequency-based continuous-time diffusion processes has not been thoroughly studied. In fact, several studies have been conducted to fill the gap between the discrete-time volatility models and continuous-time diffusion processes (Kim and Wang, 2016; Kim and Fan, 2019; Song et al., 2021). However, these studies were based on a linear autoregressive structure, and the extension from linear to nonlinear structures is not straightforward. This fact increases the need for developing continuous-time diffusion process-based models that provide a rigorous mathematical formulation for the nonlinear autoregressive structure of realized volatilities.

In this paper, we develop a novel diffusion process to model high-frequency financial data that can accommodate a nonlinear GARCH structure of the realized volatilities. From empirical studies, we often observe that the log-realized volatility

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

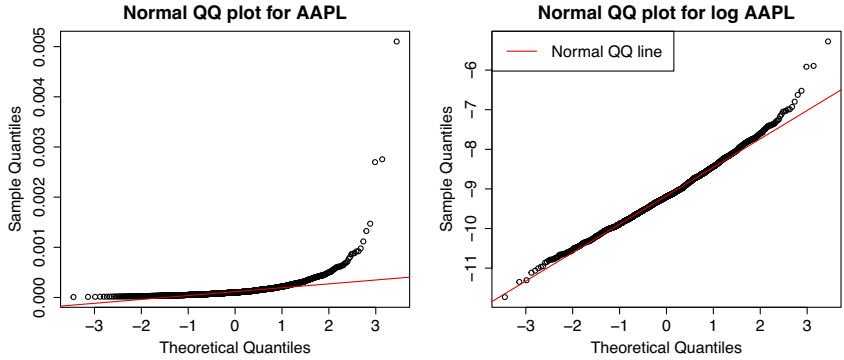

FIGURE 1. Normal QQ plots of realized volatilities and log-realized volatilities for AAPL. The red real line denotes the best linear fit line, which illustrates perfect normal distribution.

has a stronger autoregressive structure than does the original realized volatility. To reflect this, we employ the exponential GARCH structure as the nonlinear function. Specifically, the log-realized volatility follows the realized GARCH structure. To connect this low-frequency volatility structure with the continuous-time diffusion process, we develop a novel continuous instantaneous volatility process. Since the volatility process has a nonlinear structure, the linear structure of the unified GARCH-Itô (UGARCH; Kim and Wang, 2016) is not applicable. Furthermore, usual log-diffusion processes for instantaneous volatility processes cannot provide the solution. To tackle this issue, we propose a novel instantaneous volatility process based on the average integrated volatility process. The proposed instantaneous volatility process is then continuous with respect to time, and its daily integrated volatility is decomposed into the exponential realized GARCH with the daily log-integrated volatility as the innovation and exponential martingale difference. We call it the exponential realized GARCH-Itô (ERGI) model. Unlike the linear realized GARCH model, the log-realized volatility can have negative values. Thus, we allow model parameters to be negative. To estimate the model parameters, we propose a quasi-maximum likelihood estimation procedure. Specifically, we adopt the Gaussian quasi-likelihood function and use the realized volatility as the proxy of the conditional expected value. On the other hand, based on the linear relationship of the log-integrated volatility, we employ the ordinary least-squares estimation. Furthermore, we establish their asymptotic properties. To illustrate the benefit of the proposed model, we apply the ERGI model to real high-frequency trading data and find that the exponential structure helps account for the volatility dynamics.

The rest of paper is organized as follows. In Section 2, we propose the ERGI model and investigate its properties. In Section 3, we suggest the quasi-maximum likelihood estimation procedures and study their asymptotic behaviors. In Section 4, we conduct a simulation study to check the finite-sample performance

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

of the ERGI model. In Section 5, we apply the ERGI model to the top 50 high trading volume assets among the S&P 500 compositions. Section 6 contains the conclusions. All the technical proofs are collected in Section 7.

2. EXPONENTIAL REALIZED GARCH–ITÔ MODELS

In this section, we develop an exponential realized GARCH-Itô (ERGI) model as follows.

DEFINITION 1. *We call the log price X_t to follow the ERGI process if it satisfies*

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t(\theta) dB_t + J_t d\Lambda_t, \\ \sigma_t^2(\theta) &= \bar{\sigma}_t^2(\theta)\{1 + (t - [t])b_t(\theta)\}, \\ b_t(\theta) &= b_{[t]}(\theta) + (t - [t])(\omega + (\gamma - 1)b_{[t]}(\theta)) + \beta \log \bar{\sigma}_t^2(\theta) \\ &\quad - (1 - t + [t])\left(\beta + \beta^*(t - [t])\right)\log \sigma_{[t]}^2 + \nu(1 - t + [t])Z_t^2, \end{aligned}$$

where $\bar{\sigma}_t^2(\theta) = (t - [t])^{-1}\int_{[t]}^t \sigma_s^2(\theta)ds$, $[t]$ denotes the integer part of t, except that $[t] = t - 1$ when t is an integer, $Z_t = \int_{[t]}^t dW_s$, $\beta^* = \frac{1+\beta\varrho_2}{\varrho_2 - 2\varrho_3}$, where $\varrho_2 = \beta^{-2}(e^\beta - 1 - \beta)$ and $\varrho_3 = \beta^{-3}(e^\beta - 1 - \beta - \beta^2/2)$, and $\theta = (\omega, \gamma, \beta, \nu)$ is the model parameter. For the jump part, J_t is the jump size and Λ_t is the Poisson process with the intensity λ_t.

The log-average integrated volatility, $\log \bar{\sigma}_t^2(\theta)$, provides the innovation term, and Z_t is the random fluctuation. By the construction, the process $b_t(\theta)$ is continuous with respect to time t; thus, the instantaneous volatility process $\sigma_t^2(\theta)$ is also continuous. At the integer time points, we have

$$b_n(\theta) = \omega + \gamma b_{n-1}(\theta) + \beta \log \int_{n-1}^n \sigma_t^2(\theta)dt.$$

That is, $b_n(\theta)$ can be explicitly expressed by the past log-integrated volatilities, and $b_t(\theta)$ has a form of the interpolation between these values. Thus, when considering $b_t(\theta)$, the ERGI model has a similar structure to the realized GARCH-Itô model (Song et al., 2021) with the log-integrated volatility as the innovations. However, unlike the realized GARCH-Itô model, to obtain the nonlinear exponential realized GARCH form, the instantaneous volatility process has a nonlinear structure, such as $\bar{\sigma}_t^2(\theta)\{1 + (t - [t])b_t(\theta)\}$. The solution for this structure is

$$\frac{1}{t - n + 1}\int_{n-1}^t \sigma_s^2(\theta)ds = \sigma_{n-1}^2(\theta)e^{\int_{n-1}^t b_s(\theta)ds} \quad \text{a.s.}$$

Details can be found in Lemma 1. Using the above solution, we can measure the integrated volatility $\int_{n-1}^n \sigma_t^2(\theta)dt$ with $\sigma_{n-1}^2(\theta)$ and $\int_{n-1}^n b_s(\theta)ds$, which has the realized GARCH form with the log-integrated volatilities. Specific properties of the integrated volatility are shown in the following theorem.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

THEOREM 1. *Under the ERGI model, for $|\beta| < 1$ and $n \in \mathbb{N}$, the integrated volatilities have the following properties.*

(a) *We have*

$$\int_{n-1}^{n} \sigma_t^2(\theta)dt = \exp(h_n(\theta) + D_n) \quad a.s.,$$

$$h_n(\theta) = \omega^* + \gamma h_{n-1}(\theta) + \beta^g \log \int_{n-2}^{n-1} \sigma_t^2(\theta)dt,$$

where

$$\omega^* = \{(1-\gamma)\varrho_2 + \varrho\}\omega + (1-\gamma)\nu(\varrho_2 - 2\varrho_3), \quad \beta^g = \varrho\beta, \quad \varrho = \varrho_1 + (\gamma-1)\varrho_2,$$
$$\varrho_1 = \beta^{-1}(e^\beta - 1), \quad \varrho_2 = \beta^{-2}(e^\beta - 1 - \beta), \quad \varrho_3 = \beta^{-3}(e^\beta - 1 - \beta - \beta^2/2),$$

and

$$D_n = 2\nu \int_{n-1}^{n} \{(n-t)\beta^{-1}e^{\beta(n-t)} - (e^{\beta(n-t)} - 1)\beta^{-2}\}Z_t dW_t$$

is a martingale difference.

(b) *We have*

$$\int_{n-1}^{n} \sigma_t^2(\theta)dt = \exp(H_n(\theta))M_n,$$
$$\mathbb{E}\left[\int_{n-1}^{n} \sigma_t^2(\theta)dt \middle| \mathcal{F}_{n-1}\right] = \exp(H_n(\theta)) \quad a.s., \tag{2.1}$$

where

$$H_n(\theta) = \omega^g + \gamma H_{n-1}(\theta) + \beta^g \log \int_{n-2}^{n-1} \sigma_t^2(\theta)dt,$$

$$\omega^g = \omega^* + (1-\gamma)\log \mathbb{E}[\exp(D_n)],$$

and

$$M_n = \exp(D_n - \log \mathbb{E}[\exp(D_n)])$$

is an exponential martingale difference.

Theorem 1(a) shows that the log-integrated volatility is decomposed into the realized GARCH with the log-integrated volatility innovations, $h_n(\theta)$, and the martingale difference D_n. Thus, the log-realized GARCH, $h_n(\theta)$, is the conditional expected value of the log-integrated volatility, but it is not the conditional expected value of the original integrated volatility. In Theorem 1(b), we show that the integrated volatility is decomposed into the exponential function of the realized GARCH with the log-integrated volatility innovations, $H_n(\theta)$, which has the additional interceptor term from the martingale difference term D_n, and the exponential martingale difference M_n. Since the expectation is a linear function, the additional interceptor term does not appear in the linear realized GARCH. However, the ERGI

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

is nonlinear, and thus we have the additional interceptor term. The main purpose of this paper is to develop a model for analyzing the original integrated volatility, so we develop a statistical inference based on (2.1). Theorem 1 indicates that the proposed model has a nonlinear exponential GARCH structure. From the empirical study, we find that this nonlinear structure helps explain the volatility dynamics as compared to the usual linear realized GARCH. Details can be found in Section 5.

Remark 1. Klüppelberg, Lindner, and Maller (2004) suggested continuous-time processes whose instantaneous volatility process has the ARCH and GARCH structures, which is called the continuous-time GARCH (COGARCH), and Haug and Czado (2007) extended this to the exponential GARCH structure. These models tried to develop a Lévy diffusion process whose instantaneous volatility has the GARCH-type form, and they do not link low-frequency dynamics and continuous diffusion process. In other words, the COGARCH model was developed to explain continuous-time dynamics. However, the aim of the proposed ERGI is to explain low-frequency dynamics using high-frequency information. Thus, its daily integrated volatility has the exponential GARCH form. The relationship with the daily log returns is presented in the following section.

2.1. Relationship with the Daily Log Returns

The traditional discrete GARCH models are models of close-to-close volatilities for log returns. In this section, we discuss the relationship between the proposed ERGI and close-to-close volatilities for log returns.

We first consider the continuous part. That is, we assume that the log-price process does not have a jump component. By Itô's lemma and Theorem 1(b), we then have

$$\mathbb{E}\left[\left(X_n - X_{n-1} - \int_{n-1}^{n} \mu_t dt\right)^2 \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left[\left(\int_{n-1}^{n} \sigma_t(\theta) dB_t\right)^2 \middle| F_{n-1}\right]$$
$$= \exp\left(H_n(\theta)\right) \quad \text{a.s.}$$

Thus, the exponential GARCH volatility, $\exp(H_n(\theta))$, is the conditional volatility of the daily log return. Unfortunately, in practice, we do not have observations during the close-to-open period; thus, in order to investigate the close-to-close volatility, we need to impose a structure on the overnight period. For example, we can simply use squared close-to-open log returns as a proxy of the integrated volatility for the close-to-open period. We can then apply the proposed ERGI model to the realized volatility plus squared close-to-open log return. On the other hand, we can assume that the close-to-open volatility dynamics are the same as in the open-to-close period. We then only need to match the scale. To do this, we can calculate the averages for the open-to-close realized volatilities and the squared close-to-open log returns, and we multiply the inverse of the proportion of the average of the realized volatility. In the empirical study, we investigate the whole-day market dynamics (see Section 5.1). The above methods are practical solutions

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

without theoretical justifications. Thus, it would be interesting and important to develop a diffusion process that can accommodate the close-to-close period. We leave this for a future study.

To investigate the jump diffusion process, we assume that the jump sizes J_t's are i.i.d. with mean μ_J and variance σ_J^2, and the intensity is constant over time; that is, $\lambda_t = \lambda$. Furthermore, we assume the continuous and jump parts are not correlated. Then, the conditional volatility of the daily log return is

$$\mathbb{E}\left[\left(X_n - X_{n-1} - \int_{n-1}^{n} \mu_t dt\right)^2 \middle| \mathcal{F}_{n-1}\right] = \exp\left(H_n(\theta)\right) + \lambda\sigma_J^2 + \lambda^2\mu_J^2 \text{ a.s.}$$

The squared log return has the exponential GARCH, $\exp\left(H_n(\theta)\right)$, and additional expected jump variation. The jump variation part depends on the assumption of the jump structure. Thus, it would be interesting and important to investigate the jump variation dynamics and to model the jump component. We leave this for a future study.

3. ESTIMATION PROCEDURE

3.1. A Model Setup

We assume that the log prices follow the ERGI process defined in Definition 1. The intraday log prices for the dth day are observed at $t_{d,i}, i = 1, \ldots, m_d$, where $d - 1 = t_{d,0} < t_{d,1} < \cdots < t_{d,m_d} = 1 + d - 1$. We denote m as the average number of the high-frequency observations; that is, $m = \frac{1}{n}\sum_{d=1}^{n} m_d$. Unfortunately, true high-frequency observations, $X_{t_{d,i}}$'s, are not observed due to market microstructure noises. To accommodate the market microstructure noises, we assume that the observed log prices $Y_{t_{d,i}}$ have the following structure:

$$Y_{t_{d,i}} = X_{t_{d,i}} + \epsilon_{t_{d,i}}, \quad \text{for } d = 1, \ldots, n, i = 1, \ldots, m_d,$$

where X_t is the true log price and $\epsilon_{t_{d,i}}$'s are microstructure noises with mean zero.

