



# Realized stochastic volatility models with generalized Gegenbauer long memory

Manabu Asai<sup>a,\*</sup>, Michael McAleer<sup>b,c,d,e,f</sup>, Shelton Peiris<sup>g</sup>

<sup>a</sup>Faculty of Economics, Soka University, Hachioji, Tokyo 192-8577, Japan

<sup>b</sup>Department of Finance, Asia University, Taichung 41354, Taiwan

<sup>c</sup>Discipline of Business Analytics, University of Sydney Business School, Sydney 2006, Australia

<sup>d</sup>Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3000 Rotterdam, The Netherlands

<sup>e</sup>Department of Economic Analysis and ICAE, Complutense University of Madrid, 28040 Madrid, Spain

<sup>f</sup>Institute of Advanced Sciences, Yokohama National University, Yokohama 240-8501, Japan

<sup>g</sup>School of Mathematics and Statistics, University of Sydney, Sydney 2006, Australia

## ARTICLE INFO

### Article history:

Received 5 May 2018

Revised 21 December 2018

Accepted 21 December 2018

Available online 9 January 2019

### Keywords:

Stochastic volatility

Realized volatility measure

Long memory

Gegenbauer polynomial

Seasonality

Whittle likelihood

## ABSTRACT

Fractionally differenced processes have received a great deal of attention due to their flexibility in financial applications with long memory. In this paper, new realized stochastic volatility (RSV) models are developed: one is a RSV model with general Gegenbauer long memory (GGLM), while the other is a RSV model with seasonal long memory (SLM). The RSV model uses the information from returns and realized volatility measures simultaneously. The long memory structure of both models can describe unbounded peaks, apart from the origin in the power spectrum. For estimating the RSV–GGLM model, a two step method is suggested: the location parameters for the peaks of the power spectrum are estimated in the first step, while the remaining parameters are estimated based on the Whittle likelihood in the second step. Monte Carlo experiments give results for investigating the finite sample properties of the estimators, with a quasi-likelihood ratio test of the RSV–SLM model against the RSV–GGLM model. The RSV–GGLM and RSV–SLM models are applied to three stock market indices, for which the estimation and forecasting results indicate the adequacy of considering general long memory.

© 2019 EcoSta Econometrics and Statistics. Published by Elsevier B.V. All rights reserved.

## 1. Introduction

For purposes of modeling financial time series, a stylized fact is that volatility has long memory. One of the popular approaches is to apply an autoregressive fractionally-integrated moving-average (ARFIMA) process to (log-)volatility. In the class of generalized autoregressive conditional heteroskedasticity (GARCH) models, Baillie et al. (1996) and Bollerslev and Mikkelsen (1996) developed the fractionally integrated GARCH (FIGARCH) and fractionally integrated Exponential GARCH (FIEGARCH) models, respectively. For stochastic volatility (SV) models, Breidt et al. (1998) developed the long memory stochastic volatility (LMSV) model for unobserved log-volatility using asset return series, while Andersen et al. (2001), Andersen et al. (2003), Pong et al. (2004), Koopman et al. (2005), and Asai et al. (2012) estimated LMSV models using

\* Corresponding author.

E-mail addresses: [m-asai@soka.ac.jp](mailto:m-asai@soka.ac.jp) (M. Asai), [michael.mcaleer@gmail.com](mailto:michael.mcaleer@gmail.com) (M. McAleer), [shelton.peiris@sydney.edu.au](mailto:shelton.peiris@sydney.edu.au) (S. Peiris).

daily realized volatility (RV). As an alternative to the ARFIMA model, [Corsi \(2009\)](#) suggested a heterogeneous autoregressive (HAR) model to approximate long memory using RV.

As extensions of the long memory structure in an ARFIMA process, seasonal (periodical) long memory, Gegenbauer processes, and their general class are considered. The Gegenbauer process is based on Gegenbauer polynomials, developed by [Gray et al. \(1989\)](#) (see the survey paper of [Dissanayake et al., 2018](#) for instance). While the spectral density of the ARFIMA process is unbounded at the origin, the Gegenbauer process has a peak at a different frequency, which is referred to as the Gegenbauer frequency. As suggested in [Woodward et al. \(1998\)](#), general (or multifactor) Gegenbauer process has multiple (unbounded) peaks. The general Gegenbauer process encompasses seasonal long memory as a special case. While [Bordignon et al. \(2009\)](#) extended the FIGARCH and FIEGARCH models by accommodating seasonal long memory, [Bordignon et al. \(2007\)](#) developed the general Gegenbauer GARCH model. Although their focus is on investigating the long memory structure within a day, it may also be worth examining periodic/aperiodic long memory such as the weekly, monthly, and annual span for instance, using the general Gegenbauer process with daily realized volatility measure.

For modeling asset returns and realized volatility measure simultaneously, [Hansen et al. \(2012\)](#) suggested a realized GARCH framework (see also [Hansen and Huang, 2016](#)). The corresponding structure for SV is often referred to as the 'realized SV' (RSV) model, which is considered by [Takahashi et al. \(2009\)](#), [Koopman and Scharth \(2013\)](#), [Shirota et al. \(2014\)](#), and [Asai et al. \(2017\)](#), among others. [Shirota et al. \(2014\)](#) and [Asai et al. \(2017\)](#) accommodated the ARFIMA process in the volatility process. While [Shirota et al. \(2014\)](#) use the Markov chain Monte Carlo technique, [Asai et al. \(2017\)](#) estimated their model using the Whittle likelihood.

In this paper, we develop an RSV model with general Gegenbauer long memory. If the Gegenbauer frequencies of log-volatility are predetermined, we can use the Whittle likelihood estimator of [Hosoya \(1997\)](#) and [Zaffaroni \(2009\)](#) to estimate the RSV model, as in [Asai et al. \(2017\)](#). However, the Gegenbauer frequencies are unknown, and so we use the non-parametric estimator of [Hidalgo and Soulier \(2004\)](#), as in [Artiach and Arteche \(2012\)](#), who investigated the long memory property of the level and variance of the number of sunspots.

The organization of the paper is as follows. [Section 2](#) develops the RSV model with general Gegenbauer long memory, and discusses the differences from the model with seasonal long memory. [Section 3](#) explains the estimation method based on the Whittle likelihood under predetermined Gegenbauer frequencies, and shows the approach of [Hidalgo and Soulier \(2004\)](#) for estimating and selecting the Gegenbauer frequencies. [Section 3](#) provides the finite sample properties of these estimators, and the likelihood ratio statistic for testing the seasonal long memory against the general long memory. [Section 4](#) presents empirical results using the daily returns and realized volatility measures of three stock indices, namely Standard & Poors 500, FTSE 100, and Nikkei 225. [Section 5](#) provides some concluding remarks.

## 2. Realized SV with generalized Gegenbauer long memory

Let  $y_t$  and  $x_t$  denote the return and the log of realized volatility measure of a financial asset, respectively. We present the new realized SV model with generalized Gegenbauer long memory (RSV-GGLM), as follows:

$$y_t = \varepsilon_t \exp(h_t/2), \quad \varepsilon_t \sim N(0, 1), \quad t = 1, \dots, T, \quad (1)$$

$$x_t = h_t + v_t, \quad v_t \sim N(0, \sigma_v^2), \quad (2)$$

$$\phi(L)P(L)(h_{t+1} - \mu) = \theta(L)\eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad (3)$$

where

$$P(L) = \prod_{l=1}^k (1 - 2 \cos(\omega_l)L + L^2)^{d_l} (1 - L)^d,$$

$\varepsilon_t$ ,  $v_t$ , and  $\eta_t$  are independent processes,  $L$  is the lag operator,  $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$ , and  $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$ . The log-volatility process  $h_t$  is latent, while  $y_t$  is observed. We assume the roots of  $\phi(z)$  and  $\theta(z)$  lie outside the unit circle to ensure stationarity and invertibility of  $\{h_t\}$ , respectively. Equation (3) is known as the  $k$ -factor Gegenbauer process or generalized exponential model (see [Woodward et al., 1998](#); [Hsu and Tsai, 2009](#); [Mc Elroy and Holan, 2012](#) for instance). For the stationarity condition on  $P(L)$ , we assume  $|d| < 1/2$ ,  $|d_l| < 1/2$ , and  $0 < \omega_l < \pi$  (see [Woodward et al., 1998](#); [Mc Elroy and Holan, 2012](#)). Regarding the long memory property, we need to assume  $\max\{d, d_1, \dots, d_k\} > 0$ , as shown by Theorem 1 of [Mc Elroy and Holan \(2012\)](#). We can include  $(1 + L)^a$  explicitly in  $P(L)$ . However, it can be expressed by setting  $\omega_k = \pi$  and  $d_k = a/2$  on  $(1 - 2 \cos(\omega_k)L + L^2)^{d_k}$ , and hence it is redundant. Furthermore, it is rare to find such a case in the analysis of volatility of financial time series, as investigated by [Ding and Granger \(1996\)](#), [Baillie et al. \(1996\)](#) and [Bollerslev and Mikkelsen \(1996\)](#). For these reasons, we do not use  $(1 + L)^a$  explicitly.

