

## Article

# Generalized Fractional Processes with Long Memory and Time Dependent Volatility Revisited

M. Shelton Peiris <sup>1,\*</sup> and Manabu Asai <sup>2</sup><sup>1</sup> School of Mathematics and Statistics, The University of Sydney, Sydney 2006, Australia<sup>2</sup> Faculty of Economics, Soka University, Tokyo 192-8577, Japan; m-asai@soka.ac.jp

\* Correspondence: shelton.peiris@sydney.edu.au; Tel.: +612-9351-5764

Academic Editor: Kerry Patterson

Received: 15 January 2016; Accepted: 15 August 2016; Published: 5 September 2016

**Abstract:** In recent years, fractionally-differenced processes have received a great deal of attention due to their flexibility in financial applications with long-memory. This paper revisits the class of generalized fractionally-differenced processes generated by Gegenbauer polynomials and the ARMA structure (GARMA) with both the long-memory and time-dependent innovation variance. We establish the existence and uniqueness of second-order solutions. We also extend this family with innovations to follow GARCH and stochastic volatility (SV). Under certain regularity conditions, we give asymptotic results for the approximate maximum likelihood estimator for the GARMA-GARCH model. We discuss a Monte Carlo likelihood method for the GARMA-SV model and investigate finite sample properties via Monte Carlo experiments. Finally, we illustrate the usefulness of this approach using monthly inflation rates for France, Japan and the United States.

**Keywords:** GARMA; GARCH; stochastic volatility; long-memory; fractional differencing

**JEL Classification:** C18, C40, C58

## 1. Introduction

Consider the well-known ARFIMA( $p, d, q$ ) model given by:

$$\phi(B)Y_t = \theta(B)\epsilon_t, \quad (1)$$

where  $Y_t = (1 - B)^d X_t$ ,  $d \in (-1, 0.5)$ ,  $\{\epsilon_t\}$  is a sequence of uncorrelated (not necessarily independent) random variables, such that  $\text{Var}(\epsilon_t) = \sigma^2$ , and  $\phi(B)$  (stationary AR( $p$ )) and  $\theta(B)$  (invertible MA( $q$ )) polynomials respectively.

This standard case of constant variance innovations has been considered in many traditional time series analysis with applications. However, in recent years, there has been a great number of developments based on time-dependent instantaneous innovation variance (or volatility) such that  $\text{Var}(\epsilon_t) = \sigma_t^2$ . In particular, the following cases have been considered:

- (i)  $\sigma_t^2$  is a deterministic function of  $t$  or  $\sigma_t^2 = f(t)$ ,
- (ii)  $\epsilon_t$  follows the family of (G)ARCH process (see, [1,2]),
- (iii)  $\log(\sigma_t^2)$  is another stochastic process.

These cases (i) to (iii) can be analysed with emphasis on different practical issues. However, in applications, we need additional assumptions on  $\sigma_t^2$ , such as:

- (a)  $0 < m < \sigma_t^2 < M < \infty$  to ensure  $\text{Var}(Y_t)$  is finite, in (i)
- (b) stationarity or stability of both  $\{\epsilon_t^2\}$  and  $\log(\{\epsilon_t^2\})$  in (ii) and (iii)