Without the presence of price jumps, several nonparametric realized volatility estimators have been constructed that take advantage of subsampling and local-averaging techniques to remove the effect of market microstructure noises so that the integrated volatility can be estimated consistently and efficiently (Zhang, 2006; Barndorff-Nielsen et al., 2008; Jacod et al., 2009; Xiu, 2010; Fan and Kim, 2018; Shin et al., 2021). To identify the jump locations, given noisy high-frequency data, Fan and Wang (2007) and Zhang, Kim, and Wang (2016) proposed wavelet methods to detect jumps and applied the MSRV method to jump-adjusted data. They showed that the estimator of jump variation has the convergence rate of $m^{-1/4}$, and the estimator of integrated volatility achieves the optimal convergence rate of $m^{-1/4}$. On the other hand, Aït-Sahalia and Xiu (2016) proposed jump robust pre-averaging methods by employing a truncation method. They also demonstrated that the estimators of jump variation and integrated volatility achieve the convergence rate of $m^{-1/4}$, which is known as the optimal rate with

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

the presence of the microstructure noise. In this paper, for the ith day, we let RV_i be the corresponding estimator of daily integrated volatility that is robust to microstructure noises and price jumps. In the numerical study, we employ the jump robust pre-averaging method.

3.2. GARCH Parameters Estimation

We first fix notations. For a given vector $\mathbf{x} = (x_i)_{i=1,\dots,k}$, we define $\|\mathbf{x}\|_{\max} = \max_i |x_i|$. Let C's be generic constants whose values do not depend on θ^g, n, and m and may change from occurrence to occurrence.

3.2.1. *Quasi-Maximum Likelihood Estimation with Gaussian Likelihood Function.* In this section, we develop an estimation procedure for the true GARCH model parameters $\theta_0^g = (\omega_0^g, \gamma_0, \alpha_0^g)$.

Theorem 1 indicates that the integrated volatility is decomposed into the exponential GARCH term $\exp(H_i(\theta^g))$ and the exponential martingale difference term M_i, which implies

$$\frac{\int_{i-1}^{i} \sigma_t^2(\theta_0^g)dt - \exp\left(H_i(\theta_0^g)\right)}{\exp\left(H_i(\theta_0^g)\right)} = M_i - 1 \quad \text{a.s.}$$

Since M_i is the exponential martingale difference, $M_i - 1$ is a martingale difference and stationary. This inspires us to use integrated volatility as a proxy for exponential GARCH volatility. On the other hand, for low-frequency time series models, such as GARCH-type models, we often use the squared log returns as a proxy, and to estimate the model parameters, we often employ the Gaussian likelihood function. In other words, the QMLE with the Gaussian likelihood function is widely used. In terms of the squared log return, the Gaussian likelihood function has the squared log returns as the proxy of the conditional volatility. This motivates us to use the integrated volatility instead of the squared log returns as the proxy of the conditional volatility in the Gaussian likelihood function as follows:

$$\widehat{L}_n(\theta^g) = -\frac{1}{n}\sum_{i=1}^{n}\left\{H_i(\theta^g) + \frac{\int_{i-1}^{i}\sigma_t^2(\theta_0^g)dt}{\exp(H_i(\theta^g))}\right\}.$$

We can estimate the parameter θ^g by maximizing the above quasi-likelihood function. However, in practice, the integrated volatility is not observable, so we need to estimate it first. We employ the jump robust realized volatility estimator (Fan and Wang, 2007; Aït-Sahalia and Xiu, 2016; Zhang et al., 2016). We then estimate the log-conditional expectation of the integrated volatilities as follows:

$$\widehat{H}_i(\theta^g) = \omega^g + \gamma\widehat{H}_{i-1}(\theta^g) + \beta^g \log RV_{i-1}, \tag{3.1}$$

where the initial value $\widehat{H}_1(\theta^g)$ is set to be $\log RV_1$. The effect of the initial value is negligible with the rate of n^{-1} (see Lemma 1 in Kim and Wang, 2016), so its choice does not have a significant effect on the parameter estimation. With the estimated

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

conditional expected volatility function, we define the following quasi-likelihood function:

$$\widehat{L}_{n,m}(\theta^g) = -\frac{1}{n}\sum_{i=1}^{n}\left\{\widehat{H}_i(\theta^g) + \frac{RV_i}{\exp\left(\widehat{H}_i(\theta^g)\right)}\right\}.$$

Using the proof of Lemma 2, we can show the following uniform convergence:

$$\sup_{\theta\in\Theta}\left|\widehat{L}_{n,m}(\theta) - \widehat{L}_n(\theta)\right| = O_p(m^{-1/4}),$$

where Θ is the parameter space defined in Assumption 1(a). Thus, the asymptotic results for $\widehat{L}_{n,m}(\theta^g)$ are the same as those for $\widehat{L}_n(\theta^g)$ with additional $m^{-1/4}$ order. We then obtain the estimator for the GARCH parameters θ_0^g by maximizing the above quasi-likelihood function,

$$\widehat{\theta^g} = \arg\max_{\theta^g\in\Theta}\widehat{L}_{n,m}(\theta^g),$$

where Θ is the parameter space of θ^g. To establish its asymptotic properties, we need the following technical conditions.

Remark 2. Even if the effect of the initial value is negligible, for the finite sample, the random variable $\log RV_1$ happens to be far from the true initial value. To handle this practical issue, we can assume that the initial value is a long-term average. Under this condition, we can use the theoretical average value $\frac{\omega^g}{1-\gamma-\beta^g}$ as the initial value. That is, we additionally assume that the initial value is $H_1(\theta_0^g) = \frac{\omega_0^g}{1-\gamma_0-\beta_0^g}$. With this condition, we can obtain the same asymptotic result derived in Theorem 2.

Assumption 1.

(a) $\theta_0^g \in \Theta = \{(\omega^g, \gamma, \beta^g); \omega_l < |\omega^g| < \omega_u, \gamma_l < |\gamma| < \gamma_u < 1, \beta_l < |\beta| < \beta_u < 1, |\gamma + \beta^g| < 1\}$, where $\omega_l, \omega_u, \gamma_l, \gamma_u, \beta_l, \beta_u$ are some known constants;

(b) $\sup_i \mathbb{E}\left[\left|RV_i - \int_{i-1}^{i}\sigma_t^2(\theta_0^g)dt\right|^4\right]^{1/4} \le Cm^{-1/4}$ and $\sup_i \mathbb{E}\left[\left|\log RV_i - \log\int_{i-1}^{i}\sigma_t^2(\theta_0^g)dt\right|^4\right]^{1/4} \le Cm^{-1/4}$;

(c) $\mathbb{E}\left[\left|\int_{i-1}^{i}\sigma_t^2(\theta_0^g)dt\right|^4\right] \le C$, $\mathbb{E}\left[\sup_{\theta^g\in\Theta}\exp(4|H_i(\theta^g)|)\right] \le C$, $\mathbb{E}\left[\sup_{\theta^g\in\Theta}\exp(4|\widehat{H}_i(\theta^g)|)\right] \le C$, and $\mathbb{E}\left[M_i^4\right] \le C$ for all i.

Remark 3. Under Assumption 1(a), unlike the linear GARCH-Itô models (Kim and Wang, 2016; Song et al., 2021), we allow the parameters to be negative. The condition $|\gamma + \beta^g| < 1$ provides stationary properties of conditional volatilities. There exist realized volatility estimators satisfying Assumption 1(b) under some finite moment conditions (see Tao et al., 2011; Kim, Wang, and Zou, 2016). The sufficient condition for Assumption 1(c) is that $\mathbb{E}\left[\exp(s|\log RV_i|)\right] \le C$ and $\mathbb{E}\left[\exp(s|\log\int_{i-1}^{i}\sigma_t^2(\theta_0^g)dt|)\right] \le C$ for $s \ge 4/(1-\gamma_u)$.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

In the following theorem, we establish the asymptotic properties of the proposed QMLE.

THEOREM 2. *Under Assumption 1, we have*

$$\|\widehat{\theta^g} - \theta_0^g\|_{\max} = O_p(n^{-1/2} + m^{-1/4}).$$

Furthermore, we suppose that $nm^{-1/2} \to 0$ *and Assumption 1 is satisfied. Then, we have*

$$\sqrt{n}(\widehat{\theta^g} - \theta_0^g) \xrightarrow{d} N(0, AV^{-1}),$$

where $A = \mathbb{E}\left[(1 - M_i)^2\right]$ *and* $V = \mathbb{E}\left[\frac{\partial H_i(\theta^g)}{\partial \theta^g} \frac{\partial H_i(\theta^g)}{\partial (\theta^g)^\top}\Big|_{\theta^g = \theta_0^g}\right]$.

Remark 4. Theorem 2 shows that the QMLE $\widehat{\theta^g}$ has the convergence rate $n^{-1/2} + m^{-1/4}$. The $n^{-1/2}$ term is the usual convergence rate due to the low-frequency errors, $M_i - 1$. The $m^{-1/4}$ term is the cost to estimate the integrated volatility, which is known as the optimal rate with the presence of the microstructure noise. Specifically, by Theorem 1, we have the following relationship:

$$\int_{n-1}^{n} \sigma_t^2(\theta^g) dt = \exp\left(H_n(\theta^g)\right) M_n \quad \text{a.s.},$$

and, additionally, due to the estimation error of the latent integrated volatility, we have

$$RV_n = \int_{n-1}^{n} \sigma_t^2(\theta^g) dt + E_n = \exp\left(H_n(\theta^g)\right) M_n + E_n \quad \text{a.s.},$$

where E_n is the estimation error of the latent integrated volatility. The error rate of E_n is $m^{-1/4}$, and its asymptotic variance is specified in the literature on estimating integrated volatility (Zhang, 2006; Barndorff-Nielsen et al., 2008; Jacod et al., 2009; Aït-Sahalia et al., 2010; Xiu, 2010). Then, the asymptotic variance of $\widehat{\theta^g}$ in Theorem 2 has an additional term that is a function of variance of E_n. For example, we have

$$\|\widehat{\theta^g} - \theta_0^g\|_{\max} \approx C_1 \frac{\sqrt{A}}{n^{1/2}} + C_2 \frac{\sqrt{Avar_{RV}}}{m^{1/4}},$$

where $Avar_{RV}$ is the asymptotic variance of RV, and C_1 and C_2 are functions of $H_i(\theta_0^g)$.

Remark 5. The condition $nm^{-1/2} \to 0$ is required to remove the effect from the estimation error of the realized volatility when establishing the asymptotic normality. However, as in the realized GARCH model (Hansen et al., 2012), if we assume that the conditional volatility is a function of the realized volatility estimator RV_i, this assumption is not required.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

3.2.2. *Least-Squares Estimation.* In the previous section, we developed the parameter estimation procedure for the integrated volatility. In this section, we study an estimation procedure for the log-integrated volatility. Theorem 1(a) indicates that the log-integrated volatility has the following relationship:

$$\log \int_{i-1}^{i} \sigma_t^2(\theta)dt = h_i(\theta) + D_i \quad \text{a.s.}, \tag{3.2}$$

where D_i's are i.i.d. random noise with mean zero. Since $\log \int_{i-1}^{i} \sigma_t^2(\theta)dt$ and $\int_{i-1}^{i} \sigma_t^2(\theta)dt$ have different conditional expected values, the parameter of interest is not the same as θ^g. Specifically, for the log-integrated volatility, the target parameter is $\theta_0^* = (\omega_0^*, \gamma_0, \alpha_0^g)$. We note that the difference between ω_0^* and ω_0^g is $(1-\gamma)\log \mathbb{E}[\exp(D_i)]$. The relationship (3.2) inspires us to employ the ordinary least-squares (OLS) estimation procedure. To harness the OLS procedure, we need to estimate the integrated volatility. For example, we estimate the log-conditional expectation of the integrated volatilities as follows:

$$\widehat{h}_i(\theta^*) = \omega^* + \gamma \widehat{h}_{i-1}(\theta^g) + \beta^g \log RV_{i-1}, \tag{3.3}$$

where the initial value $\widehat{h}_1(\theta^*)$ is set to be $\log RV_1$. Then, with the realized volatility, we have

$$\log RV_i = \widehat{h}_i(\theta_0^*) + \{\log RV_i - \log \int_{i-1}^{i} \sigma_t^2(\theta)dt\} + \{h_i(\theta_0^*) - \widehat{h}_i(\theta_0^*)\} + D_i \quad \text{a.s.}$$

As we discussed in the previous section, the error terms $\log \int_{i-1}^{i} \sigma_t^2(\theta)dt - \log RV_i$ and $h_i(\theta_0^*) - \widehat{h}_i(\theta_0^*)$ have the error rate $m^{-1/4}$. We can consider these errors as new random noises such as $\widetilde{D}_i = \{\log RV_i - \log \int_{i-1}^{i} \sigma_t^2(\theta)dt\} + \{h_i(\theta_0^*) - \widehat{h}_i(\theta_0^*)\} + D_i$. We then have

$$\log RV_i = \widehat{h}_i(\theta_0^*) + \widetilde{D}_i \quad \text{a.s.},$$

and we can obtain the OLS estimator as follows:

$$\widehat{\theta^*} = \text{argmin}_{\theta^* \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \left\{\log RV_i - \widehat{h}_i(\theta^*)\right\}^2.$$

Similar to Theorem 2, we can establish the following asymptotic properties of the proposed OLS estimator.

THEOREM 3. *Under Assumption 1, we have*

$$\|\widehat{\theta^*} - \theta_0^*\|_{\max} = O_p(n^{-1/2} + m^{-1/4}).$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Furthermore, we suppose that $nm^{-1/2} \to 0$ *and Assumption 1 is satisfied. Then, we have*

$$\sqrt{n}(\widehat{\theta^*} - \theta_0^*) \xrightarrow{d} N(0, A^* V^{*-1}),$$

where $A^* = \mathbb{E}\left[D_i^2\right]$ *and* $V^* = \mathbb{E}\left[\frac{\partial h_i(\theta^*)}{\partial \theta^*} \frac{\partial h_i(\theta^*)}{\partial (\theta^*)^\top}\big|_{\theta^*=\theta_0^*}\right]$.