By excluding the data of  $x_t$ , the model reduces to the class of generalized long memory SV models, encompassing the long memory SV model of [Breidt et al. \(1998\)](#) with  $k = 0$ , and the Gegenbauer ARMA SV (GARMASV) model of [Artiach and Arteche \(2012\)](#) with  $k = 1$  and  $d = 0$ . Furthermore, the RSV-GGLM model extends the long memory part of the realized SV models of [Shirota et al. \(2014\)](#) and [Asai et al. \(2017\)](#). Note that we do not consider asymmetric effects and heavy-tails,

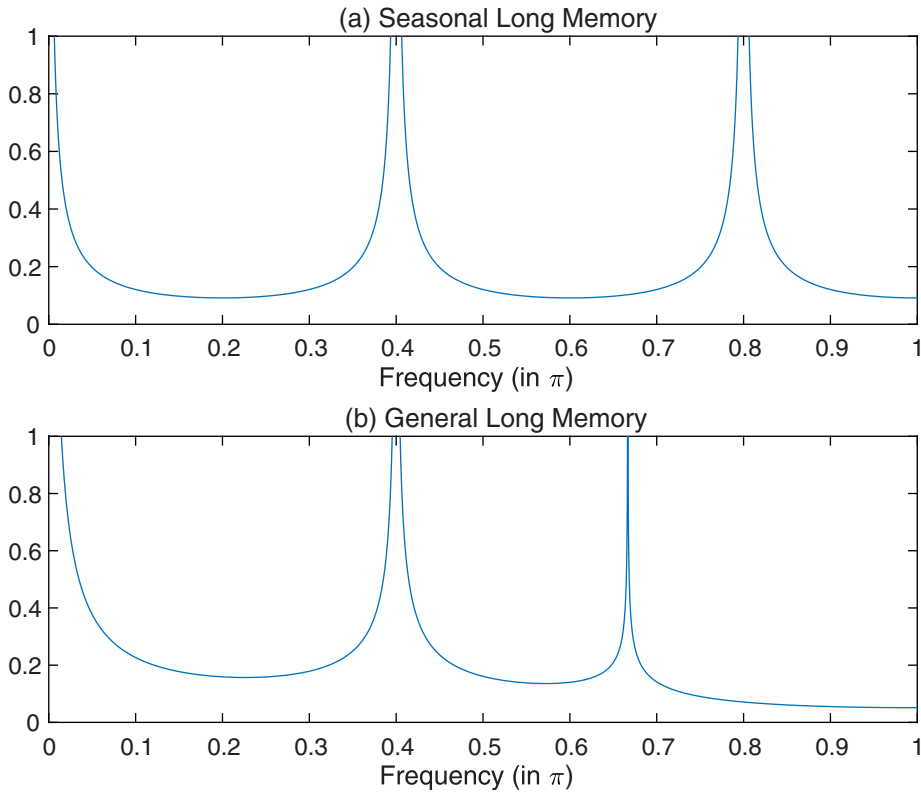


Fig. 1. Power spectrum of long memory processes.

unlike [Shirota et al. \(2014\)](#) and [Asai et al. \(2017\)](#), in order to concentrate on the specification and estimation of various long memory structures.

To consider the structure of the RSV–GGLM model, we start from a simple Gegenbauer process. When  $k = 1$  and  $d = 0$ , we can write [Eq. \(3\)](#) as:

$$h_{t+1} = \mu + (1 - 2 \cos(\omega_1)L + L^2)^{-d_1} [\phi(L)]^{-1} \theta(L) \eta_t, \tag{4}$$

which is known as the Gegenbauer process, named after the Gegenbauer polynomials defined by  $(1 - 2 \cos(\omega_1)z + z^2)^{-d_1} = \sum_{j=0}^{\infty} b_j z^j$ . The power spectrum of the Gegenbauer process is given by:

$$f_h(\lambda) = \frac{\sigma_\eta^2}{2\pi} [2(\cos \lambda - \cos \omega_1)]^{-2d_1} g_h(\lambda), \quad -\pi < \lambda < \pi, \tag{5}$$

where  $g_h(\lambda) = \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}$  corresponds to the autoregressive and moving-average (ARMA) part. The power spectrum shows the long memory feature characterized by an unbounded spectrum at the Gegenbauer frequency  $\omega_1$ . By the structure, the RSV–GGLM model accommodates conventional long memory and multi-factor Gegenbauer long memory.

At this stage, we should consider the difference between  $P(L)$  and the seasonal long memory filter,  $(1 - L^s)^d$  (see [Porter-Hudak, 1990](#)). As discussed in [Bordignon et al. \(2007\)](#), we can decompose the seasonal filter, as in  $P(L)$ . For instance, if we consider a weekly pattern for daily data ( $s = 5$ ), we obtain:

$$(1 - L^5)^d = (1 - L)^d \left(1 - 2 \cos\left(\frac{2\pi}{5}\right)L + L^2\right)^d \left(1 - 2 \cos\left(\frac{2\pi}{2.5}\right)L + L^2\right)^d.$$

Hence, generalized Gegenbauer processes encompass seasonal ARFIMA models. [Fig. 1](#) shows the power spectrum of a seasonal long memory process,  $(1 - L^5)^{0.4} h_t = \eta_{t-1}$ , and a general Gegenbauer process,  $(1 - L)^{0.4} (1 - 2 \cos(2\pi/5)L + L^2)^{0.3} (1 - 2 \cos(2\pi/3)L + L^2)^{0.2} h_t = \eta_{t-1}$ . Note that the GGLM process includes multiple seasonal ARFIMA structure, as considered in [Chan and Tsai \(2012\)](#).

### 3. Estimation and forecasting

#### 3.1. Whittle likelihood estimation of short and long memory parameters

Following Zaffaroni (2009) and Asai et al. (2017), we consider the log of the square of  $y_t$  as  $\underline{y}_t = \ln(y_t^2)$ . By the transformation, we obtain the linearized model:

$$\underline{y}_t = c + \alpha_t + u_t, \quad x_t = \mu + \alpha_t + v_t, \quad \alpha_t = [P(L)]^{-1}[\phi(L)]^{-1}\theta(L)\eta_{t-1},$$

where  $c = \mu + E(\ln \varepsilon_t^2)$ ,  $\alpha_t = h_t - \mu$ , and  $u_t = \ln \varepsilon_t^2 - E(\ln \varepsilon_t^2)$ . By Harvey et al. (1994), it is known that  $E(\ln \varepsilon_t^2) = -1.2703$  and  $V(\ln \varepsilon_t^2) = \pi^2/2$ . Since  $u_t$  is independent with mean zero and variance,  $\sigma_u^2 = \pi^2/4$ ,  $\underline{y}_t$  follows the long memory process with additive noise. Furthermore, we consider the mean subtracted series,  $\mathbf{z}_t = (y_t^\dagger, x_t^\dagger)'$ , where  $y_t^\dagger = \underline{y}_t - c$  and  $x_t^\dagger = x_t - \mu$ , in order to obtain:

$$\mathbf{z}_t = \begin{pmatrix} u_t + \sum_{j=0}^{\infty} \psi_j \eta_{t-j-1} \\ v_t + \sum_{j=0}^{\infty} \psi_j \eta_{t-j-1} \end{pmatrix} = \sum_{j=0}^{\infty} G_j \mathbf{e}_{t-j}, \tag{6}$$

where  $\sum_{j=0}^{\infty} \psi_j z^j = [P(z)]^{-1}[\phi(z)]^{-1}\theta(z)$ ,  $\mathbf{e}_t = (u_t, v_t, \eta_t)'$ , and

$$G_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad G_j = \begin{pmatrix} 0 & 0 & \psi_j \\ 0 & 0 & \psi_j \end{pmatrix} \quad (j \geq 1),$$

with  $E(\mathbf{e}_t) = 0$  and  $V(\mathbf{e}_t) = \Sigma_e = \text{diag}(\sigma_u^2, \sigma_v^2, \sigma_\eta^2)$ . Although the process  $\{u_t\}$  is non-Gaussian, a reasonable estimation procedure is to maximize the quasi-likelihood, or the likelihood computed as if  $\{u_t\}$  was Gaussian. Note that we estimate  $\mu$  by the sample mean of  $x_t$ , to reduce the number of parameters.

Before applying the method of Zaffaroni (2009) and Asai et al. (2017), we return to the estimation of a simple Gegenbauer ARMA process when  $h_t$  is observed and  $k = 1$  and  $d = 0$ . The asymptotic results of the ML estimator of Chung (1996a,b) and Peiris and Asai (2016) indicate that the ML estimator of the location parameter,  $\omega_1$ , is  $T$ -consistent rather than  $\sqrt{T}$ -consistent, and that the estimator of  $\omega_1$  and the remaining parameters are asymptotically independent. Since the Whittle likelihood (WL) estimator has the same limiting distribution as the quasi-maximum likelihood (QML) estimator in the time domain (Taniguchi and Kakizawa, 2000, Chapter 5), it is reasonable to consider estimation of  $(\omega_1, \dots, \omega_k)$  and the remaining parameters separately in the RSV-GGLM model. We will explain in this section the semiparametric estimation technique of  $\omega_l$  ( $l = 1, \dots, k$ ) for  $k$ -factor Gegenbauer processes.