Assumption (a) is imposed as it is natural to set bounds for the deterministic function  $\sigma_t^2 = f(t)$ . In fact, Assumption (b) is required when  $\sigma_t^2$  is a stochastic process. For case (i), the effect of non-stochastic and time-dependent instantaneous variance when  $d = 0$  was studied by Niemi [3] under the standard AR and MA regularity conditions on the zeros of  $\phi(B)$  and  $\theta(B)$ , respectively. In his work, Peiris [4] argues that the results of Niemi [3] can be extended to the ARFIMA family when  $d \in (-0.5, 0.5)$ . In (ii), it is known that  $X_t$  is strictly stationary if  $\epsilon_t$  is strictly stationary, and in particular,  $X_t$  is  $2m$ -th order stationary if  $\epsilon_t$  is  $2m$ -th order stationary. For the Integrated GARCH (IGARCH) model, Nelson [5] and Bougerol and Picard [6] have argued that IGARCH is strictly stationary under additional regularity conditions. For case (iii) with stochastic volatility, it is obvious that if  $\log \sigma_t^2$  is  $m$ -th order stationary, then  $\epsilon_t$  is  $2m$ -th order stationary. Therefore, it can be argued that when  $\sigma_t^2$  is a stationary stochastic process, the conditional likelihood estimation carried out by taking  $\sigma_t^2 = \sigma^2$  cannot be significant since  $\sigma_t^2$  is bounded. This would be useful in the estimation of parameters, especially in both cases (ii) and (iii). See the survey papers of McAleer [7] and Shephard [8] for the various extensions of (ii) and (iii).

An alternative way of modelling time-dependent volatilities has been extensively studied using ARMA models with time-dependent coefficients driven by constant variance innovations. See, for example, [4,9–13] and the references therein for details. However, this approach is not very attractive in applications, as it involves too many parameters to estimate.

Turning to applications in economic and financial time series, there are many popular directions of modelling and analysis of long-memory. Among others, the analysis of long-memory in inflation has been considered by Backus and Zin [14], Hassler and Wolters [15], Baillie, Chung and Tieslau [16], and Caporale and Gil-Alana [17]. In their paper, Delgado and Robinson [18] considered a number of methods for the analysis of long-memory time series using ARFIMA. An alternative and a general approach is to use the ARFIMA family with conditional and stochastic volatility as considered by Baillie et al. [19], Bollerslev and Mikkelsen [20], Ling and Li [21], Breidt et al. [22], Deo and Hurvich [23], and Bos et al. [24]. In their recent paper, Bos et al. [24] accommodate the stochastic volatility in ARFIMA modelling. Empirical evidence confirms that such models are very satisfactory in practice. Therefore, the aim of this paper is to extend the ARFIMA models with time-varying volatility to a general flexible class of time series models based on Gegenbauer polynomials together with the ARMA structure. The Gegenbauer ARMA (GARMA) model is a generalization of the ARFIMA model. Clearly, the former encompasses the latter as a special case. We will also extend the class of  $k$ -factor Gegenbauer process following Woodward et al. [25], Ferrara and Gueganand [26], and Caporale and Gil-Alana [17], by accommodating time-dependent volatility.

The organization of the paper is as follows. Section 2 reviews the family of GARMA with constant variance (volatility). Section 3 shows the existence and uniqueness of second-order solutions for the GARMA model with time-dependent volatility and develops new classes of GARMA-GARCH and GARMA-SV. Section 4 presents the asymptotic results for the maximum likelihood estimator for the GARMA-GARCH model and reports a Monte Carlo likelihood method for estimating the GARMA-SV model. Section 5 presents an illustrative example via simulation data, while Section 6 demonstrate an empirical example using inflation data in France, Japan and the United States. Section 7 gives concluding remarks.

## 2. Basic Results on GARMA with Constant Volatility

In this section, we review the family of GARMA processes with constant volatility. Based on the work of Gray et al. [27] and Chung [28], consider the family of time series generated by:

$$\phi(B)(1 - 2uB + B^2)^d X_t = \theta(B)\epsilon_t, \quad (2)$$

where the polynomials  $\phi(B)$  and  $\theta(B)$  are as defined before,  $|u| \leq 1$ ,  $|d| < 1$  are real parameters and  $\epsilon_t$  is white noise with zero mean and variance  $\sigma_\epsilon^2$ .

This process in (2) is known as Gegenbauer ARMA of order  $(p, d, q)$  or GARMA  $(p, d, q; u)$  and has the following properties:

- The power spectrum is given by:

$$f_X(\omega) = C(\omega) \times [4(\cos \omega - u)^2]^{-d}, \quad -\pi < \omega < \pi, \quad (3)$$

where  $C(\omega) = \frac{\sigma_\epsilon^2}{2\pi} \left| \frac{\theta(e^{-i\omega})}{\phi(e^{-i\omega})} \right|^2$  and  $i = \sqrt{-1}$ .