Since this estimation procedure is developed based on the log-integrated volatility, to predict the future volatility, we need a convexity adjustment. Using Theorem 1, we can adjust the bias as follows:

$$\exp(\widehat{h}_n(\widehat{\theta^*})) \times \frac{1}{n}\sum_{i=1}^{n} \exp(\log RV_i - \widehat{h}_i(\widehat{\theta^*})).$$

3.3. Hypothesis Tests

In financial practice, we are interested in statistical inferences about the GARCH parameters, such as hypothesis tests. In this section, we discuss how to conduct hypothesis tests for the GARCH parameters.

First, we consider the integrated volatility. Theorem 2 implies that

$$\sqrt{n}(\widehat{\theta^g} - \theta_0^g) \xrightarrow{d} N(0, AV^{-1}),$$

where $A = \mathbb{E}\left[(1 - M_i)^2\right]$ and $V = \mathbb{E}\left[\frac{\partial H_i(\theta^g)}{\partial \theta^g} \frac{\partial H_i(\theta^g)}{\partial (\theta^g)^\top}\big|_{\theta^g=\theta_0^g}\right]$. To evaluate the asymptotic variances of the GARCH parameter estimators, we first need to estimate A and V. We use the following estimators:

$$\widehat{A} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{RV_i - \widehat{H}_i(\widehat{\theta^g})}{\widehat{H}_i(\widehat{\theta^g})}\right)^2 \quad \text{and} \quad \widehat{V}(\theta^g) = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \widehat{H}_i(\theta^g)}{\partial \theta^g} \frac{\partial \widehat{H}_i(\theta^g)}{\partial (\theta^g)^\top},$$

where $\widehat{H}_i(\theta^g)$ is defined in (3.1). Under some stationary conditions, we can establish their consistency. Then, by Slutsky's theorem, we can obtain

$$T_{i,n} = \frac{\sqrt{n}(\widehat{\theta^g}_i - \theta_{0i}^g)}{\sqrt{\widehat{A}\widehat{V}_{ii}^{-1}(\widehat{\theta^g})}} \xrightarrow{d} N(0,1),$$

where $\widehat{\theta^g}_i$ and θ_{0i}^g are the ith elements of $\widehat{\theta^g}$ and θ_0^g, respectively, and $\widehat{V}_{ii}^{-1}(\widehat{\theta^g})$ is the ith diagonal element of $\widehat{V}^{-1}(\widehat{\theta^g})$.

Similarly, we can estimate the asymptotic variances for the log-integrated volatility as follows:

$$\widehat{A}^* = \frac{1}{n}\sum_{i=1}^{n}\left(\log RV_i - \widehat{h}_i(\widehat{\theta^*})\right)^2 \quad \text{and} \quad \widehat{V}^*(\theta^*) = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \widehat{h}_i(\theta^*)}{\partial \theta^*} \frac{\partial \widehat{h}_i(\theta^*)}{\partial (\theta^*)^\top},$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

where $\widehat{h}_i(\theta^*)$ is defined in (3.3). Then, we can obtain

$$T^*_{i,n} = \frac{\sqrt{n}(\widehat{\theta^*}_i - \theta^*_{0i})}{\sqrt{\widehat{A}^* \widehat{V}^{*-1}_{ii}(\widehat{\theta^*})}} \xrightarrow{d} N(0,1),$$

where $\widehat{\theta^*}_i$ and $\theta^g *_{0i}$ are the ith elements of $\widehat{\theta^*}$ and θ^*_0, respectively, and $\widehat{V}^{*-1}_{ii}(\widehat{\theta^*})$ is the ith diagonal element of $\widehat{V}^{*-1}(\widehat{\theta^*})$. Thus, using the proposed Z-statistics $T_{i,n}$ and $T^*_{i,n}$, we can conduct the hypothesis tests based on the standard normal distribution.

4. A SIMULATION STUDY

We conducted Monte Carlo simulations to check the finite-sample performance of the ERGI model. The log prices were generated from the ERGI model given in Definition 1 for n days with m high-frequency observations. The model parameters were set to be $(\omega_0, \gamma_0, \beta_0, \nu_0) = (-0.1, 0.3, 0.5, 2)$ and $\mu_t = 0$. Then, the GARCH parameters $(\omega^g_0, \gamma_0, \beta^g_0) = (0.3207, 0.3, 0.4405)$. For the jump part, we set the intensity $\lambda_t = 10$ and the jump size $|J_t| = 0.05$. The signs of the jump size were randomly generated. Let $t_{d,j} = d - 1 + j/m$, for $d = 1, \ldots, n$ and $j = 0, \ldots, m$. We generated the noisy observations as follows:

$$Y_{t_{d,j}} = X_{t_{d,j}} + \epsilon_{t_{d,j}}, \qquad \text{for } d = 1, \ldots, n \text{ and } j = 0, \ldots, m,$$

where $\epsilon_{t_{d,j}}$'s are i.i.d. normal random variables with mean zero and standard deviation $0.01\sqrt{\int_{d-1}^{d} \sigma_t^2(\theta^g)dt}$. To generate the true process, we chose $m = 11,700$. We varied n from 100 to 500 and m from 390 to 11,700, which corresponds to the number of minutes and 2 seconds during the open-to-close period, respectively. We used $Y_{t_{d,j}}$ as the high-frequency observations. To estimate the integrated volatilities, we used the jump robust pre-averaging method (Jacod et al., 2009; Aït-Sahalia and Xiu, 2016) as follows:

$$RV_i = \frac{1}{\psi K} \sum_{k=1}^{m-K+1} \left\{ \bar{Y}^2(t_{i,k}) - \frac{1}{2} \widehat{Y}^2(t_{i,k}) \right\} \mathbf{1}_{\{|\bar{Y}(t_{i,k})| \le \tau_m\}},$$

where we take the weight function $g(x) = x \wedge (1 - x)$, the bandwidth size $K = \lfloor m^{1/2} \rfloor$,

$$\bar{Y}(t_{i,k}) = \sum_{l=1}^{K-1} g\left(\frac{l}{K}\right) \left(Y_{t_{i,k+l}} - Y_{t_{i,k+l-1}}\right), \quad \psi = \int_0^1 g(t)^2 dt,$$

$$\widehat{Y}^2(t_{i,k}) = \sum_{l=1}^{K} \left\{ g\left(\frac{l}{K}\right) - g\left(\frac{l-1}{K}\right) \right\}^2 \left(Y_{t_{i,k+l-1}} - Y_{t_{i,k+l-2}}\right)^2,$$

$\mathbf{1}_{\{\cdot\}}$ is an indicator function, and $\tau_m = c_\tau m^{-0.235}$ is a truncation level for the constant c_τ. We chose c_τ as four times the sample standard deviation of the pre-averaged

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

TABLE 1. MSEs for the parameter estimates of the QMLE and OLS estimators with $n = 100, 200, 500$ and $m = 390, 1170, 11,700$

		QMLE			OLS		
n	m	ω^g	γ	β^g	ω^*	γ	β^g
100	390	0.0854	0.1312	0.0468	0.0168	0.0509	0.0226
	1,170	0.0852	0.1217	0.0415	0.0145	0.0395	0.0159
	11,700	0.0865	0.1204	0.0408	0.0156	0.0401	0.0122
200	390	0.0435	0.0714	0.0309	0.0050	0.0207	0.0132
	1,170	0.0428	0.0690	0.0280	0.0052	0.0168	0.0080
	11,700	0.0453	0.0720	0.0274	0.0056	0.0157	0.0054
500	390	0.0296	0.0484	0.0213	0.0017	0.0122	0.0106
	1,170	0.0249	0.0395	0.0177	0.0016	0.0076	0.0053
	11,700	0.0244	0.0367	0.0171	0.0017	0.0053	0.0024

prices $m^{1/4}\bar{Y}(t_{d,k})$. We estimated the parameters using the procedure in Section 3. We repeated the whole procedure 500 times.

To check the performance of the realized volatility estimator, we calculated squared relative errors as follows:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{RV_i - \int_{i-1}^{i}\sigma_t^2(\theta_0)dt}{RV_i}\right)^2.$$

We then calculated the sample average of squared relative errors over 500 simulations. We have the average errors 0.0117, 0.0463, and 0.10751 for $m = 11,700, 1,170,$ and 390, respectively. As m increases, the average errors decrease. This result supports the theoretical findings in the realized volatility estimator literature (Jacod et al., 2009; Aït-Sahalia and Xiu, 2016).

Table 1 reports the mean squared errors (MSEs) of the QMLE and OLS estimates $\widehat{\theta^g}$ and $\widehat{\theta^*}$ with $n = 100, 200, 500$ and $m = 390, 1,170, 11,700$. In Table 1, MSEs usually decrease as the number of high-frequency observations or daily observations increases. When comparing the QMLE and OLS procedures, the OLS shows better performance. This may be because the OLS is based on the log-integrated volatility structure, which can better explain the ERGI model structure. This result supports the theoretical findings in Section 3.

To check the asymptotic normality of the GARCH parameters $(\omega^g, \gamma, \beta^g)$, we calculated the Z-statistics defined in Section 3.3. In Figures 2 and 3, we draw the standard normal QQ plots of the Z-statistics estimates of the QMLE and OLS procedures for $m = 11,700$ and $n = 100, 200, 500$. In Figures 2 and 3, we find that

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

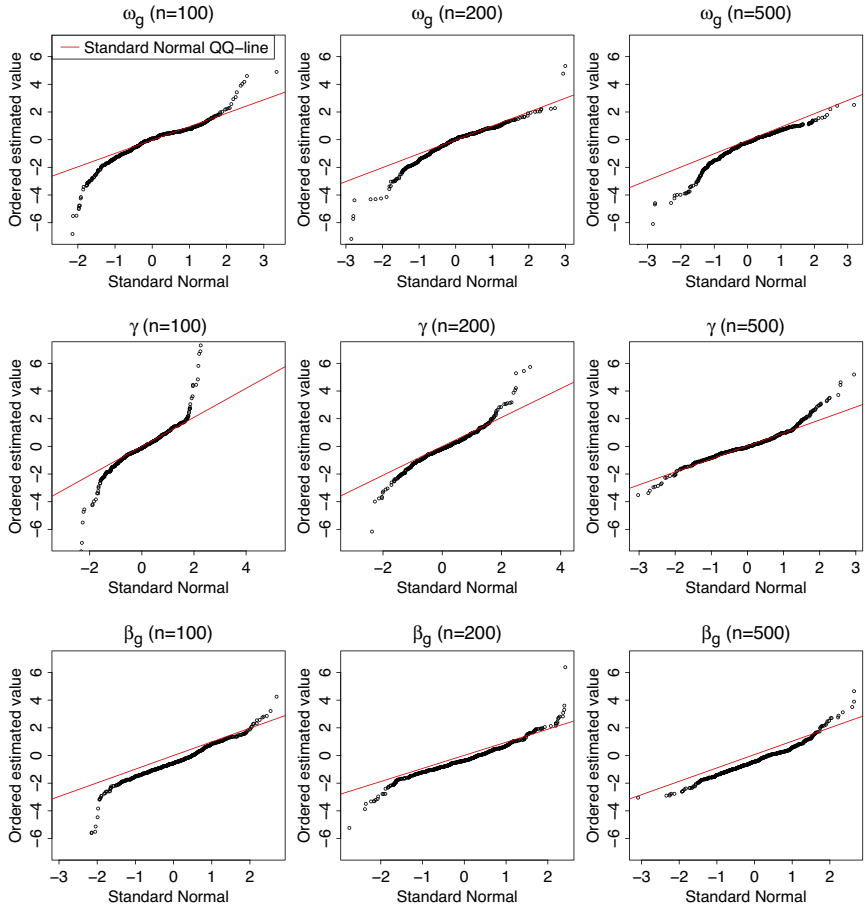

FIGURE 2. Standard normal QQ plots of the Z-statistics estimates of the QMLE estimators ω^g, γ, and β^g for $m = 11,700$ and $n = 100, 200, 500$. The red real line denotes the best linear fit line, which illustrates perfect standard normal distribution.

the Z-statistics usually become close to the standard normal distribution as the sample period increases. This result supports the theoretical findings in Section 3. Thus, based on the proposed Z-statistics, we can conduct hypothesis tests using the standard normal distribution.

We examined the out-of-sample performance of estimating the 1-day-ahead GARCH volatility $\exp(H_{n+1}(\theta_0^g))$. To estimate future GARCH volatility, we employed the proposed conditional ERGI estimator with QMLE (ERGI-QMLE) $\exp(\widehat{H}_{n+1}(\widehat{\theta^g}))$, the ERGI estimator with OLS (ERGI-OLS) $\exp(\widehat{h}_{n+1}(\widehat{\theta^*})) \times \frac{1}{n}\sum_{i=1}^{n}\exp(\log RV_i - \widehat{h}_i(\widehat{\theta^*}))$, realized GARCH volatility estimator (Hansen et al., 2012; Song et al., 2021), and PRV of the previous day. For example, the realized

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

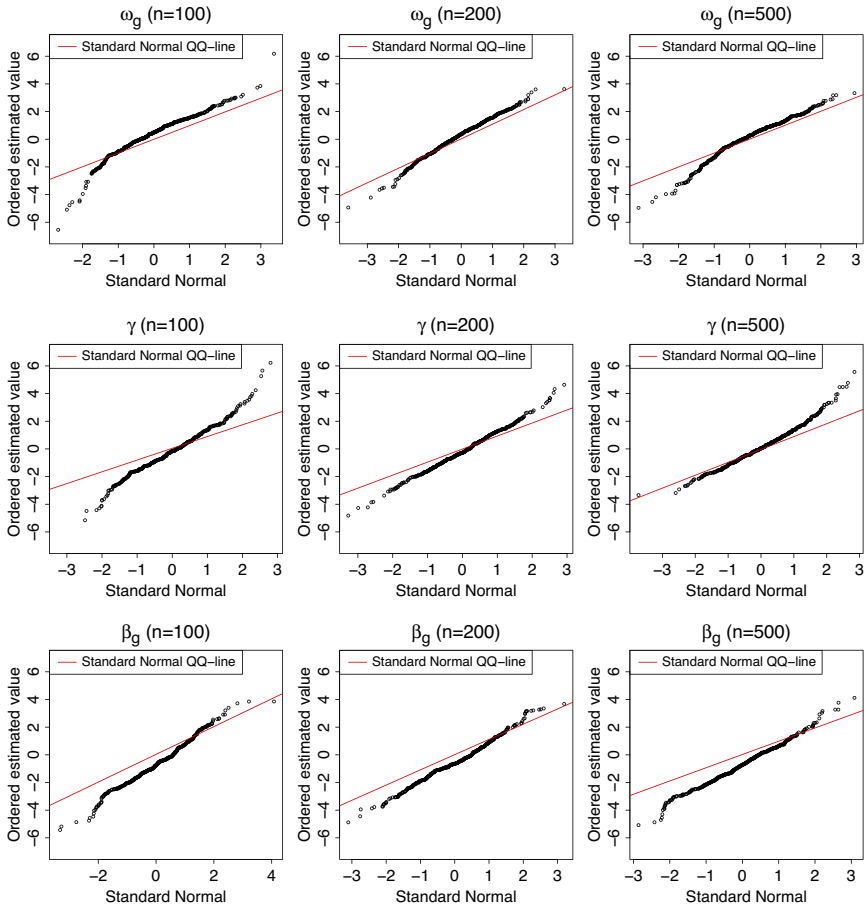

FIGURE 3. Standard normal QQ plots of the Z-statistics estimates of the OLS estimators ω^*, γ, and β^g for $m = 11,700$ and $n = 100, 200, 500$. The red real line denotes the best linear fit line, which illustrates perfect standard normal distribution.