Define  $\boldsymbol{\delta} = (d, d_1, \dots, d_k, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_\eta^2, \sigma_v^2)'$  and  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_k)'$  as two vectors of parameters, where it is assumed  $(\omega_1, \dots, \omega_k)$  is known. By the specification, the process  $\{\mathbf{z}_t\}$  in (6) is a second-order stationary process and has a spectral density matrix defined by  $\mathbf{f}(\lambda) = \frac{1}{2\pi} \mathbf{k}(\lambda; \boldsymbol{\delta}) \Sigma_e \mathbf{k}(\lambda; \boldsymbol{\delta})^*$ , where  $\mathbf{k}(\lambda; \boldsymbol{\delta}) = \sum_{j=0}^{\infty} G_j e^{i\lambda j}$ , which yields:

$$\mathbf{f}(\lambda) = \frac{1}{2\pi} \begin{pmatrix} K_{11}(\lambda) & K_{12}(\lambda) \\ K_{12}(\lambda)^* & K_{22}(\lambda) \end{pmatrix}, \tag{7}$$

with

$$K_{11}(\lambda) = \sigma_v^2 + \sigma_\eta^2 |\psi(e^{i\lambda})|^2, \quad K_{12}(\lambda) = \sigma_\eta^2 |\psi(e^{i\lambda})|^2, \quad K_{22}(\lambda) = \sigma_u^2 + \sigma_\eta^2 |\psi(e^{i\lambda})|^2.$$

Note that we can write  $|\psi(e^{i\lambda})|^2 = |P(e^{i\lambda})|^{-2} g_h(\lambda)$ ,

$$|P(e^{i\lambda})|^2 = [2 \sin(\lambda/2)]^{2d} \prod_{l=1}^k [2(\cos(\lambda) - \cos(\omega_l))]^{2d_l},$$

and  $g_h(\lambda)$  is defined by equation (5). The (1,1)th element of  $\mathbf{f}(\lambda)$  is the spectral density of  $x_t$ , which can be interpreted as the conventional signal plus noise process. The (2,2)th element of  $\mathbf{f}(\lambda)$  is the spectral density of  $\log y_t^2$ , and corresponds to the result of Breidt et al. (1998).

Let  $\mathbf{I}_T(\mathbf{z}, \lambda)$  be the periodogram matrix defined by:

$$\mathbf{I}_T(\mathbf{z}, \lambda) = \mathbf{w}_T(\lambda) \mathbf{w}_T(\lambda)^*, \quad -\pi < \lambda \leq \pi,$$

where  $\mathbf{w}_T(\lambda)$  is the finite Fourier transform, defined by:

$$\mathbf{w}_T(\lambda) = \frac{1}{\sqrt{2\pi}} \sum_{t=1}^T \mathbf{z}_t e^{it\lambda}.$$

For purposes of deriving the quasi-likelihood function, we treat the process  $\mathbf{z}_t$  as Gaussian. Choose the frequencies  $\lambda_j$ ,  $j = 1, \dots, n$ , equi-spaced in the region  $(-\pi, \pi)$  so that  $\mathbf{f}(\lambda)$  is continuous at  $\lambda = \lambda_j$ . Then the finite Fourier transform  $\mathbf{w}_T(\lambda_j)$ ,  $j = 1, \dots, n$ , will have a complex-valued multivariate normal distribution which, for large  $T$ , is approximately independent, each with probability density function given by:

$$\pi^{-1} \{ \det \mathbf{f}(\lambda_j; \boldsymbol{\delta}) \}^{-1/2} \exp \left[ -\frac{1}{2} \text{tr} \{ \mathbf{f}^{-1}(\lambda_j; \boldsymbol{\delta}) \mathbf{w}_T(\lambda_j) \mathbf{w}_T(\lambda_j)^* \} \right], \quad j = 1, \dots, T.$$

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

$$\bar{L}_T(\boldsymbol{\delta}) = -\frac{1}{2} \sum_{j=1}^T \left[ \log \det \mathbf{f}(\lambda_j; \boldsymbol{\delta}) + \text{tr} \{ \mathbf{f}^{-1}(\lambda_j; \boldsymbol{\delta}) \mathbf{I}_T(\mathbf{z}, \lambda_j) \} \right]. \tag{8}$$

In integral form, Eq. (8) has the expression:

$$-\frac{T}{4\pi} \left[ \int_{-\pi}^{\pi} \log \det \mathbf{f}(\lambda; \boldsymbol{\delta}) d\lambda + \int_{-\pi}^{\pi} \text{tr} \{ \mathbf{f}^{-1}(\lambda; \boldsymbol{\delta}) \mathbf{I}_T(\mathbf{z}, \lambda) \} d\lambda \right]. \tag{9}$$

The function  $\bar{L}_T(\boldsymbol{\delta})$  is called the quasi-log-likelihood function. The approximation was originally proposed by Whittle (1952) for scalar-valued stationary processes (see also Dunsmuir and Hannan, 1976; Taniguchi and Kakizawa, 2000). Define the Whittle likelihood (WL) estimator,  $\hat{\boldsymbol{\delta}}_T$ , which is obtained by minimizing  $-\bar{L}_T(\boldsymbol{\delta})$ . In practice, we use the discrete quasi-log-likelihood (8) with frequency  $\lambda_j = 2\pi j/T$  ( $j = 1, \dots, \lfloor (T-1)/2 \rfloor$ ), for the symmetry of the Fourier transform, as in standard empirical analysis.

Following Hosoya (1997), define the quantity:

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

where

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

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

Noting that

$$\det \mathbf{f}(\lambda; \boldsymbol{\delta}) = \frac{1}{2\pi} \left\{ \sigma_v^2 \sigma_u^2 + (\sigma_v^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\eta^2) |\psi(e^{i\lambda})|^2 \right\},$$

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

As discussed in Asai et al. (2017), we can obtain the asymptotic results of the WL estimator by checking the conditions of Hosoya (1997). If the vector of frequency parameters,  $\boldsymbol{\omega}$ , is known, we can apply the approach which was used to prove Chan and Tsai (2012, Theorem 2) and Tsai et al. (2015, Theorem 1), in order to verify Hosoya (1997, Assumptions A, C, and D) to show the consistency and asymptotic normality of the WL estimator. Then we obtain:

$$\sqrt{T}(\hat{\boldsymbol{\delta}}_T - \boldsymbol{\delta}_0) \xrightarrow{d} N(0, W^{-1}U(W^*)^{-1}), \tag{10}$$

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

$$U_{jl} = 4\pi \int_{-\pi}^{\pi} \text{tr} [\mathbf{h}_j(\lambda; \boldsymbol{\delta}_0) \mathbf{f}(\lambda) \mathbf{h}_l(\lambda; \boldsymbol{\delta}_0) \mathbf{f}(\lambda)] d\lambda \Big|_{\boldsymbol{\delta}=\boldsymbol{\delta}_0} + C_1 \left\{ \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\lambda_1) \mathbf{h}_j(\lambda_1; \boldsymbol{\delta}_0) k(\lambda_1) \Big|_{\boldsymbol{\delta}=\boldsymbol{\delta}_0} d\lambda_1 \right]_{11} \right\}^2, \tag{11}$$

and

$$\int_{-\pi}^{\pi} k^*(\lambda) \mathbf{h}_j(\lambda; \boldsymbol{\delta}_0) k(\lambda) \Big|_{\boldsymbol{\delta}=\boldsymbol{\delta}_0} d\lambda = 0 \quad \text{for } \delta_j \in (d, d_1, \dots, d_k, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q), \tag{12}$$

with  $C_1$  as the fourth cumulant of  $u_t$ , given by

$$C_1 = E(u_t^4) - 3\{E(u_t^2)\}^2 = \psi^{(3)}\left(\frac{1}{2}\right) - \frac{3\pi^2}{16},$$

where  $\psi^{(3)}(z)$  is the penta gamma function (see Abramovits and Stegun, 1970, Eq. (26.4.36) for the result of the fourth moment of  $u_t$ ). Although  $U_{jl}$  defined by Hosoya (1997, Theorem 2.2) is based on the fourth-order spectral density, it can be simplified as in Hosoya (1997, (11) under Assumption F) (see also Taniguchi and Kakizawa, 2000, Eq. (5.3.22) and Zaffaroni, 2009, Theorem 2), which can be verified straightforwardly by the structure of the RSV-GGLM models.

We can allow a non-Gaussian distribution for  $\varepsilon_t$  by setting  $\sigma_u^2$  as a free parameter (see Harvey et al., 1994, for instance), and by adding it in  $\boldsymbol{\delta}$ . As explained above, we use a two step procedure rather than estimating  $(\boldsymbol{\delta}, \boldsymbol{\omega})$  simultaneously. In the first step, we obtain a consistent estimate of  $\boldsymbol{\omega}$ ,  $\hat{\boldsymbol{\omega}}$ , by using a semiparametric method suggested by Harvey et al. (1994). In the second step, we obtain the WL estimate,  $\hat{\boldsymbol{\delta}}$ , by minimizing  $-\bar{L}_T(\boldsymbol{\delta}, \hat{\boldsymbol{\omega}})$ . Since we use  $\hat{\boldsymbol{\omega}}$  instead of the true  $\boldsymbol{\omega}$ , no asymptotic results are yet available for this case.