- The process in (2) is stationary with long-memory when  $|u| < 1$  and  $0 < d < 1/2$  or  $|u| = 1$  and  $d < 1/4$ . The long-memory features are characterized by:
  - hyperbolic decay of the autocorrelation function (ACF) superimposed with a sinusoidal,
  - unbounded spectrum at the Gegenbauer frequency,  $\omega = \omega_g = \cos^{-1}(u)$ .

In their recent papers, a modified class of generalized fractional processes has been studied by Shitan and Peiris [29,30].

Now, consider the following special case or the GARMA(0,  $d$ , 0;  $u$ ) process and its properties for later reference. That is, when  $\phi(B) = \theta(B) = 1$ , we have:

$$(1 - 2uB + B^2)^d X_t = \epsilon_t. \quad (4)$$

Suppose that the following regularity conditions are satisfied:

R1: AR regularity:  $|u| < 1$  and  $d < 1/2$  or  $|u| = 1$  and  $d < 1/4$ .

R2: MA regularity:  $|u| < 1$  and  $d > -1/2$  or  $|u| = 1$  and  $d > -1/4$ .

#### A Stationary Solution to GARMA(0, $d$ , 0; $u$ ) Model

Under the regularity conditions in R1, there exists the Wold representation to (4) given by:

$$X_t = \psi(B)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \quad (5)$$

where  $\psi(B) = (1 - 2uB + B^2)^{-d} = \sum_{j=0}^{\infty} \psi_j B^j$  with  $\psi_0 = 1$  and the Gegenbauer coefficients  $\psi_j$  have the explicit representation:

$$\psi_j = \sum_{q=0}^{[j/2]} \frac{(-1)^q (2u)^{j-2q} \Gamma(d-q+j)}{q!(j-2q)!\Gamma(d)}$$

such that  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$  ( $\Gamma(\cdot)$  is the Gamma function; see [31] for details). These coefficients  $\psi_j$ ,  $j \geq 2$  are recursively related by:

$$\psi_j = 2u \left( \frac{d-1+j}{j} \right) \psi_{j-1} - \left( \frac{2d-2+j}{j} \right) \psi_{j-2} \quad (6)$$

with initial values  $\psi_0 = 1$  and  $\psi_1 = 2du$ .

#### An Invertible Solution to GARMA(0, $d$ , 0; $u$ ) Model

Under the MA regularity conditions, there exists an invertible solution to (4) given by:

$$\epsilon_t = (1 - 2uB + B^2)^d X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad (7)$$

where  $\pi_j$ ,  $j \geq 0$  are obtained from (6) replacing  $d$  by  $-d$  with corresponding initial values.

The next section develops the class of GARMA(0,  $d$ , 0;  $u$ ) driven by time-dependent or stochastic innovations for later reference.

### 3. GARMA with Time-Dependent Innovations

Suppose that  $\text{Var}(\epsilon_t)$  in (4) is time-dependent. Consider the class of regular (in both AR and MA) GARMA(0,  $d$ , 0;  $u$ ) driven by time-dependent innovations satisfying:

$$\epsilon_t = \sigma_t z_t, \quad z_t \sim NID(0, 1), \quad (8)$$

where  $\sigma_t$  is the time-dependent volatility.

Below, we establish the existence and uniqueness of second-order solutions to (4) with innovations in (8) under certain additional regularity conditions.