GARCH volatility estimator is estimated based on the following conditional volatility:

$$h_n(\theta^g) = \omega + \gamma h_{n-1}(\theta^g) + \beta RV_{n-1}.$$

In other words, the realized GARCH volatility estimator has the usual linear GARCH structure with the realized volatilities. We measured the MSEs with the 1-day-ahead sample period over 500 samples as follows:

$$\frac{1}{500}\sum_{i=1}^{500}\left[\widehat{\text{var}}_{n+1,i} - \exp\left(H_{n+1,i}(\theta_0^g)\right)\right]^2,$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

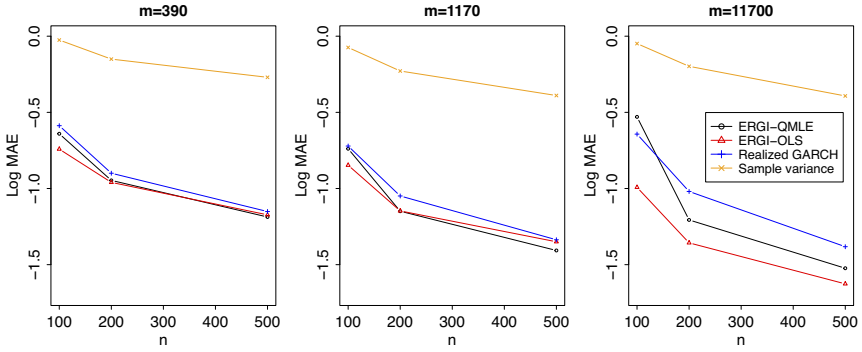

FIGURE 4. Log MSEs for the ERGI-QMLE, ERGI-OLS, realized GARCH, and PRV with $n = 100, 200, 500$, and $m = 390, 1,170, 11,700$.

where $\widehat{\text{var}}_{n+1,i}$ is one of the above future volatility estimators at the ith sample path given the available information at time n. Figure 4 depicts the MSEs for the ERGI-QMLE, ERGI-OLS, realized GARCH, and PRV against varying the numbers of low- and high-frequency observations, n and m. In Figure 4, we find that the ERGI models show the best performance. The interesting finding is that the realized GARCH model can also capture some volatility dynamics. This may be because even if the volatility dynamic structure is nonlinear, it could have some linear dynamics. Especially when the log-volatility quantities are small, by Taylor's expansion, the linear model can capture some nonlinear dynamics. However, by using only the linear structure, it cannot fully explain the nonlinear dynamic structure. From these results, we can conjecture that modeling appropriate dynamic structure helps account for market dynamics. When comparing the ERGI-QMLE with the ERGI-OLS, the ERGI-OLS shows better performance. This result is consistent with the parameter estimation result. That is, the OLS can better explain the ERGI model structure.

5. EMPIRICAL STUDY

We applied the proposed ERGI model to real trading high-frequency data. We obtained the top 50 trading volume assets intraday data from January 2010 to December 2016 from the trade and quote (TAQ) database in the Wharton Research Data Services system. We used the log prices and employed the jump robust PRV estimation procedure defined in Section 4 to estimate open-to-close integrated volatility. In the empirical study, we chose the tuning parameter c_τ as 10 times the sample standard deviation of pre-averaged prices $m^{1/4}\bar{Y}(t_{d,k})$. To check the accuracy of the PRV estimator, we calculated standard errors (SE) as follows. We first calculated the asymptotic variance, proposed by Aït-Sahalia and Xiu (2016), and divided the square root of the asymptotic variance estimator by the square root of the number of high-frequency observations. We report the data summary in

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

TABLE 2. Averages of the number of high-frequency observations, realized volatility (RV), SE of the realized volatility estimator (SE), and jump variation (JV).

Stock	# of obs	RV$\times 10^4$	SE$\times 10^4$	JV$\times 10^4$	Stock	# of obs	RV$\times 10^4$	SE$\times 10^4$	JV$\times 10^4$
PG	16,912.4	0.5637	0.0414	0.4000	MO	23,667.5	0.6886	0.0075	0.2074
HBAN	15,941.4	2.7678	0.0183	0.3550	QCOM	40,986.1	1.3487	0.0059	0.2306
FCX	37,196.1	5.4242	0.0222	0.8249	MRK	36,710.1	0.9756	0.0198	0.2643
MRO	31,241.5	3.8938	0.0203	0.6569	GILD	43,938.5	1.8619	0.0132	0.4030
ORCL	44,489.4	1.3062	0.0062	0.1915	DAL	35,653.1	4.2791	0.0219	0.6506
AMD	22,373.9	5.9256	0.0345	0.9735	LUV	23,756.9	2.3442	0.0120	0.6057
AMAT	33,556.6	2.0554	0.0079	0.3043	T	38,606.7	0.7138	0.0037	0.1288
XRX	18,877.3	2.3137	0.0172	1.4115	CSCO	40,328.4	1.2632	0.0079	0.2128
WFC	44,023.8	1.4846	0.0073	0.3143	DIS	16,682.3	1.0435	0.0108	0.1842
NFLX	30,613.5	5.4658	0.0293	0.5557	NVDA	27,743.1	3.2549	0.0151	0.4210
F	34,610.1	2.2033	0.0177	0.4130	SLB	31,761.7	2.0279	0.0094	0.4409
GE	46,327.4	1.2444	0.0153	0.2523	BMY	27,565.6	1.1346	0.0157	0.2655
INTC	45,515.7	1.4031	0.0076	0.2464	ATVI	25,432.0	2.0102	0.0100	0.3804
XOM	45,802.3	0.9642	0.0072	0.1490	MU	38,907.5	5.1948	0.0301	0.9009
RF	19,662.3	3.7024	0.0236	0.4208	JPM	40,898.6	1.6448	0.0729	0.4199
DOW	28,093.0	1.9877	0.0190	0.3006	CVX	32,592.7	1.1472	0.0067	0.2205
NEM	29,263.8	3.3862	0.0118	0.4225	MSFT	61,219.5	1.2213	0.0035	0.1855
CSX	16,106.8	1.6852	0.0153	0.4427	BAC	63,492.7	2.4282	0.0098	0.3718
TXN	25,727.4	1.3822	0.0072	0.3163	WMT	30,549.0	0.6412	0.0033	0.1288
JNJ	31,877.1	0.5137	0.0062	0.1689	WMB	26,970.9	3.8400	0.0394	0.8323
VZ	32,309.9	0.7749	0.0185	0.2155	AAPL	90,386.9	1.4419	0.0118	0.3017
HST	21,239.2	2.2687	0.0135	0.2591	BSX	24,163.9	2.3119	0.0145	0.4101
MGM	16,642.6	4.5845	0.0309	0.6648	PFE	43,600.5	1.0821	0.0160	0.2296
KO	37,024.6	0.6041	0.0034	0.1091	HAL	39,739.1	3.2834	0.0136	0.4913
SCHW	27,159.3	2.0437	0.1963	0.3522	GLW	25,151.0	1.8426	0.0155	0.3186

Table 2. The number of high-frequency data is from 16,000 to 90,000 on average, and we find that the proportion of the jump variation is approximately 8%–40% of the total variation on average. The SE is less than 10% of the realized volatility.

We first estimated the model parameters using the last 1,000 days of data. From the estimated model parameters, we obtained the following conditional expected volatilities for each asset:

$$\widehat{H}_{n+1}(\widehat{\theta^g}) = \widehat{\omega}^g + \widehat{\gamma}\widehat{H}_n(\widehat{\theta^g}) + \widehat{\alpha}^g \log(RV_n),$$
$$\widehat{h}_{n+1}(\widehat{\theta^*}) = \widehat{\omega}^* + \widehat{\gamma}\widehat{H}_n(\widehat{\theta^*}) + \widehat{\alpha}^g \log(RV_n).$$

Table 3 reports the estimation results. From Table 3, we find that dynamic structures can be explained by the past log PRV, and the coefficients of realized

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

TABLE 3. ERGI model estimation results for the QMLE and OLS. In the parentheses, we report the p-values.

	QMLE			OLS		
Stock	ω^g	γ	β^g	ω^*	γ	β^g
PG	−1.18 (0.00)	0.33 (0.00)	0.55 (0.00)	−0.50 (0.00)	0.48 (0.00)	0.47 (0.00)
HBAN	−1.06 (0.00)	0.35 (0.00)	0.52 (0.00)	−1.10 (0.00)	0.38 (0.00)	0.50 (0.00)
FCX	−0.07 (0.00)	0.48 (0.00)	0.50 (0.00)	−0.14 (0.00)	0.50 (0.00)	0.48 (0.00)
MRO	−0.12 (0.00)	0.39 (0.00)	0.59 (0.00)	−0.16 (0.00)	0.43 (0.00)	0.55 (0.00)
ORCL	−1.61 (0.00)	0.18 (0.00)	0.64 (0.00)	−0.53 (0.00)	0.44 (0.00)	0.51 (0.00)
AMD	−0.50 (0.00)	0.50 (0.00)	0.43 (0.00)	−0.74 (0.00)	0.47 (0.00)	0.43 (0.00)
AMAT	−1.83 (0.26)	0.19 (0.00)	0.59 (0.00)	−0.82 (0.00)	0.46 (0.00)	0.44 (0.00)
XRX	−0.57 (0.00)	0.48 (0.00)	0.45 (0.00)	−1.28 (0.00)	0.42 (0.00)	0.44 (0.00)
WFC	−1.17 (0.00)	0.24 (0.00)	0.63 (0.00)	−0.67 (0.00)	0.42 (0.00)	0.51 (0.00)
NFLX	−0.70 (0.05)	0.29 (0.00)	0.61 (0.02)	−0.72 (0.00)	0.38 (0.00)	0.53 (0.00)
F	−1.56 (0.00)	0.26 (0.03)	0.56 (0.00)	−0.77 (0.00)	0.47 (0.00)	0.45 (0.00)
GE	−1.40 (0.00)	0.19 (0.11)	0.66 (0.00)	−0.57 (0.00)	0.46 (0.00)	0.48 (0.00)
INTC	−1.14 (0.00)	0.27 (0.00)	0.60 (0.00)	−1.31 (0.00)	0.31 (0.00)	0.55 (0.00)
XOM	−0.62 (0.00)	0.34 (0.00)	0.59 (0.00)	−0.67 (0.00)	0.38 (0.00)	0.55 (0.00)
RF	−0.98 (0.00)	0.42 (0.00)	0.46 (0.00)	−0.86 (0.00)	0.46 (0.00)	0.44 (0.00)
DOW	−1.06 (0.00)	0.32 (0.00)	0.56 (0.00)	−0.84 (0.00)	0.42 (0.00)	0.49 (0.00)
NEM	−0.74 (0.00)	0.35 (0.00)	0.55 (0.00)	−0.72 (0.00)	0.38 (0.00)	0.53 (0.00)
CSX	−0.83 (0.00)	0.32 (0.00)	0.58 (0.00)	−0.87 (0.00)	0.34 (0.00)	0.57 (0.00)
TXN	−1.31 (0.00)	0.24 (0.00)	0.62 (0.04)	−0.52 (0.00)	0.43 (0.00)	0.52 (0.00)
JNJ	−1.03 (0.00)	0.38 (0.00)	0.51 (0.00)	−0.50 (0.00)	0.50 (0.00)	0.45 (0.00)
VZ	−1.91 (0.00)	0.21 (0.00)	0.58 (0.00)	−0.58 (0.00)	0.55 (0.00)	0.39 (0.00)
HST	−0.62 (0.10)	0.47 (0.00)	0.46 (0.00)	−0.53 (0.00)	0.50 (0.00)	0.44 (0.00)
MGM	−1.04 (0.00)	0.32 (0.00)	0.55 (0.00)	−0.72 (0.00)	0.38 (0.00)	0.53 (0.00)
KO	−1.46 (0.00)	0.26 (0.00)	0.59 (0.00)	−0.70 (0.00)	0.47 (0.00)	0.46 (0.00)
SCHW	−1.14 (0.00)	0.33 (0.00)	0.53 (0.00)	−0.68 (0.00)	0.46 (0.00)	0.46 (0.00)
MO	−1.37 (0.00)	0.33 (0.00)	0.52 (0.00)	−0.62 (0.00)	0.49 (0.00)	0.44 (0.00)
QCOM	−1.67 (0.00)	0.22 (0.00)	0.59 (0.00)	−0.64 (0.00)	0.51 (0.00)	0.42 (0.00)
MRK	−0.99 (0.00)	0.33 (0.00)	0.56 (0.00)	−0.47 (0.00)	0.32 (0.00)	0.63 (0.00)
GILD	−0.86 (0.00)	0.33 (0.00)	0.56 (0.00)	−0.73 (0.00)	0.38 (0.00)	0.54 (0.00)
DAL	−1.91 (0.00)	0.18 (0.02)	0.57 (0.00)	−0.76 (0.00)	0.50 (0.00)	0.41 (0.00)
LUV	−1.45 (0.00)	0.33 (0.00)	0.49 (0.00)	−1.07 (0.00)	0.43 (0.00)	0.45 (0.00)
T	−1.98 (0.01)	0.29 (0.00)	0.51 (0.00)	−0.67 (0.00)	0.53 (0.00)	0.40 (0.00)
CSCO	−1.83 (0.00)	0.18 (0.00)	0.62 (0.00)	−0.79 (0.00)	0.46 (0.00)	0.46 (0.00)
DIS	−1.41(0.00)	0.23 (0.00)	0.62 (0.00)	−1.41 (0.00)	0.30 (0.00)	0.56 (0.00)
NVDA	−1.31 (0.02)	0.29 (0.00)	0.55 (0.00)	−0.94 (0.00)	0.44 (0.00)	0.45 (0.00)