### 3.2. Semiparametric estimation of location frequency parameters and identification of $k$

We explain the semiparametric technique of [Harvey et al. \(1994\)](#) for estimating the parameters of  $\omega$ . We assume  $k$  is known until we discuss the identification of parameters. For purposes of introducing the approach of [Harvey et al. \(1994\)](#), we consider a simple case of a univariate process which produces  $I_T(\lambda)$ , with the assumptions  $d = 0$ ,  $\omega_1 \neq 0$ ,  $\omega_2 \neq 0$ ,  $d_1 \geq d_2$ , and  $k = 2$ . Then we can estimate  $\omega_1$  and  $\omega_2$  consistently as:

$$\hat{\omega}_1 = \frac{2\pi}{T} \arg \max_{1 \leq j \leq m} I_T(\lambda_j), \quad \hat{\omega}_2 = \frac{2\pi}{T} \arg \max_{\substack{1 \leq j \leq m \\ |\lambda_j - \hat{\omega}_1| \geq z_T/T}} I_T(\lambda_j),$$

where  $z_T = T \exp(-\sqrt{\ln(T)})$ , and  $m$  is an integer between 1 and  $\lfloor (T - 1)/2 \rfloor$ , satisfying at least:

$$\frac{1}{m} + \frac{m}{T} \rightarrow 0 \text{ as } T \rightarrow \infty.$$

After we estimate  $\omega_1$ , it is possible to estimate the second location parameter,  $\omega_2$ , which has a sufficient distance from the first location. For general  $k$ , we can estimate  $(\omega_1, \dots, \omega_k)$  sequentially.

[Harvey et al. \(1994\)](#) modified the GPH estimator of [Geweke and Porter-Hudak \(1983\)](#), which was originally suggested to estimate long memory parameter,  $d$ , using a log-periodogram regression, in order to estimate  $d_l$  at the Gegenbauer frequency  $\omega_l$ . To identify the number of location frequencies,  $k$ , we follow the approach of [Harvey et al. \(1994\)](#), based on their modified GPH estimator for  $d_1, \dots, d_k$ , which is defined by:

$$\hat{d}_l = \sum_{\substack{1 \leq |j| \leq m \\ 0 < \hat{\omega}_l + \lambda_j \leq \pi}} \xi_k \ln \{ I_T(\hat{\omega}_l + \lambda_j) \}, \tag{13}$$

where  $\xi_k = s_m^{-2} (\zeta(\lambda_j) - \bar{\zeta}_m)$ ,  $\zeta(\lambda) = -\ln(|1 - e^{i\lambda}|)$ ,  $\bar{\zeta}_m = m^{-1} \sum_{j=1}^m \zeta(\lambda_j)$ , and  $s_m^2 = \sum_{j=1}^m (\zeta(\lambda_j) - \bar{\zeta}_m)^2$ . [Harvey et al. \(1994\)](#) show that  $m^{1/2}(\hat{d}_l - d_l)$  converges weakly to  $N(0, \pi^2/12)$ , under the assumption of a Gaussian process. The procedure of [Harvey et al. \(1994\)](#) consists of the following steps: (i) Find the largest periodogram ordinate; (ii) if the corresponding estimate of  $d_l$  is significant, add the respective Gegenbauer filter to the model, otherwise terminate the procedure; (iii) Exclude the neighborhood of the last pole from the periodogram, and repeat the procedure from (i) onward. For the assumption of Gaussianity of the procedure, we use the data of  $x_t$ , which produces  $I_T(x_t, \lambda)$ , excluding  $y_t$ .

### 3.3. Estimating and forecasting volatility

Using the WL estimates above, we can obtain the minimum mean square linear estimator (MMSLE) of  $h_t$  from the work of [Harvey \(1998\)](#) and [Asai et al. \(2017\)](#). Define  $\mathbf{x}^\dagger = (x_1^\dagger, \dots, x_T^\dagger)'$ ,  $\mathbf{y}^\dagger = (y_1^\dagger, \dots, y_T^\dagger)'$ ,  $\mathbf{h} = (h_1, \dots, h_T)'$ ,  $\mathbf{v} = (v_1, \dots, v_T)'$ , and  $\mathbf{u} = (u_1, \dots, u_T)'$  in order to obtain:

$$\mathbf{x}^\dagger = \mathbf{h} - \mu \mathbf{1}_T + \mathbf{v}, \quad \mathbf{y}^\dagger = \mathbf{h} - \mu \mathbf{1}_T + \mathbf{u},$$

where  $\mathbf{1}_T$  is an  $T \times 1$  vector of ones. Then, the minimum mean square linear estimator of  $\mathbf{h}$  is given by:

$$\tilde{\mathbf{h}} = \mu \mathbf{1}_T + \tau^{-1} (I_T - \Sigma_\tau^{-1}) (\sigma_v^{-2} \mathbf{x}^\dagger + \sigma_u^{-2} \mathbf{y}^\dagger),$$

where  $\tau = \sigma_v^{-2} + \sigma_u^{-2}$ ,  $\Sigma_\tau = I_T + \tau \Sigma_h$ , and  $V(\mathbf{h}) = \Sigma_h$ . We obtain  $\Sigma_h$  via the algorithm of [Mc Elroy and Holan \(2012\)](#) (see the [Appendix](#) for details). [Harvey \(1998\)](#) recommends using the volatility estimate:

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

where  $\tilde{\sigma}_y^2 = T^{-1} \sum_{t=1}^T \tilde{y}_t^2$ , and  $\tilde{y}_t = y_t \exp(-0.5\tilde{h}_t)$  are the heteroskedasticity-corrected observations.

For predicting the observations for  $x_t^\dagger$  and  $y_t^\dagger$  for  $t = T + 1, \dots, T + l$ , denote  $\mathbf{x}_l^\dagger$  and  $\mathbf{y}_l^\dagger$  as the  $l \times 1$  vectors of predicted values, respectively. Then the corresponding MMSLEs are given by:

$$\tilde{\mathbf{x}}_l^\dagger = \mathbf{R}_x \Sigma_x^{-1} \mathbf{x}_l^\dagger, \quad \tilde{\mathbf{y}}_l^\dagger = \mathbf{R}_y \Sigma_y^{-1} \mathbf{y}_l^\dagger,$$

where  $\Sigma_x = \Sigma_h + \sigma_v^2 I_T$ ,  $\Sigma_y = \Sigma_h + \sigma_u^2 I_T$ ,  $R_x$  ( $R_y$ ) is the  $l \times T$  matrix of covariances between  $\mathbf{x}_l^\dagger$  and  $\mathbf{x}^\dagger$  ( $\mathbf{y}_l^\dagger$  and  $\mathbf{y}^\dagger$ ). Using  $\tilde{\mathbf{h}}_l = \mu \mathbf{1}_l + \tau^{-1} (\sigma_v^{-2} \tilde{\mathbf{x}}_l^\dagger + \sigma_u^{-2} \tilde{\mathbf{y}}_l^\dagger)$ , the predictions of  $\sigma_{T+j}^2$  ( $j = 1, \dots, l$ ) are given by exponentiating the elements of  $\tilde{\mathbf{h}}_l$ , and multiplying by  $\tilde{\sigma}_y^2$ .

### 3.4. Finite sample properties

We conducted Monte Carlo experiments for investigating the finite sample properties of the WL estimator of  $\delta$  and the semiparametric estimator of  $\omega$  for the RSV-GGLM model. We consider two kinds of long memory components:



**Table 1**  
Finite sample performance of selection procedures.

(a) Relative frequencies of selecting $k$				
$k$	DGP1		DGP2	
	$T = 1024$	$T = 2048$	$T = 1024$	$T = 2048$
0	0.0000	0.0000	0.0000	0.0000
1	0.0114	0.0000	0.3892	0.0002
2	0.6032	0.2354	0.3058	0.2984
3	0.2418	0.0870	0.0384	0.0236
4	0.1436	0.3910	0.2480	0.2030
5	0.0000	0.2866	0.0186	0.4276
6	0.0000	0.0000	0.0000	0.0382
Mean	2.5176	3.7288	2.2010	3.8740

Note: The entries show the relative frequencies of selected values of  $k$  by the procedure of Harvey et al. (1994). The true value is  $k = 2$ .