#### 3.1. Unique Stable Solutions

We use the following general approach:

Let  $(\Omega, \mathcal{A}, \mathcal{P})$  be a probability space, and let  $L_0^r(\Omega, \mathcal{A}, \mathcal{P})$  be the space of all real-valued random variables on  $(\Omega, \mathcal{A}, \mathcal{P})$  with finite  $r$ -th order moments. Suppose that  $\{\zeta_t\}$  is sequence of random variables in  $L_0^r(\Omega, \mathcal{A}, \mathcal{P})$  and  $\mathcal{M}_s(\zeta)$  is the closed linear subspace of  $L_0^r(\Omega, \mathcal{A}, \mathcal{P})$  spanned by the elements  $\zeta_t, t \leq s$ . Let  $L_0^r(\Omega, \mathcal{A}, \mathcal{P})$  be the closed linear subspace spanned by all of the elements  $\zeta_t, t \geq 1$  by  $\mathcal{M}(\zeta)$ .

In the case of  $r = 2$ , let  $\xi$  and  $\zeta$  be any two random elements in  $L_0^2(\Omega, \mathcal{A}, \mathcal{P})$  such that the inner product and the norm satisfy

$$\langle \xi, \zeta \rangle = \text{Cov}(\xi, \zeta) \text{ and } \|\xi\|^2 = E(\xi^2)$$

respectively. It is easy to verify that  $L_0^2(\Omega, \mathcal{A}, \mathcal{P})$  is a Hilbert space.

Intuitively, if the volatility process is stationary, it guarantees the existence of the second moment of  $\epsilon_t$ , which enables us to establish the following two lemmas.

**Lemma 1.** *Under the AR regularity conditions R1, if the innovation process is stationary, then the solution in (5) belongs to  $L_0^2(\cdot)$ , i.e.,  $X_t \in L_0^2(\cdot)$ .*

**Proof.** Since the innovation process  $\epsilon_t$  is stationary,  $\text{var}(\epsilon_t) = \sigma_t^2$  is bounded. Hence the solution  $X_t = \psi(B)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \in L_0^2(\cdot)$ .  $\square$

**Lemma 2.** *Under the MA regularity conditions R2, if the innovation process is stationary, then an invertible solution to (7) belongs to  $L_0^2(\cdot)$  or  $\epsilon_t \in L_0^2(\cdot)$ .*

The proof is similar to that of Lemma 1.

**Lemma 3.** *Under the both AR and MA regularity conditions and stationarity of the innovation process, one has  $\mathcal{M}_t(X) = \mathcal{M}_t(\epsilon)$ .*

The proof to this follows from Lemmas 1 and 2.

Next, we consider two special cases useful in applications.

#### 3.2. Two Special Cases

Consider two popular cases where the innovations  $\epsilon_t$  follow GARCH or SV processes.

## 3.2.1. GARMA(0, d, 0; u)-GARCH(r, s)

Suppose that  $\sigma_t^2$  in (8) follows a GARCH(r, s) process, such that

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2, \quad (9)$$

where  $r$  and  $s$  are positive integers (both not simultaneously zero),  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  and  $\beta_j \geq 0$ . It is well-known that an equivalent representation of the GARCH(r, s) process is:

$$\epsilon_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^s \beta_j \epsilon_{t-j}^2 + \eta_t - \sum_{j=1}^s \beta_j \eta_{t-j},$$

where  $\eta_t = \epsilon_t^2 - \sigma_t^2 = (z_t^2 - 1)\sigma_t^2$  and are serially uncorrelated with mean zero. Now, we state the following lemma:

**Lemma 4.** Let  $\{X_t\}$  be generated by GARMA(0, d, 0; u) and  $\sigma_t^2$  follows (9).

- (a) If the AR regularity conditions and  $\sum_{i=1}^r \alpha_i + \sum_{j=1}^s \beta_j < 1$  are satisfied, then  $\{X_t\}$  is second-order stationary and  $X_t \in L_0^2(\cdot)$ .
- (b) If the MA regularity conditions are satisfied, then  $\{X_t\}$  is invertible and  $\epsilon_t \in L_0^2(\cdot)$ .

Proofs follow from Lemmas 1 and 2.

The next section develops the class of GARMA(0, d, 0; u) driven by SV innovations or GARMA-SV.