(Continues)

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

TABLE 3. Continued

	QMLE			OLS		
Stock	ω^g	γ	β^g	ω^*	γ	β^g
SLB	−0.63 (0.00)	0.37 (0.00)	0.56 (0.00)	−0.58 (0.00)	0.42 (0.00)	0.51 (0.00)
BMY	−1.72 (0.00)	0.23 (0.00)	0.50 (0.00)	−0.82 (0.00)	0.49 (0.00)	0.42 (0.00)
ATVI	−0.83 (0.00)	0.43 (0.00)	0.47 (0.00)	−1.01 (0.00)	0.46 (0.00)	0.42 (0.00)
MU	−1.10 (0.00)	0.34 (0.07)	0.51 (0.00)	−0.77 (0.00)	0.42 (0.00)	0.48 (0.00)
JPM	−1.42 (0.00)	0.16 (0.00)	0.68 (0.00)	−0.55 (0.00)	0.42 (0.00)	0.52 (0.00)
CVX	−0.39 (0.00)	0.38 (0.00)	0.57 (0.00)	−0.51 (0.00)	0.40 (0.00)	0.55 (0.00)
MSFT	−1.29 (0.00)	0.30 (0.00)	0.55 (0.00)	−1.19 (0.00)	0.38 (0.00)	0.49 (0.00)
BAC	−1.26 (0.00)	0.27 (0.00)	0.58 (0.00)	−0.55 (0.00)	0.41 (0.00)	0.53 (0.00)
WMT	−0.71 (0.00)	0.52 (0.00)	0.40 (0.00)	−0.75 (0.00)	0.53 (0.00)	0.40 (0.00)
WMB	−0.09 (0.00)	0.41 (0.00)	0.57 (0.00)	−0.21 (0.00)	0.52 (0.00)	0.46 (0.00)
AAPL	−2.01 (0.00)	0.09 (0.00)	0.68 (0.00)	−1.67 (0.00)	0.20 (0.00)	0.62 (0.00)
BSX	−1.40 (0.00)	0.29 (0.00)	0.54 (0.00)	−0.63 (0.00)	0.50 (0.00)	0.43 (0.00)
PFE	−0.89 (0.00)	0.33 (0.00)	0.57 (0.00)	−0.45 (0.00)	0.43 (0.00)	0.53 (0.00)
HAL	−0.70 (0.00)	0.31 (0.00)	0.60 (0.04)	−0.72 (0.00)	0.38 (0.00)	0.54 (0.00)
GLW	−1.92 (0.00)	0.18 (0.00)	0.60 (0.00)	−1.46 (0.00)	0.34 (0.00)	0.50 (0.00)

volatilities are statistically significant. Thus, the proposed exponential model is valid.

For comparison, we employed the realized GARCH (Hansen et al., 2012; Song et al., 2021), UGARCH (Kim and Wang, 2016), and HAR (Corsi, 2009) models. We used the jump robust PRV as the proxy. Thus, the above models account for the volatility of the continuous part of the log-price process. For example, the UGARCH model employs the squared open-to-close log returns as the risk factor and the jump robust PRV as the proxy. The realized GARCH and HAR models employ the jump robust PRV as the risk factor and proxy. To measure the performance of the volatility, we used the mean squared prediction errors (MSPE) and QLIKE (Patton, 2011) as follows:

$$MSPE = \frac{1}{n}\sum_{i=1}^{n}(Vol_i - RV_i)^2 \quad \text{and} \quad QLIKE = \frac{1}{n}\sum_{i=1}^{n}\log Vol_i + \frac{RV_i}{Vol_i},$$

where Vol_i is one of the ERGI-QMLE, ERGI-OLS, HAR, realized GARCH, and UGARCH. We used RV_i as the nonparametric daily volatility estimator. Furthermore, we calculated the out-of-sample R-square (OSR; Campbell and Thompson, 2008) as follows:

$$OSR = 1 - \frac{\sum_{i=1}^{n}\left(RV_i - Vol_i^*\right)^2}{\sum_{i=1}^{n}\left(RV_i - Vol_i\right)^2},$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Table 4. Average rank of MSPEs and QLIKEs for the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period. In the parentheses, we report the number of those ranked first among competitors.

	ERGI-Q	ERGI-O	Real	HAR	UGARCH
Period 1					
MSPE	2.12 (10)	1.78 (25)	3.10 (4)	3.00 (11)	5.00 (0)
QLIKE	2.22 (9)	2.02 (20)	3.12 (5)	2.64 (16)	5.00 (0)
Period 2					
MSPE	1.60 (26)	1.72 (20)	3.48 (1)	3.54 (3)	4.66 (0)
QLIKE	1.82 (19)	2.20 (15)	2.58 (11)	3.40 (5)	5.00 (0)
Whole period					
MSPE	1.70 (25)	1.72 (19)	3.50 (1)	3.36 (5)	4.72 (0)
QLIKE	1.70 (24)	2.06 (14)	2.86 (4)	3.38 (8)	5.00 (0)

where Vol_i^* is the proposed ERGI-QMLE model, and Vol_i is one of the ERGI-OLS, realized GARCH, HAR, and unified GARCH-Itô models and sample mean of the in-sample RV_i's. We predicted the 1-day-ahead conditional expected volatility by the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models using the in-sample period data. We fixed the in-sample period as 500 days and used the rolling window scheme to estimate the parameters. The number of out-of-sample observations was 1,262. To check the period dependency, we split the period into two equal parts, denoted by Period 1 and Period 2. Table 4 reports the average rank and the number of those ranked first of MSPEs and QLIKEs for the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period over the 50 assets. Figure 5 depicts the relative MSPE and QLIKE for the ERGI-OLS, realized GARCH, HAR, and UGARCH models with respect to the ERGI-QMLE model for Period 1, Period 2, and the whole period. Figure 6 draws the OSR for the ERGI-QMLE model with respect to the ERGI-OLS, realized GARCH, HAR, and UGARCH models, and sample mean for Period 1, Period 2, and the whole period. From Table 4 and Figures 5 and 6, we find that the realized volatility-based models, such as the ERGI, realized GARCH, and HAR models, perform better than the UGARCH model, which incorporates the squared open-to-close returns as the innovation. That is, incorporating the realized volatility helps account for the volatility dynamics. When comparing the realized volatility-based models, the proposed ERGI model shows the best performance. For the ERGI models, the QMLE and OLS estimation procedures show similar performance. From this result, we can conjecture that the nonlinear exponential form with realized volatilities helps explain the market dynamics.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

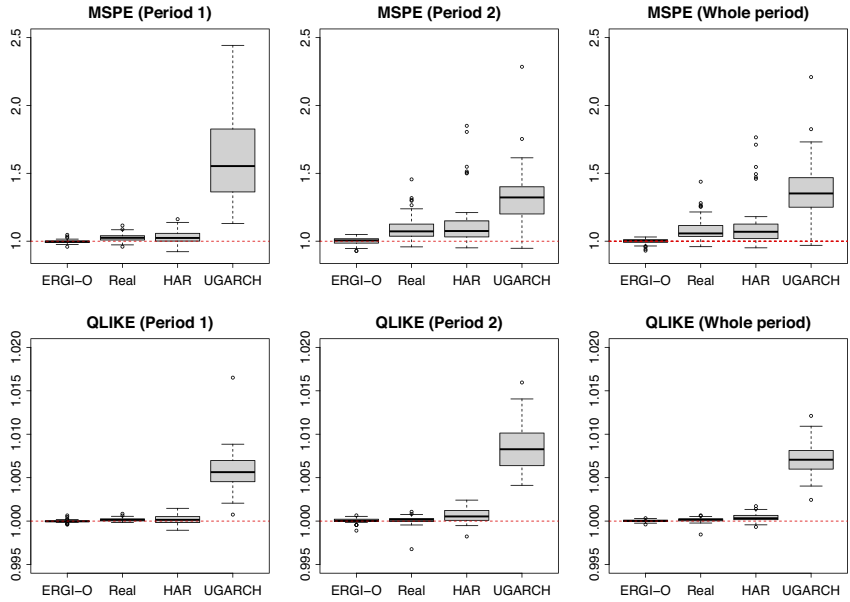

FIGURE 5. Box plots of relative MSPE and QLIKE for the realized GARCH, HAR, and UGARCH models with respect to the ERGI model for Period 1, Period 2, and the whole period.

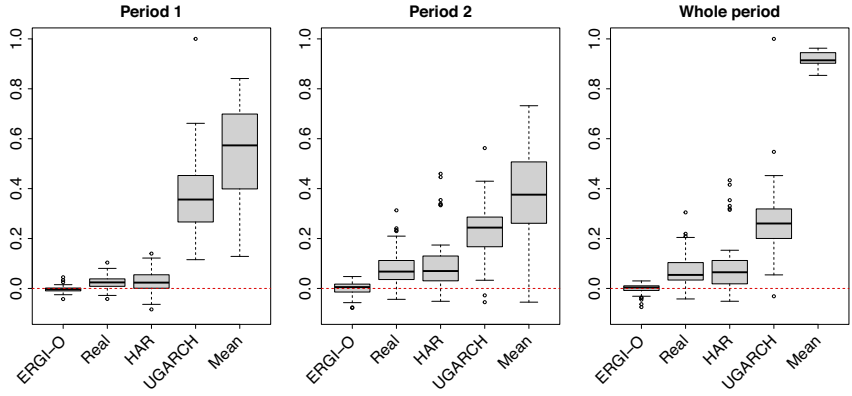

FIGURE 6. Box plots of OSR for the ERGI-QMLE model with respect to the ERGI-OLS, realized GARCH, HAR, and UGARCH models, and sample mean for Period 1, Period 2, and the whole period.

To further compare the predictive accuracy among the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models, we conducted Diebold–Mariano (DM) tests (Diebold and Mariano, 2002) as follows. We first calculated the residuals for the four models as follows:

$$e_i = RV_i - Vol_i,$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

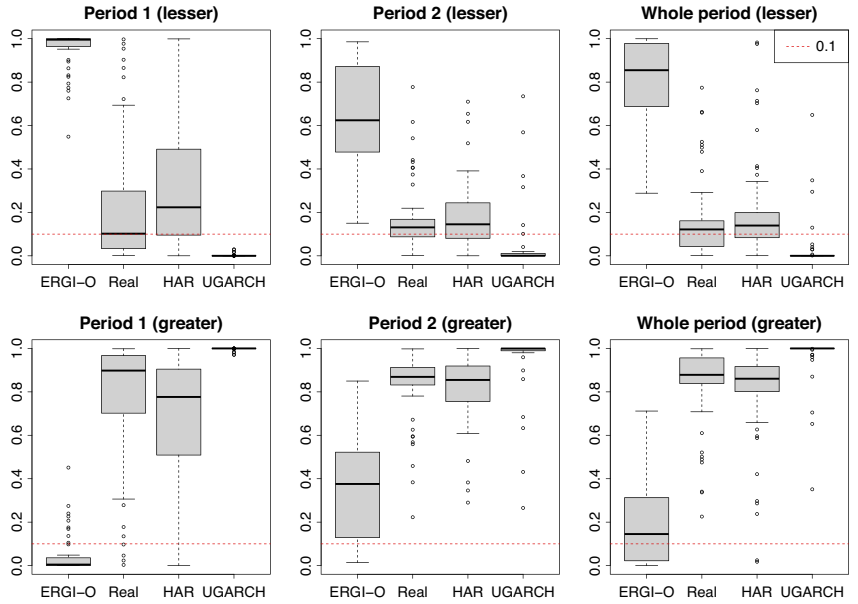

FIGURE 7. Box plots of the p-values of the lesser and greater DM tests for the ERGI-QMLE versus one of the ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period.

where Vol_i is one of the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models and RV_i is the nonparametric realized volatility. We defined

$$d_i = e_i^{*2} - e_i^2,$$

where e_i^* is the residuals from the ERGI-QMLE model and e_i is the residuals from one of the ERGI-OLS, realized GARCH, HAR, and UGARCH models. We then conducted hypothesis tests for

$$H_0' : \mathbb{E}[d_i] = 0 \quad \text{vs.} \quad H_1 : \mathbb{E}[d_i] < 0 \text{ (or } \mathbb{E}[d_i] > 0).$$

The first alternative statement ($\mathbb{E}[d_i] < 0$) is to test whether the ERGI-QMLE model is better, whereas the second alternative statement ($\mathbb{E}[d_i] > 0$) is to test whether another model is better than the ERGI-QMLE model. We call these tests "lesser" and "greater," respectively. Figure 7 depicts the box plots of the p-values of the lesser and greater DM tests for the ERGI-QMLE model versus one of the ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period. In Figure 7, the lesser tests show that the p-values of 22, 16, and 46 assets for the realized GARCH, HAR, and UGARCH models, respectively, are less than 10% over the whole period. In contrast, the greater tests indicate that a couple of assets for the HAR model have significant p-values

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

TABLE 5. Average rank of the first-order autocorrelation for the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period. In the parentheses, we report the number of those ranked first among competitors.