(b) Relative frequencies of containing true location parameters				
Parameters				
$\omega_1$	0.9990	1.0000	0.9858	0.9972
$\omega_2$	0.9798	0.9988	0.0730	0.3320

Note: The entries show the relative frequencies of containing true location parameters such that  $|\hat{\omega}_j - \omega_j| < \exp(-\sqrt{\ln(T)})$  for each selected value of  $k$ .

$$(d_0, d_1, d_2, \omega_1, \omega_2) = \begin{cases} (0.4, 0.4, 0.4, 2\pi/5, 2\pi/2.5) & \text{for Seasonal Long Memory (DGP1)} \\ (0.4, 0.3, 0.2, 2\pi/5, 2\pi/3) & \text{for General Gegenbauer Long Memory (DGP2),} \end{cases}$$

for which the power spectra are shown in Fig. 1. Note that the original specification of the RSV–SLM (Seasonal Long Memory) model is given by Eqs. (1)–(3), with  $P(L) = (1 - L^s)^d$ , and  $s = 5$  corresponds to the above DGP. For the remaining parameters, we specify  $(\sigma_v, \sigma_\eta, \phi, \mu) = (0.2, 0.4, 0.6, -0.1)$ . We consider sample sizes  $T = \{1024, 2048\}$ , with  $R = 5000$  replications.

The first experiment considers selection of the number of location parameters,  $k$ , for the semiparametric procedure. Table 1(a) shows the relative frequencies for selecting the number of long memory parameters via the procedure of Harvey et al. (1994), with  $m = 0.5T^{0.7}$ . The mean selected value indicates that there is an upward bias in the procedure for the sample sizes, which may be caused by over-rejection of the modified GPH estimator. Table 1(b) presents the relative frequencies of containing the true location parameters, such that  $|\hat{\omega}_j - \omega_j| < z_T/T = \exp(-\sqrt{\ln(T)})$  for each selected value of  $k$ . The frequencies of selecting true values increase as the sample size and/or the true value of long memory parameter increases. Compared to DGP1, the relative frequencies for containing the true value are smaller for  $\omega_2$  in DGP2. It may be caused by the tickness of the pole  $\omega_2$ , as shown by Fig. 1(b). As a result, the approach of Harvey et al. (1994) tends to select larger values of  $k$  for  $T = 2048$ , but the location parameter estimates chosen by the approach tend to include the true parameters.

The second experiment examines the finite sample properties of the WL estimator under the true values of  $\omega$  for the RSV–GGLM models. Table 2 reports the sample mean, standard deviation, and root mean squared error (RMSE) of the WL estimator of  $\delta$ . For  $\sigma_u, \phi, d, d_1$ , and  $d_2$ , the bias of the estimator is negligible for both  $T = 1024$  and  $T = 2048$ . While the bias for  $\sigma_v$  is upward, that of  $\sigma_\eta$  is downward. Compared with the case  $T = 1024$ , there is no improvement in the biases of  $\sigma_v$  and  $\sigma_\eta$ . However, the results for  $T = 2048$  have smaller standard deviations and RMSEs. Table 2 also shows the sample mean, standard deviation, and root mean squared error of the estimator of  $\mu$  by the sample mean of  $x_t$ , with the same implications.

By the structure of the RSV–GGLM model, we consider the quasi-likelihood ratio (QLR) statistic for testing the RSV–SLM model against the RSV–GGLM model. As shown in Taniguchi and Kakizawa (2000, Theorem 3.1.3) in the general framework, the QLR test under known  $\omega$  has the asymptotic  $\chi^2(2)$  distribution. The last entries of Table 2 report the rejection frequencies of the QLR statistic at the five percent significance level, indicating that the rejection frequency under the null model approaches the nominal size of 5% as  $T$  increases. Under the alternative model, the sample size of  $T = 1024$  is sufficient to reject the null hypothesis for the parameter set.

#### 4. Empirical analysis

The empirical analysis focuses on estimating and forecasting the RSV–GGLM model for three sets of stock indices, namely Standard & Poors 500 (S&P), FTSE 100 (FTSE), and Nikkei 225 (Nikkei). For each return computed for 1-min intervals of the trading day at  $t$  between 9:30 a.m. and 4:00 p.m., we calculated the daily volatility using the realized kernel (RK) estimator of Barndorff-Nielsen et al. (2008), which is consistent and robust to microstructure noise and jumps. We also calculate the corresponding open-close returns for the three assets.

We denote the return and log of the RK estimate at day  $t$  as  $r_t$  and  $x_t$ , respectively. The sample period is from March 23, 2007 to September 19, 2017, to obtain the last 2548 observations, excluding holidays and weekends. We use the first

**Table 2**  
Finite sample performance of WL estimator for RSV–GGLM models.

Parameters	DGP1: RSV–SLM			DGP2: RSV–GGLM		
	True	T = 1024	T = 2048	True	T = 1024	T = 2048
$\sigma_u$	2.2214	2.2192 (0.0846) [0.0847]	2.2192 (0.0605) [0.0606]	2.2214	2.2195 (0.0846) [0.0847]	2.2195 (0.0605) [0.0605]
$\sigma_v$	0.02	0.0352 (0.0407) [0.0434]	0.0375 (0.0377) [0.0416]	0.02	0.0298 (0.0323) [0.0337]	0.0294 (0.0285) [0.0300]
$\sigma_\eta$	0.4	0.3063 (0.0236) [0.0966]	0.3032 (0.0164) [0.0982]	0.4	0.3047 (0.0228) [0.0980]	0.3025 (0.0159) [0.0988]
$\phi$	0.6	0.6318 (0.0828) [0.0887]	0.6226 (0.0631) [0.0670]	0.6	0.6190 (0.0861) [0.0881]	0.6110 (0.0661) [0.0670]
$d$	0.4	0.3808 (0.0927) [0.0947]	0.3897 (0.0647) [0.0655]	0.4	0.3843 (0.0920) [0.0933]	0.3945 (0.0661) [0.0663]
$d_1$	0.4	0.4235 (0.0353) [0.0424]	0.4221 (0.0277) [0.0355]	0.3	0.3128 (0.0319) [0.0344]	0.3120 (0.0240) [0.0268]
$d_2$	0.4	0.4172 (0.0328) [0.0370]	0.4118 (0.0222) [0.0252]	0.2	0.2031 (0.0310) [0.0312]	0.2039 (0.0220) [0.0224]
$\mu$	−0.1	−0.1010 (0.3673) [0.3673]	−0.1010 (0.3386) [0.3386]	−0.1	−0.1081 (0.5085) [0.5085]	−0.1006 (0.4692) [0.4692]
QLR Test		0.0806	0.0510		1.0000	1.0000

Note: Except for ‘ $\mu$ ’ and ‘LR test’, entries show the means of the WL estimates under true  $\omega$ , and  $\mu$  is estimated by the sample mean of  $x_t$ . Standard errors are in parentheses, and root mean squared errors are in brackets. ‘QLR Test’ reports the rejection frequencies of the QLR statistic for testing the null hypothesis of the RSV–SLM model. The critical value of the QLR test with true  $\omega$  is given by 5.9915, which is the upper five percentile of  $\chi^2(2)$  distribution.

**Table 3**  
Descriptive statistics of return and log-RK.

Data	Average	Std. dev.	Skewness	Kurtosis
Return				
S&P	0.0186	1.2149	−0.2762	13.385
FTSE	−0.0275	0.9062	−0.0612	8.4948
Nikkei	−0.0267	1.1981	−0.7717	18.750
Log-RK				
S&P	−0.7221	1.1983	0.5285	3.3759
FTSE	−0.8334	0.9910	0.6429	3.4408
Nikkei	−0.5297	0.9702	0.6928	4.1009

$T = 2048$  returns for estimating the RSV–GGLM models, and the remaining 500 series for forecasting. The estimation period includes the Global Financial Crisis starting from September 15, 2008.

Table 3 presents the descriptive statistics of the returns and log-volatility for the whole sample. The empirical distribution of the returns is highly leptokurtic, and is skewed to the left. Compared with the returns series, the distribution of log-volatility is closer to the normal distribution, but is skewed to the right, and the kurtosis exceeds three. As our interest is on volatility, we use the mean subtracted returns,  $y_t = r_t - \bar{r}$ . Fig. 2 shows the sample spectral density for log-volatility. There is a clear evidence that the spectral density is unbounded at the origin,  $\omega = 0$ . Since there are several peaks apart from the origin, it is worth investigating the general pattern for the structure of long memory.

Before considering long memory process, we conduct the augmented Dickey–Fuller (ADF) test, based on the regression:

$$\Delta x_t = \alpha^{adf} x_{t-1} + \sum_{i=1}^{p-1} \beta_i^{adf} \Delta x_{t-i} + c^{adf} + \text{error}.$$

The ADF test examines the negative of the parameter  $\alpha^{adf}$  based on its  $t$  ratio, which is called the  $\tau$  statistic. Table 4 gives the results of the ADF tests for level and differenced series. All tests reject the null hypothesis of a unit root, indicating that log-volatility has no unit roots.