## 3.2.2. GARMA(0, d, 0; u)-SV

Suppose that  $h_t = \ln(\sigma_t^2)$  satisfies the following recursion

$$\rho(B)h_t = \kappa^* + \nu(B)\xi_{t-j}, \quad (10)$$

where  $\rho(B) = 1 - \rho_1 B - \rho_2 B^2 - \dots - \rho_l B^l$  and  $\nu(B) = 1 + \nu_1 B + \nu_2 B^2 + \dots + \nu_m B^m$ ,  $\kappa^*$  is a constant,  $\xi_t \sim NID(0, \sigma_\xi^2)$ , and the disturbances  $z_t$  and  $\xi_t$  are mutually independent for all  $t$ . Further assume that the roots of  $\rho(z) = 0$  and  $\nu(z) = 0$  lie outside the unit circle. Note that  $\sigma_\xi^2$  measures the conditional volatility of the log-volatility.

Let  $\Psi^*(B)\rho(B) = \nu(B)$ . Then, it is known that there exists a sequence  $\{\psi_j^*\}$ , such that

$$\Psi^*(B) = \sum_{j=0}^{\infty} \psi_j^* B^j$$

with  $\sum_{j=0}^{\infty} (\psi_j^*)^2 < \infty$ . Now, we have the following lemma:

**Lemma 5.** Under the AR regularity condition on  $\rho(\theta)$ , the log-volatility process in (10) has uniquely determined  $L_0^2(\cdot)$  solution, such that

$$h_t = \kappa + \sum_{j=0}^{\infty} \psi_j^* \xi_{t-j}, \quad (11)$$

where  $\kappa = E(h_t) = \frac{\kappa^*}{\rho(1)}$  is the mean of the log-volatility process.

It is clear from (11) that the log-volatility process  $\{h_t\}$  converges in the mean square to  $\kappa$  and:

$$E[(h_t - \kappa)^2] \rightarrow \sigma_\xi^2 \sum_{j=0}^{\infty} (\psi_j')^2 < \infty.$$

In their paper, Chesney and Scott [32] have considered the case where  $h_t$  follows an AR(1) process, such that

$$h_t = \kappa^* + \rho h_{t-1} + \xi_t, \quad (12)$$

where  $\kappa^*$  is a positive constant, and  $|\rho| < 1$ .

**Lemma 6.** *The process in (12) is equivalent to*

$$\ln(\epsilon_t^2) = \underline{\kappa} + \rho \ln(\epsilon_{t-1}^2) + w_t, \quad (13)$$

where  $\underline{\kappa} = a(1 - \rho) + \kappa^*$ ,  $a = E[\ln(z_t^2)] = -1.2704$ ,  $w_t = u_t - \rho u_{t-1} + \xi_t$  and  $u_t = \ln(z_t^2) - a$ . Note that  $E[w_t] = 0$ ;  $\text{Var}[w_t] = 4.93(1 + \rho^2) + \sigma_\xi^2$ ;  $E(w_t w_{t-1}) = -\rho$ ;  $E(w_t w_{t-j}) = 0$  ( $j \geq 2$ ).

**Proof.** Since  $\ln(\epsilon_t^2) = \ln(\sigma_t^2) + \ln(z_t^2)$ , we have  $\ln(\epsilon_t^2) = h_t + \ln(z_t^2)$ . Substituting for  $h_t = \ln(\epsilon_t^2) - a - u_t$  in (12) and noting that  $\text{Var}[\ln(z_t^2)] = 4.93$ , the lemma follows.  $\square$

**Remark.** Clearly,  $\{w_t\}$  in (12) is not a martingale difference series, and hence, it is not useful in applications. However, this can be written as an ARMA(1,1) in the form:

$$\ln(\epsilon_t^2) = \underline{\kappa} + \rho \ln(\epsilon_{t-1}^2) + v_t - \theta v_{t-1}, \quad (14)$$

where  $v_t \sim WN(0, \sigma_v^2)$ ;  $\theta$  and  $\sigma_v^2$  are given by  $\frac{1+\theta^2}{\theta} = \frac{(1+\rho^2)\sigma_u^2 + \sigma_\xi^2}{\rho\sigma_u^2}$ ,  $\sigma_u^2 = \text{Var}[\log(z_t^2)] = 4.93$ .