	ERGI-Q	ERGI-O	Real	HAR	UGARCH
Period 1	1.72 (18)	2 (23)	2.52 (8)	3.78 (1)	4.98 (0)
Period 2	2.24 (12)	2.58 (13)	2.88 (14)	2.76 (10)	4.54 (1)
Whole period	2 (17)	2.4 (14)	2.78 (12)	3.18 (6)	4.64 (1)

over the whole period. When comparing the estimation methods, the ERGI-OLS shows better performance. From these results, although the ERGI does not give significant better predictive accuracy for all assets, we can conclude that for most assets, the ERGI is at least not worse than other models, and, for some assets, the ERGI shows significantly better performance than the other models.

We also conducted the model confidence set (MCS) procedure (Hansen, Lunde, and Nason, 2011). We used the R package provided by Bernardi and Catania (2018) and followed the default setup with the absolute error loss. For example, we chose the confidence level of the test as 20% and used the Tmax statistics that are described in Section 3.1.2 of Hansen et al. (2011). According to the MCS procedure, we chose most of the estimation methods as the superior set. Specifically, the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models are included 50, 49, 48, 50, and 48 times, respectively, in the superior set. The ERGI-QMLE and HAR models show the best performance. However, it is hard to conclude that these models have significantly different performance in terms of the MCS. This result is different from the pairwise test such as the DM test. One possible explanation is that the MCS is based on the maximum value, which may reduce the power of the test.

To check the volatility persistence of the nonparametric volatility, we studied the residuals between the nonparametric volatility and estimated conditional volatilities, $Vol_i - RV_i$, where Vol_i is the predicted 1-day-ahead conditional expected volatility by the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models using the in-sample period data. We then checked their autocorrelations. Table 5 reports the average rank and number of the first rank of the first-order autocorrelation for the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period over the 50 assets. Figure 8 provides the box plots of the first-order autocorrelation for the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period over the 50 assets. From Table 5 and Figure 8, we find that the ERGI-QMLE and ERGI-OGI models have relatively small autocorrelations. That is, the ERGI models can reduce the volatility persistence. These numerical results provide evidence to conclude that the nonlinear exponential autoregressive structure helps explain the market dynamics in the volatility time series.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

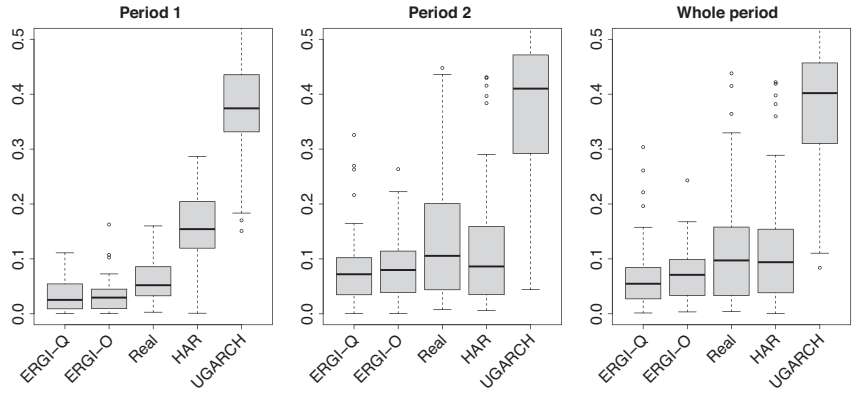

FIGURE 8. Box plots of the first-order autocorrelation for the residuals of the ERGI-QMLE, ERGI-OLS, realized GARCH, HAR, and UGARCH models for Period 1, Period 2, and the whole period.

5.1. Whole-Day Market Dynamics

Since high-frequency data are available only during trading hours, in the previous section, we only considered the open-to-close period. However, Taylor (2007) showed that overnight information is important for evaluating risk management models, so the volatility measured by the open-to-close high-frequency observations may significantly undervalue their risk. See also Martens (2002), Hansen and Lunde (2005), Andersen, Bollerslev, and Huang (2011), Tseng, Lai, and Lin (2012), Todorova and Souček (2014), and Kim, Shin, and Wang (2022). In this section, we investigated volatility dynamics for the whole day.

Recently, to explain whole-day market dynamics, Kim et al. (2022) introduced the overnight GARCH-Itô (OGI) model, which can accommodate two different instantaneous volatility processes for the open-to-close and close-to-open periods. For example, the whole-day (open-to-open) volatility has the following conditional volatility:

$$h_n(\theta) = \omega + \gamma h_{n-1}(\theta) + \alpha \int_{n-2}^{n-2+\lambda} \sigma_t^2(\theta)dt + \beta(X_{n-1} - X_{n-2+\lambda})^2,$$

where λ is the proportion of the trading hour ($\lambda = 6.5/24$). That is, the OGI model has two different risk factors, such as the open-to-close integrated volatility and squared close-to-open return. Similarly, we can propose the overnight exponential realized GARCH-Itô (OERGI) model as follows:

$$H_n(\theta) = \omega + \gamma H_{n-1}(\theta) + \alpha \log\left(\int_{n-2}^{n-2+\lambda} \sigma_t^2(\theta)dt\right) + \beta|X_{n-1} - X_{n-2+\lambda}|,$$

where $\exp(H_n(\theta))$ is the conditional expected volatility. Since the overnight return can be zero, we used the absolute value of the overnight return as the overnight

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

TABLE 6. Average rank of MSPEs and QLIKEs for the OERGI-QMLE, OERGI-OLS, adj-ERGI-QMLE, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period. In the parentheses, we report the number of those ranked first among competitors.

	MSPE							
	OERGI-Q	OERGI-O	A-ERGI-Q	A-ERGI-O	OGI	A-Real	A-HAR	A-UGARCH
Period 1	3.42 (9)	4.24 (15)	4 (1)	3.78 (4)	4.26 (13)	4.74 (1)	4.34 (5)	7.22 (2)
Period 2	5.04 (8)	3.88 (8)	3 (8)	3.46 (3)	4.76 (13)	4.58 (6)	4.58 (3)	6.7 (1)
Whole	4.64 (11)	3.7 (12)	3.26 (4)	3.58 (4)	4.94 (11)	4.64 (3)	4.42 (5)	6.82 (0)
	QLIKE							
	OERGI-Q	OERGI-O	A-ERGI-Q	A-ERGI-O	OGI	A-Real	A-HAR	A-UGARCH
Period 1	3.28 (15)	5.04 (7)	3.98 (2)	3.82 (2)	3.44 (16)	4.72 (2)	4.66 (2)	7.06 (4)
Period 2	3.4 (6)	5.28 (4)	3.8 (2)	4.06 (2)	3.62 (20)	4.36 (5)	4.18 (10)	7.3 (1)
Whole	3.26 (10)	5.16 (4)	3.8 (2)	4.12 (1)	3.58 (20)	4.58 (2)	4.46 (7)	7.04 (4)

risk factor instead of the log transformation. Thus, the OERGI model has the open-to-close log-integrated volatility and absolute of value of the overnight return as the risk factors. For the models considered in the previous section, to match the magnitude, we magnified the estimator by multiplying it with $(1 + mean[OV/RV])$, where OV is the overnight return squares and RV is the open-to-close realized volatility. We call them the adj-ERGI-QMLE, adj-ERGI-OLS, adj-realized GARCH, adj-HAR, and adj-UGARCH models. To estimate the OGI and OERGI models, we used the realized volatility estimator plus squared overnight return as the proxy and employed the QMLE method with the Gaussian likelihood function for the OGI and OERGI models and the OLS method for the OERGI model. We call them the OGI, OERGI-QMLE, and OERGI-OLS models.

We first compared the performances based on the MSPE, QLIKE, and OSR. Table 6 reports the average rank and the number of those ranked first of MSPEs and QLIKEs for the OERGI-QMLE, OERGI-OLS, adj-ERGI-QMLE, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period over the 50 assets. Figure 9 depicts the relative MSPE and QLIKE for the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models with respect to the adj-ERGI-QMLE model for Period 1, Period 2, and the whole period. Figure 10 draws the OSR for the adj-ERGI-QMLE model with respect to the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period. From Table 6 and Figures 9 and 10, we find that the ERGI-based models usually show better performance than others. This may be because the ERGI-based models can capture the nonlinear dynamics. When comparing the ERGI models, the adj-ERGI-QMLE model shows stable results, but it is hard to conclude that one model dominates the others.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

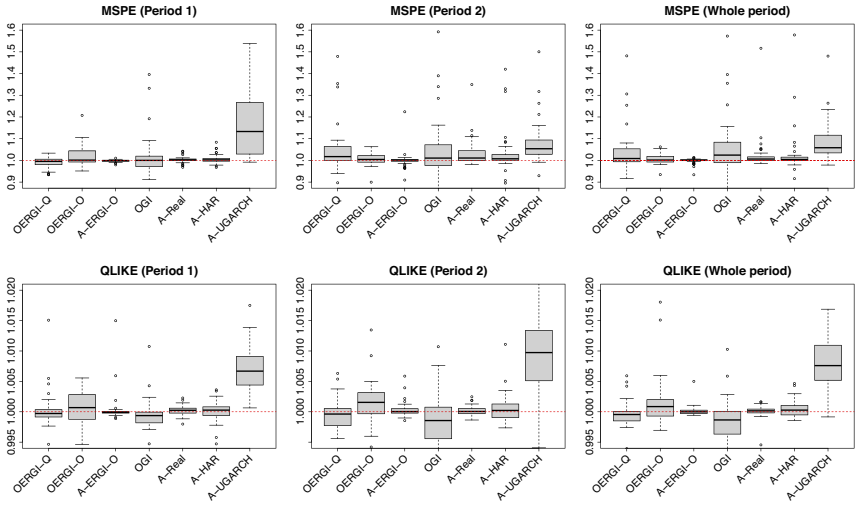

FIGURE 9. Box plots of relative MSPE and QLIKE for the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models with respect to the adj-ERGI-QMLE model for Period 1, Period 2, and the whole period.

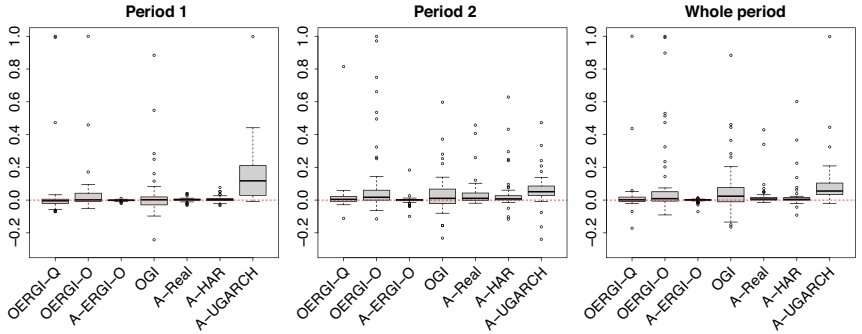

FIGURE 10. Box plots of OSR for the adj-ERGI-QMLE model with respect to the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period.

We conducted Diebold–Mariano tests. Figure 7 depicts the box plots of the p-values of the lesser and greater DM tests for the adj-ERGI-QMLE model versus one of the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period. In Figure 7, the lesser tests show that the p-values of 10, 12, 3, 11, 22, 42, and 29 assets for the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models, respectively, are less than

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

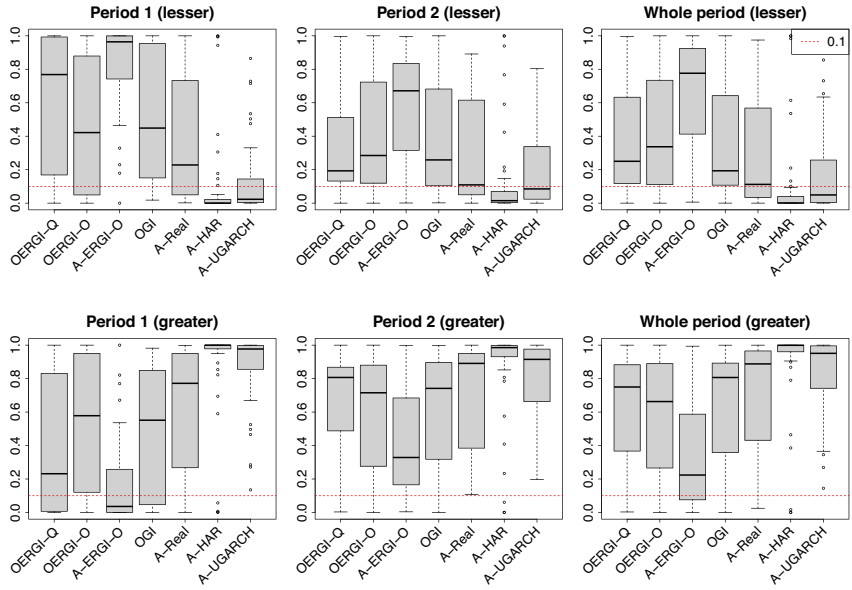

FIGURE 11. Box plots of the p-values of the lesser and greater DM tests for the adj-ERGI-QMLE model versus one of the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period.

10% over the whole period. In contrast, the greater tests show that the p-values of 4, 7, 17, 5, 2, 3, and 0 assets for the OERGI-QMLE, OERGI-OLS, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models, respectively, are less than 10% over the whole period. From these results, although the ERGI-based models do not give significantly better predictive accuracy for all assets, we can conclude that for most assets, the ERGI-based models are at least not worse than the other models, and, for some assets, the ERGI-based models show significantly better performance than the other models.