Table 5 gives the semiparametric estimates of the location parameter  $\omega$ , accompanied by the results of the procedure of Harvey et al. (1994) for selecting the number  $k$ . The data of S&P, FTSE, and Nikkei selected the number of components,  $k$ , as 4, 2, and 3, respectively. In the following analysis, we set  $\hat{\omega}_0 = 0$  from Fig. 2. Table 5 shows that the periods of frequencies



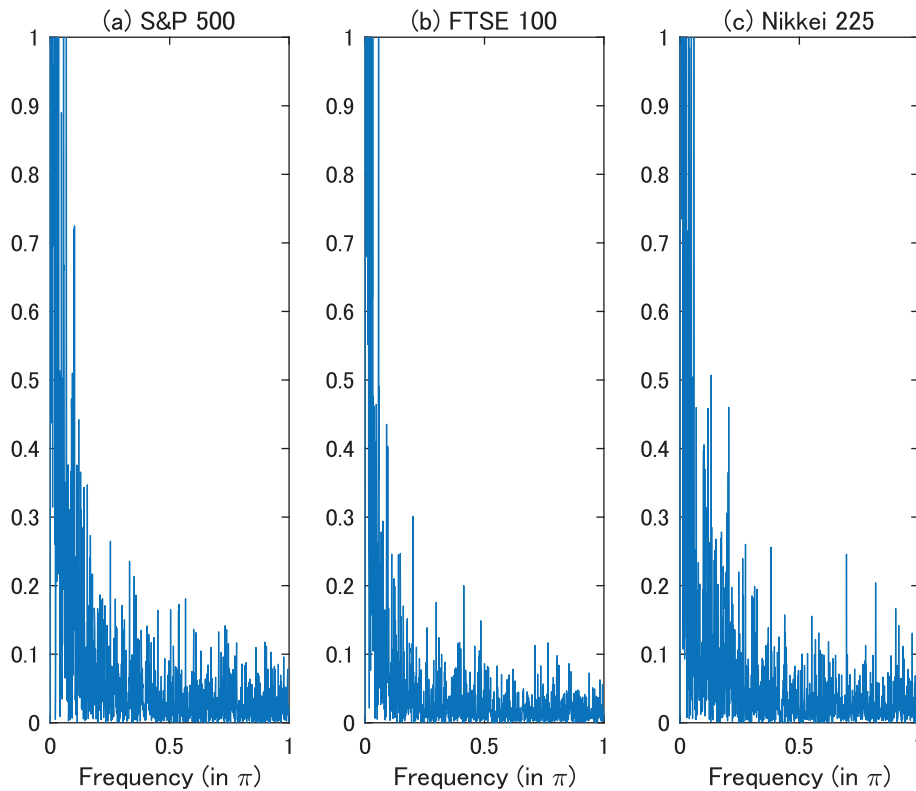


Fig. 2. Sample spectral density of log-volatility.

**Table 4**  
Augmented Dickey Fuller tests for Log-RK.

Data	Level series	Differenced series
S&P	-3.7137*(24)	-11.452*(24)
FTSE	-3.1124*(24)	-12.140*(23)
Nikkei	-4.4171*(22)	-13.855*(18)

Note: The entries show the values of the  $\tau$  statistic for the augmented Dickey Fuller test with constant term. The parentheses give the selected values of the lag length under the maximum value of 25. "\*" indicates the significance at the 5% level.

**Table 5**  
Semiparametric estimates of location parameters.

$l$	S&P			FTSE			Nikkei		
	$\hat{\omega}_l$	Days	$P$ -value	$\hat{\omega}_l$	Days	$P$ -value	$\hat{\omega}_l$	Days	$P$ -value
0	0.0010	–	0.0003*	0.0010	–	0.0001*	0.0020	–	0.0005*
1	0.0674	29.681	0.0004*	0.0898	22.261	0.0001*	0.1309	15.284	0.0013*
2	0.1553	14.124	0.0011*	0.2002	9.9902	0.0004*	0.2041	9.7990	0.0032*
3	0.2520	7.9380	0.0047*	0.4131	4.8416	0.0752	0.2744	7.2883	0.0050*
4	0.3320	6.0235	0.0189*				0.3809	5.2513	0.0553
5	0.5664	3.5310	0.6334						

Note: The estimates of  $\omega_l$  are reported with the unit of  $\pi$ . 'Days' indicates the period corresponding to  $\hat{\omega}_l$ . 'P-value' shows the  $P$ -value for the modified GPH estimates of  $d_l$ , and "\*" indicates the significance at the five percent level.

are close to (20,10,5) for FTSE, implying that  $P(L) = (1 - L^{20})^d$  is another candidate for specifying the long memory structure. As an alternative specification, we also consider  $P(L) = (1 - L^{30})^d$  and  $P(L) = (1 - L^{15})^d$  for S&P and Nikkei, respectively.

Table 6 gives the WL estimates for the RSV (without long memory), RSV-GGLM and RSV-SLM models. For the RSV model, the estimate of  $\sigma_u$  is close to  $\pi/\sqrt{2}$ , which is obtained by the standard normal distribution for  $\varepsilon_t$ . The estimates of  $\sigma_\eta$  are about 0.3, while those of  $\phi$  are greater than 0.9. These values are typical in the literature of SV models. For the RSV-SLM model, we set  $s = 30, 20, 15$  for S&P, FTSE, and Nikkei, respectively, as stated above. The QLR tests for the RSV model against

**Table 6**  
WL estimates of RSV–GGLM and RSV–SLM models.

Parameters	S&P			FTSE			Nikkei		
	RSV	RSV–SLM	RSV–GGLM	RSV	RSV–SLM	RSV–GGLM	RSV	RSV–SLM	RSV–GGLM
$\sigma_u$	2.2420 (0.0703)	2.2354 (0.0372)	2.2164 (0.0484)	2.3140 (0.0468)	2.3075 (0.1254)	2.3163 (0.0502)	2.2348 (0.0396)	2.2300 (0.1346)	2.2195 (0.0398)
$\sigma_v$	0.4015 (0.0069)	0.3728 (0.0059)	0.2231 (0.0102)	0.3358 (0.0060)	0.3073 (0.0140)	0.3412 (0.0081)	0.3614 (0.0178)	0.3545 (0.0166)	0.3398 (0.0103)
$\sigma_\eta$	0.3166 (0.0100)	0.3428 (0.0099)	0.3686 (0.0080)	0.2200 (0.0057)	0.2481 (0.0048)	0.1223 (0.0021)	0.3367 (0.0069)	0.3208 (0.0149)	0.2943 (0.0068)
$\phi$	0.9578 (0.0261)	0.9091 (0.0279)	−0.2167 (0.0057)	0.9736 (0.0265)	0.9009 (0.0143)	0.7844 (0.0011)	0.9213 (0.0446)	0.8927 (0.0238)	0.6039 (0.0154)
{s, k}		{30, 0}	{1, 4}		{20, 0}	{1, 2}	{15, 0}	{1, 3}	
d		0.0091 (0.0022)	0.4595 (0.0238)		0.1304 (0.0069)	0.4258 (0.0091)		0.1200 (0.0025)	0.4555 (0.0174)
$d_1$			0.0851 (0.0022)			−0.0131 (0.0002)			−0.0599 (0.0013)
$d_2$			0.0220 (0.0004)			0.0660 (0.0013)			0.0192 (0.0006)
$d_3$			0.0175 (0.0004)						0.0128 (0.0004)
$d_4$			0.0350 (0.0014)						
QLR Test									
$H_0$ : RSV		1.1380 [0.2861]	87.541 [0.0000]		3.7643 [0.0523]	19.344 [0.0017]		11.354 [0.0008]	50.454 [0.0000]
$H_0$ : RSV- rSLM			96.403 [0.0000]			53.109 [0.0000]			39.100 [0.0000]
$H_0$ : d = 0.5			10.648 [0.0011]			5.3205 [0.0210]			13.780 [0.0002]

Note: Standard errors are in parentheses. ‘QLR Test ( $H_0$ : RSV)’ reports the statistic for testing the null hypothesis of the RSV model, which has the asymptotic  $\chi^2(k+1)$  distribution. ‘QLR Test ( $H_0$ : RSV–SLM)’ reports the statistic for testing the null hypothesis of the RSV–SLM model, which has the asymptotic  $\chi^2(k)$  distribution. ‘QLR Test ( $H_0$ : d = 0.5)’ gives the statistic for testing  $d = 0.5$  under RSV–GGLM model, which has the asymptotic  $\chi^2(1)$  distribution. *P*-values are given in brackets.

the RSV–SLM rejected the null hypothesis of the RSV model at 5% significance level only for Nikkei. Except for  $d$  and  $\phi$ , the estimates of the RSV–SLM and RSV models are similar. The estimates of  $\phi$  in the RSV–SLM model are smaller than those of the RSV model, while the estimates of  $d$  in the RSV–SLM model are close to 0.1 and significant at the 5% level. Regarding the RSV–GGLM model, the QLR test for RSV–SLM against RSV–GGLM rejected the null hypothesis for all cases, favoring the RSV–GGLM model. The estimate of  $d$  is close to 0.45, which is dominant compared with other estimates of long memory parameters,  $d_l$  ( $l = 1, 2, 3, 4$ ). All the estimates of the long memory parameters are significant at five percent level, rejecting the RSV model with the ARFIMA(1,  $d$ , 0) specification. Compared to the RSV and RSV–SLM models, the estimates of  $\phi$  becomes smaller. Since the estimates of  $d_1$  in FTSE and Nikkei are negative and significant, they work for anti-persistent. However, as shown by [Mc Elroy and Holan \(2012, Theorem 1\)](#), the largest long memory parameter determines the long-range persistence in the autocorrelation and autocovariance functions. As the largest long memory parameter in the RSV–GGLM model is greater than 0.4 for all cases, the estimated process for log-volatility has long memory. These result suggest that the log-volatility prefers the GGLM specification, implying that, apart from the conventional long memory, the process can be explained by aperiodic long range dependencies on the days shown in [Table 5](#).