This result can be extended to any general ARMA structure for the log-volatility process.

**Lemma 7.** *Suppose that the SV process follows an ARMA( $l, m$ ) model as in (10). Then, the corresponding  $\{\ln(\epsilon_t^2)\}$  process satisfies an ARMA( $k, k$ ) in the form:*

$$\ln(\epsilon_t^2) = \underline{\kappa} + \rho_1 \ln(\epsilon_{t-1}^2) + \rho_2 \ln(\epsilon_{t-2}^2) + \cdots + \rho_k \ln(\epsilon_{t-k}^2) + v_t + \theta_1 v_{t-1} + \cdots + \theta_r v_{t-k},$$

where  $\underline{\kappa} = \kappa^* + a(1 - \rho_1 - \rho_2 - \cdots - \rho_p)$  and  $v_t = \xi_t - u_t$ ,  $v_{t-j} = \theta_j \xi_{t-j} - \rho_j u_{t-j}$  for  $j = 1, 2, \dots, k$  with  $k = \max(l, m)$ .

**Proof.** The proof follows from extending the approach in Lemma 6.  $\square$

Section 4 discusses the estimation of parameters.

#### 4. Estimation of Parameters

This section discusses the estimation for the GARMA( $p, d, q; u$ ) model with time-dependent volatility. We divide the section into two parts, namely, (i) the GARMA-GARCH and (ii) the GARMA-SV.

##### 4.1. GARMA-GARCH Model

Suppose that  $\mathbf{X} = \{X_1, \dots, X_n\}$  is generated by the GARMA( $p, d, q; u$ )-GARCH( $r, s$ ) model (2), (8) and (9). Define  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)'$ ,  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)'$ ,  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_r)'$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_s)'$ ,  $\boldsymbol{\gamma} = (d, \boldsymbol{\phi}', \boldsymbol{\theta}')$  and  $\boldsymbol{\delta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')$ . Let  $\boldsymbol{\lambda} = (u, \boldsymbol{\delta}', \boldsymbol{\gamma}')'$ , and let  $\boldsymbol{\lambda}_0 = (u_0, \boldsymbol{\delta}'_0, \boldsymbol{\gamma}'_0)'$  be the true value of  $\boldsymbol{\lambda}$  in the interior of the compact set  $\Lambda$ . The approximate log-likelihood function (excluding the constant) is given by

$$L(\boldsymbol{\lambda}) = \frac{1}{n} \sum_{t=1}^n l_t, \quad l_t = -\frac{1}{2} \log(\sigma_t^2) - \frac{\epsilon_t^2}{2\sigma_t^2},$$

where

$$\begin{aligned}\epsilon_t &= [\theta(B)]^{-1} \phi(B) (1 - 2uB + B^2)^d X_t, \\ \sigma_t^2 &= [\beta(B)]^{-1} [\alpha_0 + \alpha(B)\epsilon_t^2]\end{aligned}$$

and  $\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_r B^r$ ,  $\beta(B) = 1 - \beta_1 B - \beta_2 B^2 - \dots - \beta_s B^s$ .

Assuming the initial values of  $X_0, X_{-1}, X_{-2}, \dots$  are zero, an approximate maximum likelihood estimator  $\hat{\lambda}$  of  $\lambda$  in  $\Lambda$  is obtained by maximizing the above function, which is asymptotically equivalent to the maximum likelihood (ML) estimator. For the case of a non-normal distribution, we can still use the same approach with the corresponding quasi-maximum likelihood (QML) estimator.