Finally, we checked the volatility persistence of the nonparametric volatility. Figure 12 provides the box plots of the first-order autocorrelation for the OERGI-QMLE, OERGI-OLS, adj-ERGI-QMLE, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period over the 50 assets. From Figure 12, we cannot find a significantly better model. This may be because as considering the whole-day market dynamics, the proxy is noisier. To sum up, for the whole-day market dynamics, we find some evidence that the log-transformation form can improve prediction performance. However, compared with the open-to-close volatility analysis, the result is less significant. This may be because the proposed model cannot fully explain the nonlinear whole-day volatility dynamics. Thus, it is important to develop a

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

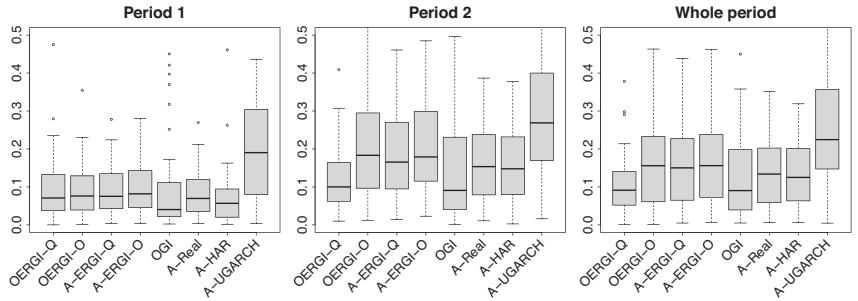

FIGURE 12. Box plots of the first-order autocorrelation for the residuals of the OERGI-QMLE, OERGI-OLS, adj-ERGI-QMLE, adj-ERGI-OLS, OGI, adj-realized GARCH, adj-HAR, and adj-UGARCH models for Period 1, Period 2, and the whole period.

nonlinear dynamic model which can account for the whole-day volatility nonlinear dynamics. We leave this for a future study.

6. CONCLUSIONS

In this paper, we propose a novel jump diffusion process to model the nonlinear autoregressive structure of the realized volatility. We employ the exponential GARCH structure. By introducing a continuous instantaneous volatility process whose integrated volatility follows the exponential realized GARCH structure, we fill the gap between the empirical discrete-time nonlinear volatility model with the realized volatility and high-frequency-based continuous-time diffusion process. That is, this paper provides a rigorous mathematical background to understand the exponential realized GARCH structure. To estimate the model parameters, we propose the quasi-maximum likelihood estimation procedure and ordinary least-squares estimation and establish their asymptotic properties. From the empirical study, we find the benefits of incorporating the nonlinear exponential realized GARCH model.

In this paper, we focus on the continuous part of log-return processes for the open-to-close period. However, it is important and interesting to study dynamic structures of the jump variation and close-to-open returns. We leave this for a future study.

7. PROOFS

7.1. Proof of Theorem 1

LEMMA 1. *Under the ERGI model in Definition 1, we have, for* $t \in (n-1, n]$,

$$\frac{1}{t-n+1}\int_{n-1}^{t} \sigma_s^2(\theta)ds = \sigma_{n-1}^2(\theta)e^{\int_{n-1}^{t} b_s(\theta)ds} \text{ a.s.} \tag{7.1}$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Proof of Lemma 1. We have

$$\begin{aligned} d((t-[t])\bar{\sigma}_t^2(\theta)) &= (t-[t])d\bar{\sigma}_t^2(\theta)+\bar{\sigma}_t^2(\theta)dt \\ &= \sigma_t^2(\theta)dt \\ &= \bar{\sigma}_t^2(\theta)\,(1+(t-[t])b_t(\theta))\,dt \text{ a.s.} \end{aligned}$$

Thus, we have

$$\begin{aligned} (t-[t])d\bar{\sigma}_t^2(\theta) &= \bar{\sigma}_t^2(\theta)\,(1+(t-[t])b_t(\theta))\,dt-\bar{\sigma}_t^2(\theta)dt \\ &= (t-[t])\bar{\sigma}_t^2(\theta)b_t(\theta)dt \text{ a.s.} \end{aligned}$$

This implies that

$$d\bar{\sigma}_t^2(\theta) = \bar{\sigma}_t^2(\theta)b_t(\theta)dt \text{ a.s.}$$

and

$$\frac{1}{t-n+1}\int_{n-1}^{t}\sigma_s^2(\theta)ds = \sigma_{n-1}^2(\theta)e^{\int_{n-1}^{t}b_s(\theta)ds} \quad \text{a.s.} \quad \text{for } t\in(n-1,n].$$

□

Proof of Theorem 1. First, we consider (a). By Itô's lemma, we have

$$\begin{aligned} R(k) &= \int_{n-1}^{n}\frac{(n-t)^k}{k!}b_t(\theta)dt \\ &= \frac{\omega}{(k+2)!}+\nu\left(\frac{1}{(k+2)!}-\frac{2}{(k+3)!}\right) \\ &\quad+\left\{\frac{1}{(k+1)!}+\frac{\gamma-1}{(k+2)!}\right\}b_{n-1}(\theta) \\ &\quad-\left\{\frac{\beta}{(k+1)!}+\frac{\beta^*-\beta}{(k+2)!}-\frac{2\beta^*}{(k+3)!}\right\}\log\sigma_{n-1}^2(\theta) \\ &\quad+2\nu\int_{n}^{n-1}\frac{(n-t)^{k+2}}{k!\,(k+2)}Z_t dW_t+\beta\int_{n-1}^{n}\frac{(n-t)^k}{k!}\log\bar{\sigma}_t^2(\theta)dt \quad \text{a.s.,} \end{aligned}$$

and, by Lemma 1, we have

$$\begin{aligned} &\int_{n-1}^{n}\frac{(n-t)^k}{k!}\log\bar{\sigma}_t^2(\theta)dt \\ &\quad= \int_{n-1}^{n}\frac{(n-t)^k}{k!}\left(\int_{n-1}^{t}b_s(\theta)ds+\log\sigma_{n-1}^2(\theta)\right)dt \\ &\quad= \frac{1}{(k+1)!}\log\sigma_{n-1}^2(\theta)+\int_{n-1}^{n}b_s(\theta)\int_{s}^{n}\frac{(n-t)^k}{k!}dtds \\ &\quad= \frac{1}{(k+1)!}\log\sigma_{n-1}^2(\theta)+R(k+1) \quad \text{a.s.} \end{aligned}$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Thus, we have

$$
\begin{aligned}
R(0) &= \int_{n-1}^{n} b_t(\theta)dt \\
&= \sum_{k=0}^{\infty} \omega\beta^{-2}\frac{\beta^{k+2}}{(k+2)!} + \nu\left(\beta^{-2}\frac{\beta^{k+2}}{(k+2)!} - \beta^{-3}\frac{\beta^{k+3}}{(k+3)!}\right) \\
&\quad + \sum_{k=0}^{\infty} b_{n-1}(\theta)\left\{\beta^{-1}\frac{\beta^{k+1}}{(k+1)!} + (\gamma-1)\beta^{-2}\frac{2\beta^{k+2}}{(k+2)!}\right\} \\
&\quad - \sum_{k=0}^{\infty} \log\sigma_{n-1}^2(\theta)\left\{\left(\beta^*-\beta\right)\beta^{-2}\frac{\beta^{k+2}}{(k+2)!} - 2\beta^*\beta^{-3}\frac{\beta^{k+3}}{(k+3)!}\right\} \\
&\quad + \sum_{k=0}^{\infty} 2\nu\int_{n-1}^{n-1} \beta^k\frac{(n-t)^{k+2}(k+1)}{(k+2)!}Z_t dW_t \\
&= \omega\varrho_2 + \nu(\varrho_2 - 2\varrho_3) - \{(\beta^*-\beta)\varrho_2 - 2\beta^*\varrho_3\}\log\sigma_{n-1}^2(\theta) \\
&\quad + \{\varrho_1 + (\gamma-1)\varrho_2\}b_{n-1}(\theta) + D_n \\
&= \omega\varrho_2 + \nu(\varrho_2 - 2\varrho_3) - \log\sigma_{n-1}^2(\theta) + \varrho b_{n-1}(\theta) + D_n \\
&= h_n(\theta) - \log\sigma_{n-1}^2(\theta) + D_n \text{ a.s.}
\end{aligned}
$$

Then, by (7.1), we have

$$
\begin{aligned}
\int_{n-1}^{n}\sigma_t^2(\theta)dt &= \sigma_{n-1}^2\exp\left(\int_{n-1}^{n} b_t(\theta)dt\right) \\
&= \exp\left(h_n(\theta) + D_n\right) \quad \text{a.s.}
\end{aligned}
$$

For (b), since $\mathbb{E}[\exp(D_n)]$ is a constant, we have

$$
\int_{n-1}^{n}\sigma_t^2(\theta)dt = \exp\left(H_n(\theta)\right)M_n,
$$

and we obtain (b). □

7.2. Proof of Theorem 2

To simplify the notation, we use θ for the GARCH model parameters θ^g. For derivatives of any given function f at x_0, we denote $\frac{\partial f(x_0)}{\partial x} = \frac{\partial f(x)}{\partial x}\big|_{x=x_0}$. Define

$$
\widehat{L}_{n,m}(\theta) = -\frac{1}{n}\sum_{i=1}^{n}\left\{\widehat{H}_i(\theta) + \frac{RV_i}{\exp\left(\widehat{H}_i(\theta)\right)}\right\} \quad \text{and} \quad \widehat{s}_{n,m}(\theta) = \frac{\partial\widehat{L}_{n,m}(\theta)}{\partial\theta};
$$

$$
\widehat{L}_n(\theta) = -\frac{1}{n}\sum_{i=1}^{n}\left\{H_i(\theta) + \frac{\int_{i-1}^{i}\sigma_t^2(\theta_0)dt}{\exp\left(H_i(\theta)\right)}\right\} \quad \text{and} \quad \widehat{s}_n(\theta) = \frac{\partial\widehat{L}_n(\theta)}{\partial\theta};
$$

$$
L_n(\theta) = -\frac{1}{n}\sum_{i=1}^{n}\left\{H_i(\theta) + \frac{\exp\left(H_i(\theta_0)\right)}{\exp\left(H_i(\theta)\right)}\right\} \quad \text{and} \quad s_n(\theta) = \frac{\partial L_n(\theta)}{\partial\theta}.
$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Since the dependence of $H_i(\theta)$ on the initial value decays exponentially (Kim and Wang, 2016), we use the true initial value $H_0(\theta_0)$ for the rest of the proofs without loss of generality.

LEMMA 2. *Under the assumptions of Theorem 2, we have* $\widehat{\theta} \xrightarrow{p} \theta_0$.

Proof of Lemma 2. We first show the uniform convergence of $\widehat{L}_{n,m}(\theta)$. That is, we need to show that

$$\begin{aligned}\sup_{\theta\in\Theta}\left|\widehat{L}_{n,m}(\theta)-L_n(\theta)\right| &\le \sup_{\theta\in\Theta}\left|\widehat{L}_{n,m}(\theta)-\widehat{L}_n(\theta)\right|+\sup_{\theta\in\Theta}\left|\widehat{L}_n(\theta)-L_n(\theta)\right| \\ &= o_p(1).\end{aligned}$$

For $\sup_{\theta\in\Theta}\left|\widehat{L}_{n,m}(\theta)-\widehat{L}_n(\theta)\right|$, we have

$$\begin{aligned}&\sup_{\theta\in\Theta}\left|\widehat{L}_{n,m}(\theta)-\widehat{L}_n(\theta)\right| \\ &\le \sup_{\theta\in\Theta}\left\{\frac{1}{n}\left|\sum_{i=1}^{n}\widehat{H}_i(\theta)-H_i(\theta)\right|+\frac{1}{n}\left|\sum_{i=1}^{n}\frac{RV_i-\int_{i-1}^{i}\sigma_t^2(\theta_0)dt}{\exp\left(\widehat{H}_i(\theta)\right)}\right|\right. \\ &\qquad \left.+\frac{1}{n}\left|\sum_{i=1}^{n}\int_{i-1}^{i}\sigma_t^2(\theta_0)dt\left(e^{-\widehat{H}_i(\theta)}-e^{-H_i(\theta)}\right)\right|\right\} \\ &= \text{(I)}+\text{(II)}+\text{(III)}.\end{aligned}$$

For (I), we have

$$\begin{aligned}\mathbb{E}[\text{(I)}] &\le \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left|\sup_{\theta\in\Theta}\sum_{k=1}^{i-1}\beta\gamma^{k-1}\left(\log RV_{i-k}-\log IV_{i-k}\right)\right|\right] \\ &\le \frac{C}{n}\sum_{i=1}^{n}\sum_{k=1}^{i-1}\gamma_u^{k-1}\mathbb{E}\left[\left|\log RV_{i-k}-\log IV_{i-k}\right|\right] \\ &\le Cm^{-1/4}, \qquad (7.2)\end{aligned}$$

where $IV_i=\int_{i-1}^{i}\sigma_t^2(\theta_0)dt$, and the last inequality is due to Assumption 1(b). Consider (II). By Assumption 1(b) and (c), we have

$$\mathbb{E}[\text{(II)}] \le \frac{C}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left|RV_i-IV_i\right|^2\right]\mathbb{E}\left[\sup_{\theta\in\Theta}\left|\exp\left(\left|\widehat{H}_i(\theta)\right|\right)\right|^2\right] \le Cm^{-1/4}.$$

For (III), we have

$$\begin{aligned}\mathbb{E}[\text{(III)}] &\le \frac{C}{n}\sum_{i=1}^{n}\mathbb{E}\left[\sup_{\theta\in\Theta}\left(e^{-\widehat{H}_i(\theta)}-e^{-H_i(\theta)}\right)^2\right]^{1/2} \\ &\le \frac{C}{n}\sum_{i=1}^{n}\mathbb{E}\left[\sup_{\theta\in\Theta}\left(e^{4|\widehat{H}_i(\theta)|}+e^{4|H_i(\theta)|}\right)\right]^{1/4}\mathbb{E}\left[\sup_{\theta\in\Theta}\left(\widehat{H}_i(\theta)-H_i(\theta)\right)^4\right]^{1/4} \\ &\le Cm^{-1/4},\end{aligned}$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

where the first and second inequalities are due to Holder's inequality and Taylor's expansion, respectively, and the last inequality can be shown as similar to the proof of (7.2) with Assumption 1(c). Thus, we have

$$\sup_{\theta\in\Theta}\left|\widehat{L}_{n,m}(\theta)-\widehat{L}_n(\theta)\right|=o_p(1). \tag{7.3}$$

We consider $\sup_{\theta\in\Theta}\left|\widehat{L}_n(\theta)-L_n(\theta)\right|$. We have

$$\widehat{L}_n(\theta)-L_n(\theta)=-\frac{1}{n}\sum_{i=1}^{n}\frac{e^{H_i(\theta_0)}}{e^{H_i(\theta)}}(M_i-1),$$

which is a martingale process for any given θ. Thus, by the martingale convergence theorem, we can show its pointwise convergence. To show its uniform convergence, we need to show the stochastic continuity for $G_n(\theta)=\widehat{L}_n(\theta)-L_n(\theta)$. By Taylor's expansion and the mean value theorem, there exists $\widetilde{\theta}$ between θ and θ' such that

$$\begin{aligned}|G_n(\theta)-G_n(\theta')| &= \left|\frac{1}{n}\sum_{i=1}^{n}\frac{e^{H_i(\theta_0)}}{e^{H_i(\widetilde{\theta})}}\frac{\partial H_i(\widetilde{\theta})}{\partial\theta}(M_i-1)(\theta-\theta')\right| \\ &\le C\frac{1}{n}\sum_{i=1}^{n}\sup_{\widetilde{\theta}\in\Theta}\left\|\frac{e^{H_i(\theta_0)}}{e^{H_i(\widetilde{\theta})}}\frac{\partial H_i(\widetilde{\theta})}{\partial\theta}(M_i-1)\right\|_{\max}\|\theta-\theta'\|_{\max}.\end{aligned}$$

By Assumption 1(c), we have $\mathbb{E}\left[e^{4|H_i(\theta)|}\right]\le C$. Then, similar to the proofs of Lemma 3 in Kim and Wang (2016), we can show

$$\frac{1}{n}\sum_{i=1}^{n}\sup_{\widetilde{\theta}\in\Theta}\left\|\frac{e^{H_i(\theta_0)}}{e^{H_i(\widetilde{\theta})}}\frac{\partial H_i(\widetilde{\theta})}{\partial\theta}(M_i-1)\right\|_{\max}=O_p(1).$$

Thus, $G_n(\theta)$ satisfies the weak Lipschitz condition, so, by Theorem 4 in Andrews (1992), we can show the uniform convergence. Therefore, we have

$$\sup_{\theta\in\Theta}\left|\widehat{L}_n(\theta)-L_n(\theta)\right|=o_p(1). \tag{7.4}$$

By (7.3) and (7.4), we show the uniform convergence of $\widehat{L}_{n,m}(\theta)$.