We examine the nonstationarity for the log-volatility of the RSV–GGLM model. By the ADF test in [Table 4](#), the null hypothesis of unit root hypothesis is rejected for all cases. As the estimate of  $d$  is close to 0.5, we consider testing  $d = 0.5$  by the QLR statistic. For estimating the RSV–GGLM model with  $d = 0.5$ , [Velasco and Robinson \(2000\)](#) suggest the White pseudo-maximum likelihood estimation based on the tapered periodogram. We calculated the QLR statistic using the quasi-likelihood evaluated at the estimates with  $d = 0.5$ . [Table 6](#) indicates that the null hypothesis  $d = 0.5$  is rejected at 5% significance level for all cases. The results indicate that the log-volatility is stationary for the data.

We investigate the performance of the out-of-sample forecasts of the RSV, RSV–GGLM, and RSV–SLM models. The benchmark model is the HAR model of [Corsi \(2009\)](#), which is given by:

$$x_t = c + \phi_d x_{t-1} + \phi_w(x_{t-1})_5 + \phi_w(x_{t-1})_{20} + \text{error},$$

where  $(x_{t-1})_h$  denotes the  $h$ -horizon average of past  $x_t$ . Note that  $(x_{t-1})_5$  and  $(x_{t-1})_{22}$  are the weekly and monthly averages, respectively. The model is interpreted as the AR(22) process with the parameter restrictions. Although the model is not technically a long memory process, it approximates the effects of longer horizons in a simple and parsimonious way. We also consider an improved forecast based on RSV–GGLM and RSV–SLM models, as in the approach of

**Table 7**  
Forecasting results.

Data	Loss fn.	HAR	RSV	RSV–GGLM	RSV–SLM	Imp. For.
S&P	MSE	19.335	3.6075	4.5405	12.972	<b>2.6838</b>
	QLIKE	–	[0.0376]	[0.0450]	[0.0795]	[0.0351]
FTSE	MSE	1.4527	1.1533	1.2276	1.4153	<b>0.8222</b>
	QLIKE	–	[0.0000]	[0.0001]	[0.0109]	[0.0002]
Nikkei	MSE	5.7199	1.5211	2.2127	4.4027	<b>1.2827</b>
	QLIKE	0.9979	0.7436	0.8410	0.9718	<b>0.7221</b>
Nikkei	MSE	–	[0.0000]	[0.0014]	[0.0346]	[0.0123]
	QLIKE	11.436	4.9928	5.5927	7.9499	<b>4.7234</b>
Nikkei	MSE	–	[0.0672]	[0.0693]	[0.0760]	[0.0622]
	QLIKE	1.2536	1.0796	1.0814	1.2034	<b>0.8775</b>
		–	[0.0256]	[0.0445]	[0.0175]	[0.0434]

Note: The values in the brackets are  $P$ -values of the Diebold-Mariano test against the forecast via the HAR model. 'Imp. For.' denotes the improved forecasts based on the forecasts of the RSV-GGLM and RSV-SLM models.

Granger and Ramanathan (1984). Using the in-sample estimates of log-volatility of two approaches, we obtain the coefficients by an artificial regression:

$$x_t = \beta_0 + \beta_1 \hat{h}_t^{(gglm)} + \beta_2 \hat{h}_t^{(slm)} + \text{error},$$

in order to produce an improved forecast:

$$\hat{h}_{T+j}^{(gglm,slm)} = \hat{\beta}_0 + \hat{\beta}_1 \hat{h}_{T+j}^{(gglm)} + \hat{\beta}_2 \hat{h}_{T+j}^{(slm)}.$$

As explained in Section 3.3, we predict true volatility,  $\sigma_{T+j}^2$ , for the four models and the improved forecast.

For comparing the out-of-sample forecasts, we use several loss functions as described in Patton (2011). Patton (2011) examined the functional form of the loss function for comparing volatility forecasts using imperfect volatility proxies, such that the forecasts are robust to the presence of noise in the proxies. A loss function is “robust” if the ranking of any two volatility forecasts,  $\hat{\sigma}_{T+j}^{2(1)}$  and  $\hat{\sigma}_{T+j}^{2(2)}$ , by expected loss is the same whether the ranking is performed using the true variance,  $\sigma_t^2$ , or some unbiased volatility proxy,  $\hat{\sigma}_t^2$ . Patton (2011) showed that MSE and QLIKE type loss functions:

$$\text{MSE} : L(\hat{\sigma}_t^2, \hat{\sigma}_{T+j}^{2(i)}) = (\hat{\sigma}_{T+j}^{2(i)} - \hat{\sigma}_t^2)^2,$$

$$\text{QLIKE} : L(\hat{\sigma}_t^2, \hat{\sigma}_{T+j}^{2(i)}) = \frac{\hat{\sigma}_{T+j}^{2(i)}}{\hat{\sigma}_t^2} - \log \frac{\hat{\sigma}_{T+j}^{2(i)}}{\hat{\sigma}_t^2} - 1$$

are robust to the forecast error,  $\hat{\sigma}_{T+j}^{2(i)} - \hat{\sigma}_t^2$ , and the standardized forecast error,  $\hat{\sigma}_{T+j}^{2(i)}/\hat{\sigma}_t^2$ , respectively. We also use the Diebold and Mariano (1995) test for equal forecast accuracy regarding these two loss functions.

We use the RK estimates,  $\hat{\sigma}_t^2 = \exp(x_{T+j})$  ( $j = 1, \dots, F$ ), as the proxy of the true volatility. Fixing the sample size at 2048 for the rolling window, we re-estimated the model and computed the one step ahead forecasts of log-volatility for the last  $F = 500$  days. We select the optimal  $k$  each time for estimating the RSV–GGLM model, and we also update OLS estimates for the improved forecasts each time. Table 7 shows the values of MSE and QLIKE for the five approaches. Table 7 indicates the HAR model has the largest value of the loss functions for all cases. The RSV–GGLM and RSV–SLM models give higher values for the MSE and QLIKE than does the simple RSV model. One of the reasons for the result is that the fluctuations in volatility are higher for the estimation period, which covers the Global Financial Crisis, as compared with the forecasting period. However, the improved forecasts based on the RSV–GGLM and RSV–SLM models produce the smallest values. For the improved forecasts, the Diebold-Mariano test rejects the forecast of the HAR model at the five percent significance level in all cases except for the MSE of Nikkei.

The empirical results show that the data for S&P, FTSE, and Nikkei prefer the more flexible structure for long memory in log-volatility than the simple ARFIMA process. More specifically, the three data series favour the RSV–GGLM model than the RSV–SLM model, implying the preference of aperiodic long-range dependence. The results of the out-of-sample forecasts indicate that the RSV without long memory gives better forecasts than the RSV–GGLM and RSV–SLM models. However, the forecasts can be improved by combining the RSV–GGLM and RSV–SLM models.

## 5. Concluding remarks

In this paper, we considered a new realized stochastic volatility model with general Gegenbauer long memory (RSV–GGLM), which encompasses the new RSV model with seasonal long memory (RSV–SLM). We suggested a two-step estimator, in which the first step estimator gives the estimates of the location parameters of the Gegenbauer frequencies, which converges faster than the speed of  $T^{1/2}$ . The second step uses the Whittle likelihood (WL) estimation method, for which

the asymptotic distribution is the same as that of the quasi-maximum likelihood estimator when the location parameter is known. Then we conducted Monte Carlo experiments for investigating the finite sample properties of both estimators, and found that the first step estimator works satisfactorily, and that the finite sample bias for the WL estimator is negligible for  $T = 2048$ .

The estimation results for S&P, FTSE, and Nikkei indicate that the simple ARFIMA process for log-volatility is rejected, favoring the RSV–GGLM model. The forecasting results indicate that combining the forecasts of the RSV–SLM and RSV–GGLM models gives improved forecasts compared with the original ones. These results indicate that RSV models with general long memory are useful additions to the existing models in the literature.

**Acknowledgments**

The authors are most grateful to Yoshi Baba, Erricos Kontoghiorghes, the associate editor, and two anonymous reviewers for very helpful comments and suggestions. The first author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, [Japan Society for the Promotion of Science \(JSPS KAKENHI JP16K03603\)](#), and the Australian Academy of Science. The second author is most grateful for the financial support of the Australian Research Council, National Science Council, Ministry of Science and Technology (MOST), Taiwan, and the Japan Society for the Promotion of Science. The third author acknowledges the support from the Faculty of Economics at Soka University

**Appendix**

We explain the calculation of the coefficients of the  $MA(\infty)$  representation of the general Gegenbauer process in equation (3), and the calculation of the autocovariance functions.

Even for the simple Gegenbauer process with ARMA parameters, it is not easy to obtain explicit formulas for the coefficients for the  $MA(\infty)$  representation and the autocovariances that are valid for all lags. Recently, [Mc Elroy and Holan \(2012, 2016\)](#) developed a computationally efficient method for calculating these values. The spectral density of the general Gegenbauer process,  $h_t$ , can be written as:

$$f_h(\lambda) = \frac{\sigma_\eta^2}{2\pi} g_h(\lambda) [2 \sin(\lambda/2)]^{-2d} \prod_{l=1}^k [2(\cos(\lambda) - \cos(\omega_l))]^{-2d_l}, \quad -\pi < \lambda < \pi,$$

where  $g_h(\omega)$  is defined by (5). For convenience, we define  $\kappa(z)$  so that  $g(\lambda) = |\kappa(e^{-i\lambda})|^2$ . Then,  $\kappa(z)$  takes the form  $\kappa(z) = \prod_l (1 - \zeta_l z)^{p_l}$  for (possibly complex) reciprocal roots,  $\zeta_l$ , of the moving average and autoregressive polynomials, where  $p_l$  is one if  $l$  corresponds to a moving average root, and minus one if  $l$  corresponds to an autoregressive root. We consider  $d > \max\{d_l\}$ , as suggested by the empirical results in [Section 4](#). For the other case such that one of  $d_1, \dots, d_k$  takes the largest value instead of  $d$ , see [Mc Elroy and Holan \(2012\)](#).

Define:

$$g_j = 2 \sum_l \frac{p_l \zeta_l^j}{j},$$

$$\beta_j = \frac{2}{j} \left\{ d + 2 \sum_{l=1}^k d_l \cos(\omega_l j) \right\} + g_j,$$

$$\tilde{\psi}_j = \frac{1}{2j} \sum_{m=1}^l m \beta_m \tilde{\psi}_{j-m}, \quad \tilde{\psi}_0 = 1.$$

[Mc Elroy and Holan \(2012\)](#) showed that the  $MA(\infty)$  representation of (3) is given by:

$$h_{t+1} = \mu + \sum_{j=0}^{\infty} \tilde{\psi}_j \eta_{t-j},$$

and the autocovariances of  $h_t$  for  $l \geq 0$  are given by:

$$\gamma_l = \sigma^2 \sum_{j=0}^{J-1} \tilde{\psi}_j \tilde{\psi}_{j+l} + R_J(l),$$

where

$$R_J(l) = \sigma^2 \left\{ J^{-1+2d} \frac{F(1-d, 1-2d; 2-2d; -l/J)}{\Gamma^2(d)(1-2d)} \right\} \{1 + o(1)\},$$

and  $F(a, b; c; z)$  is the hypergeometric function evaluated at  $z$ . Note that  $\gamma_{-l} = \gamma_l$ . [Mc Elroy and Holan \(2012\)](#) recommend using the cutoff value  $J \geq 2,000$ . We set  $J = 20T$  with  $T = \{1024, 2048\}$  in this paper.

## References

- Abramovits, M., Stegun, N., 1970. *Handbook of Mathematical Functions*. Dover Publications, N.Y.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. *J. Am. Stat. Assoc.* 96, 42–55.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. *Econometrica* 71, 529–626.
- Artiach, M., Arteche, J., 2012. Doubly fractional models for dynamic heteroscedastic cycle. *Comput. Stat. Data Anal.* 56, 2139–2158.
- Asai, M., Chang, C.L., McAleer, M., 2017. Realized stochastic volatility with general asymmetry and long memory. *J. Econometr.* 199, 202–212.
- Asai, M., McAleer, M., Medeiros, M.C., 2012. Asymmetry and long memory in volatility modeling. *J. Financ. Econometr.* 10, 495–512.
- Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econometr.* 74, 3–30.
- Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76, 1481–1536.
- Bollerslev, T., Mikkelsen, H.O., 1996. Modeling and pricing long-memory in stock market volatility. *J. Econometr.* 73, 151–184.
- Bordignon, S., Caporin, M., Lisi, F., 2007. Generalised long-memory GARCH models for intra-daily volatility. *Comput. Stat. Data Anal.* 51, 5900–5912.
- Bordignon, S., Caporin, M., Lisi, F., 2009. Periodic long-memory GARCH models. *Econometr. Rev.* 28, 60–82.
- Breidt, F.J., Crato, N., de Lima, P., 1998. The detection and estimation of long memory. *J. Econometr.* 83, 325–348.
- Chan, K.S., Tsai, H., 2012. Inference of seasonal long-memory aggregate time series. *Bernoulli* 18, 1448–1464.
- Chung, C.F., 1996a. Estimating a generalized long memory process. *J. Econometr.* 73, 237–259.
- Chung, C.F., 1996b. A generalized fractionally integrated autoregressive moving-average process. *J. Time Ser. Anal.* 17, 111–140.
- Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *J. Financ. Econometr.* 7, 174–196.
- Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. *J. Busin. Econ. Stat.* 13, 253–263.
- Ding, Z., Granger, C.W.J., 1996. Modeling volatility persistence of speculative returns: A new approach. *J. Econ.* 73, 185–215.
- Dissanayake, G., Peiris, M.S., Proietti, T., 2018. Fractionally differenced Gegenbauer processes with long memory: a review. *Stat. Sci.* 33, 413–426.
- Dunsmuir, W., Hannan, E.J., 1976. Vector linear time series models. *Adv. Appl. Probab.* 8, 339–364.
- Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long-memory time series models. *J. Time Ser. Anal.* 4, 87–104.
- Granger, C.W.J., Ramanathan, R., 1984. Improved methods of combining forecasts. *J. Forecast.* 3, 197–204.
- Gray, H.L., Zhang, N., Woodward, W.A., 1989. On generalized fractional processes. *J. Time Ser. Anal.* 10, 233–257.
- Hansen, P.R., Huang, Z., 2016. Exponential GARCH modeling with realized measures of volatility. *J. Busin. Econ. Stat.* 34, 269–287.
- Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: a complete model of returns and realized measures of volatility. *J. Appl. Econometr.* 27, 877–906.
- Harvey, A., 1998. Long memory in stochastic volatility. In: Knight, J., Satchell, S. (Eds.), *Forecasting Volatility in Financial Markets*. Oxford: Butterworth-Haineman, pp. 307–320.
- Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. *Rev. Econ. Stud.* 61, 247–264.
- Hidalgo, J., Soulier, P., 2004. Estimation of the location and exponent of the spectral singularity of a long memory process. *J. Time Ser. Anal.* 25, 55–81.
- Hosoya, Y., 1997. A limit theory for long-range dependence and statistical inference on related models. *Ann. Stat.* 25, 105–137.
- Hsu, N.J., Tsai, H., 2009. Semiparametric inference of long-memory seasonality using generalized exponential models. *J. Stat. Pl. Inference* 139, 1992–2009.
- Koopman, S.J., Jungbacker, B., Hol, E., 2005. Forecasting daily variability of the s&p 100 stock index using historical realized and implied volatility measurements. *J. Empir. Finance* 12, 445–475.
- Koopman, S.J., Scharth, M., 2013. The analysis of stochastic volatility in the presence of daily realized measures. *J. Financ. Econometr.* 11, 76–115.
- Mc Elroy, T.S., Holan, S.H., 2012. On the computation of autocovariances for generalized Gegenbauer processes. *Stat. Sinica* 22, 1661–1687.
- Mc Elroy, T.S., Holan, S.H., 2016. Computation of the autocovariances for time series with multiple long-range persistencies. *Comput. Stat. Data Anal.* 101, 44–56.
- Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. *J. Econometr.* 160, 246–256.
- Peiris, S., Asai, M., 2016. Generalized fractional processes with long memory and time dependent volatility revisited. *Econometrics* 4 (3), 121.
- Pong, S., Shackleton, M.B., Taylor, S.J., Xu, X., 2004. Forecasting currency volatility: a comparison of implied volatilities and AR(FI)MA models. *J. Bank Finance* 28, 2541–2563.
- Porter-Hudak, S., 1990. An application of the seasonal fractionally differenced model to the monetary aggregates. *J. Am. Stat. Assoc.* 85, 338–344.
- Shirota, S., Hizu, T., Omori, Y., 2014. Realized stochastic volatility with leverage and long memory. *Comput. Stat. Data Anal.* 76, 618–641.
- Takahashi, M., Omori, Y., Watanabe, T., 2009. Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Comput. Stat. Data Anal.* 53, 2404–2426.
- Taniguchi, M., Kakizawa, Y., 2000. *Asymptotic Theory of Statistical Inference for Time Series*. Springer-Verlag, New York.
- Tsai, H., Rachinger, H., Lin, E.M.H., 2015. Inference of seasonal long-memory time series with measurement error. *Scand. J. Stat.* 42, 137–154.
- Velasco, C., Robinson, P.M., 2000. Whittle pseudo-maximum likelihood estimation for nonstationary time series. *J. Am. Stat. Assoc.* 95, 1229–1243.
- Whittle, P., 1952. Some results in time series analysis. *Skandinavisk Aktuarietidskrift* 35, 48–60.
- Woodward, W.A., Cheng, Q.C., Gray, H.L., 1998. A  $k$ -factor GARMA long memory model. *J. Time Ser. Anal.* 19, 485–504.
- Zaffaroni, P., 2009. Whittle estimation of EGARCH and other exponential volatility models. *J. Econometr.* 151, 190–200.