When $\exp(H_i(\theta_0))=\exp(H_i(\theta))$ for all i, $L_n(\theta)$ is maximized. Obviously, θ_0 is one of the solutions. Suppose that there exists θ_* such that $\exp(H_i(\theta_0))=\exp(H_i(\theta_*))$ a.s. for all i. Since the exponential function is a strictly increasing function, we have $H_i(\theta_0)-H_i(\theta_*)=0$ a.s. for all i. Thus, θ_0 and $\theta_*=(\omega_*^g,\gamma_*,\beta_*^g)$ satisfy

$$\begin{pmatrix}1 & H_1(\theta_0) & \log IV_1\\ 1 & H_2(\theta_0) & \log IV_2\\ \vdots & \vdots & \vdots\\ 1 & H_{n-1}(\theta_0) & \log IV_{n-1}\end{pmatrix}\begin{pmatrix}\omega_*^g-\omega_0^g\\ \gamma_*-\gamma_0\\ \beta_*^g-\beta_0^g\end{pmatrix}\equiv M\begin{pmatrix}\omega_*^g-\omega_0^g\\ \gamma_*-\gamma_0\\ \beta_*^g-\beta_0^g\end{pmatrix}=0 \text{ a.s.}$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Since IV_i's are nondegenerating, M is of full rank. Then, $M^\top M$ is invertible, which implies $\theta_0 = \theta_*$ a.s. Therefore, $L_n(\theta)$ has the unique maximizer θ_0. Then, by Theorem 1 in Xiu (2010), with the uniform convergence of $\widehat{L}_{n,m}(\theta)$, we can show $\widehat{\theta} \xrightarrow{p} \theta_0$. □

Proof of Theorem 2. By the mean value theorem and Taylor's expansion, there exists $\widetilde{\theta}$ between $\widehat{\theta}$ and θ_0 such that

$$\widehat{s}_{n,m}(\widehat{\theta}) - \widehat{s}_{n,m}(\theta_0) = -\widehat{s}_{n,m}(\theta_0) = \nabla \widehat{s}_{n,m}(\widetilde{\theta})(\widehat{\theta} - \theta_0),$$

where $\nabla \widehat{s}_{n,m}(\widetilde{\theta}) = \frac{\partial \widehat{s}_{n,m}(\widetilde{\theta})}{\partial \theta^\top}$. We first consider $\widehat{s}_{n,m}(\theta_0)$. We have

$$\begin{aligned}
-\widehat{s}_{n,m}(\theta_0) &= \frac{1}{n}\sum_{i=1}^{n}\left\{1 - e^{-\widehat{H}_i(\theta_0)} RV_i\right\} \frac{\partial \widehat{H}_i(\theta_0)}{\partial \theta} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{1 - e^{-H_i(\theta_0)} IV_i\right\} \frac{\partial H_i(\theta_0)}{\partial \theta} + O_p(m^{-1/4}) \\
&= \frac{1}{n}\sum_{i=1}^{n}(1 - M_i)\frac{\partial H_i(\theta_0)}{\partial \theta} + O_p(m^{-1/4}) \\
&= O_p(n^{-1/2} + m^{-1/4}), \qquad \textbf{(7.5)}
\end{aligned}$$

where the second equality can be shown as similar to the proofs of Lemma 2, and the last equality is due to the martingale convergence theorem.

We consider $\nabla \widehat{s}_{n,m}(\widetilde{\theta})$. Similar to the proofs of (7.5) with the consistency of $\widehat{\theta}$, we can show

$$\begin{aligned}
\nabla \widehat{s}_{n,m}(\widetilde{\theta}) &= -\frac{1}{n}\sum_{i=1}^{n}\left\{1 - e^{-\widehat{H}_i(\widetilde{\theta})} RV_i\right\} \frac{\partial^2 \widehat{H}_i(\widetilde{\theta})}{\partial \theta \partial \theta^\top} - \frac{1}{n}\sum_{i=1}^{n} e^{-\widehat{H}_i(\widetilde{\theta})} RV_i \frac{\partial \widehat{H}_i(\widetilde{\theta})}{\partial \theta}\frac{\partial \widehat{H}_i(\widetilde{\theta})}{\partial \theta^\top} \\
&= -\frac{1}{n}\sum_{i=1}^{n} e^{-\widehat{H}_i(\widetilde{\theta})} RV_i \frac{\partial \widehat{H}_i(\widetilde{\theta})}{\partial \theta}\frac{\partial \widehat{H}_i(\widetilde{\theta})}{\partial \theta^\top} + o_p(1) \\
&= -\frac{1}{n}\sum_{i=1}^{n} M_i \frac{\partial H_i(\theta_0)}{\partial \theta}\frac{\partial H_i(\theta_0)}{\partial \theta^\top} + o_p(1).
\end{aligned}$$

Since IV_i's and M_i's are nondegenerating, $\frac{1}{n}\sum_{i=1}^{n} M_i \frac{\partial H_i(\theta_0)}{\partial \theta}\frac{\partial H_i(\theta_0)}{\partial \theta^\top}$ is positive definite. Thus, by (7.5), we have

$$\widehat{\theta}^g - \theta_0 = O_p(n^{-1/2} + m^{-1/4}).$$

We show the asymptotic normality. By Theorem 1(b), we have

$$\begin{aligned}
H_n(\theta) &= \omega^g + (\gamma + \beta^g)H_{n-1}(\theta) + \beta^g \log M_{n-1} \\
&= \frac{\omega^g}{1 - \gamma - \beta^g} + \sum_{i=1}^{\infty} \beta^g (\gamma + \beta^g)^{i-1} \log M_{n-i}.
\end{aligned}$$

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Since $\log M_i$'s are i.i.d., $(H_n(\theta), M_n)$ is strictly stationary. By Theorem 2.1 of Francq, Wintenberger, and Zakoian (2013) and Theorem 2.5 of Bougerol and Picard (1992), $(H_n(\theta), M_n)$ is ergodic. Then, applying the martingale central limit theorem, we obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(1-M_i)\frac{\partial H_i(\theta_0)}{\partial \theta} \xrightarrow{d} N(0, AV).$$

By the ergodic theorem, we can show

$$-\nabla \widehat{s}_{n,m}(\theta^*) \xrightarrow{p} V.$$

Thus, by the Slutsky theorem, we have

$$\sqrt{n}(\widehat{\theta} - \theta_0) \xrightarrow{d} N(0, AV^{-1}). \qquad \square$$

REFERENCES

Aït-Sahalia, Y., J. Fan, & D. Xiu (2010) High-frequency covariance estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association* 105(492), 1504–1517.

Aït-Sahalia, Y. & D. Xiu (2016) Increased correlation among asset classes: Are volatility or jumps to blame, or both? *Journal of Econometrics* 194(2), 205–219.

Andersen, T.G. & T. Bollerslev (1997a) Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance* 52(3), 975–1005.

Andersen, T.G. & T. Bollerslev (1997b) Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4(2–3), 115–158.

Andersen, T.G. & T. Bollerslev (1998a) Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39(4), 885–905.

Andersen, T.G. & T. Bollerslev (1998b) Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance* 53(1), 219–265.

Andersen, T.G., T. Bollerslev, F.X. Diebold, & P. Labys (2003) Modeling and forecasting realized volatility. *Econometrica* 71(2), 579–625.

Andersen, T.G., T. Bollerslev, & X. Huang (2011) A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics* 160(1), 176–189.

Andrews, D.W. (1992) Generic uniform convergence. *Econometric Theory* 8(2), 241–257.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde, & N. Shephard (2008) Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* 76(6), 1481–1536.

Bernardi, M. & L. Catania (2018) The model confidence set package for R. *International Journal of Computational Economics and Econometrics* 8(2), 144–158.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.

Bougerol, P. & N. Picard (1992) Strict stationarity of generalized autoregressive processes. *The Annals of Probability* 20(4), 1714–1730.

Campbell, J.Y. & S.B. Thompson (2008) Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies* 21(4), 1509–1531.

Corsi, F. (2009) A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7(2), 174–196.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Diebold, F.X. & R.S. Mariano (2002) Comparing predictive accuracy. *Journal of Business & Economic Statistics* 20(1), 134–144.

Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Fan, J. & D. Kim (2018) Robust high-dimensional volatility matrix estimation for high-frequency factor model. *Journal of the American Statistical Association* 113(523), 1268–1283.

Fan, J. & Y. Wang (2007) Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association* 102(480), 1349–1362.

Francq, C., O. Wintenberger, & J.-M. Zakoian (2013) GARCH models without positivity constraints: Exponential or log GARCH? *Journal of Econometrics* 177(1), 34–46.

Hansen, P.R. & Z. Huang (2016) Exponential GARCH modeling with realized measures of volatility. *Journal of Business & Economic Statistics* 34(2), 269–287.

Hansen, P.R., Z. Huang, & H.H. Shek (2012) Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics* 27(6), 877–906.

Hansen, P.R. & A. Lunde (2005) A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics* 3(4), 525–554.

Hansen, P.R., A. Lunde, & J.M. Nason (2011) The model confidence set. *Econometrica* 79(2), 453–497.

Haug, S. & C. Czado (2007) An exponential continuous-time GARCH process. *Journal of Applied Probability* 44(4), 960–976.

Jacod, J., Y. Li, P.A. Mykland, M. Podolskij, & M. Vetter (2009) Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications* 119(7), 2249–2276.

Kawakatsu, H. (2006) Matrix exponential GARCH. *Journal of Econometrics* 134(1), 95–128.

Kim, D. & J. Fan (2019) Factor GARCH–Itô models for high-frequency data with application to large volatility matrix prediction. *Journal of Econometrics* 208(2), 395–417.

Kim, D., M. Shin, & Y. Wang (2022) Overnight GARCH–Itô volatility models. *Journal of Business & Economic Statistics*, to appear.

Kim, D. & Y. Wang (2016) Unified discrete-time and continuous-time models and statistical inferences for merged low-frequency and high-frequency financial data. *Journal of Econometrics* 194, 220–230.

Kim, D., Y. Wang, & J. Zou (2016) Asymptotic theory for large volatility matrix estimation based on high-frequency financial data. *Stochastic Processes and their Applications* 126, 3527–3577.

Klüppelberg, C., A. Lindner, & R. Maller (2004) A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *Journal of Applied Probability* 41(3), 601–622.

Martens, M. (2002) Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. *Journal of Futures Markets: Futures, Options, and Other Derivative Products* 22(6), 497–518.

Nelson, D.B. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society* 59, 347–370.

Patton, A.J. (2011) Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160(1), 246–256.

Shephard, N. & K. Sheppard (2010) Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics* 25(2), 197–231.

Shin, M., Kim, D., & Fan, J. (2021) Adaptive robust large volatility matrix estimation based on high-frequency financial data. Available at SSRN 3793394.

Song, X., D. Kim, H. Yuan, X. Cui, Z. Lu, Y. Zhou, & Y. Wang (2021) Volatility analysis with realized GARCH–Itô models. *Journal of Econometrics* 222(1), 393–410.

Tao, M., Y. Wang, Q. Yao, & J. Zou (2011) Large volatility matrix inference via combining low-frequency and high-frequency approaches. *Journal of the American Statistical Association* 106(495), 1025–1040.

Taylor, N. (2007) A note on the importance of overnight information in risk management models. *Journal of Banking & Finance* 31(1), 161–180.

Todorova, N. & M. Souček (2014) Overnight information flow and realized volatility forecasting. *Finance Research Letters* 11(4), 420–428.

https://doi.org/10.1017/S0266466622000585 Published online by Cambridge University Press

Tseng, T.-C., H.-C. Lai, & C.-F. Lin (2012) The impact of overnight returns on realized volatility. *Applied Financial Economics* 22(5), 357–364.

Xiu, D. (2010) Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics* 159(1), 235–250.

Zhang, L. (2006) Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli* 12(6), 1019–1043.

Zhang, L., P.A. Mykland, & Y. Aït-Sahalia (2005) A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100(472), 1394–1411.

Zhang, X., D. Kim, & Y. Wang (2016) Jump variation estimation with noisy high frequency financial data via wavelets. *Econometrics* 4(3), 34.