For the ARFIMA  $(p, d, q)$ -GARCH  $(r, s)$  model, Breidt et al. [22] suggested the above approach. However, Ling and Li [21] established the consistency and asymptotic normality of the corresponding ML estimator and showed that the information matrix is block-diagonal. Combining the results of Ling and Li [21] and Chung [28,33], we can obtain the asymptotic results for the corresponding estimator of the GARMA-GARCH model.

**Proposition 1.** Let  $\hat{u}$  and  $(\hat{\gamma}, \hat{\delta})'$  be approximate ML estimators of  $u$  and  $(\gamma, \delta)$  based on a sample  $\{X_t\}_{t=1}^n$  from a GARMA-GARCH model under the conditions in Lemma 4. Then,  $\hat{u}$  is asymptotically independent of  $(\hat{\gamma}, \hat{\delta})'$  and:

$$n(\hat{u} - u_0) \xrightarrow{\mathcal{L}} \frac{K \sin(\omega_g)}{d} Y_0 \quad \text{if } |u| < 1 \text{ and } d \neq 0, \quad (15)$$

where  $K = E(\sigma_t^{-2}) + 2 \sum_{j=1}^{\infty} \varphi_{\sigma}(j) E(\epsilon_{t-j}^2 / \sigma_t^4)$ ,

$$Y_0 \equiv \frac{\int_0^1 \tilde{W}_1 dW_2 - \int_0^1 W_1 d\tilde{W}_2}{\int_0^1 \tilde{W}_1^2(r) dr + \int_0^1 W_1^2(r) dr}$$

and  $(\tilde{W}_1(t), \tilde{W}_2(t))$  and  $(W_1(t), W_2(t))$  are two independent Brownian motions with mean zero and covariance

$$t \begin{pmatrix} E(\sigma_t^2) & 1 \\ 1 & K \end{pmatrix}.$$

Furthermore,

$$n^2(\hat{u} \mp 1) \xrightarrow{\mathcal{L}} \pm \frac{K}{2d} Y_1 \quad \text{if } u = \pm 1 \text{ and } d \neq 0, \quad (16)$$

where  $Y_1$  is a random variable defined as

$$Y_1 \equiv \frac{\int_0^1 \left\{ \int_0^r W_1(s) ds \right\} dW_2(r)}{\int_0^1 \left\{ \int_0^r W_1(s) ds \right\}^2 dr}.$$

**Proof.** See Appendix A.2.  $\square$

As discussed in Chung [28,33], the convergence rates of  $\hat{u}$  are faster than those of the remaining parameters. The off-diagonal blocks in the information matrix (with respect to the parameter  $u$  and the remaining parameters) approach zero. Hence, the distribution of  $\hat{u}$  is asymptotically independent of the remaining parameters.

Below, we report the asymptotic result of the remaining parameters.

**Proposition 2.** Based on the sample  $\{X_t\}_{t=1}^n$  from the GARMA-GARCH model and under the conditions in Lemma 4, we have

$$\sqrt{n} \begin{pmatrix} \hat{\gamma} - \gamma_0 \\ \hat{\delta} - \delta_0 \end{pmatrix} \xrightarrow{\mathcal{L}} N \left( \mathbf{0}, \begin{bmatrix} \Omega_{\gamma}^{-1} & O \\ O & \Omega_{\delta}^{-1} \end{bmatrix} \right),$$



where  $u \neq 1$  and

$$\begin{aligned}\Omega_{\gamma} &= E \left( \frac{1}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial \gamma} \frac{\partial \epsilon_t}{\partial \gamma} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma} \right), \\ \Omega_{\delta} &= E \left( \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma} \right).\end{aligned}\quad (17)$$

**Proof.** See Appendix A.3.  $\square$

For the case of  $u$  to be one, we can use the asymptotic result of Ling and Li [21].

Following Gray et al. [27] and Chung [28], in practice, we use the grid search procedure for different value of  $u$  over the range  $[-1, 1]$  to minimize the likelihood function. For selecting the order of the GARMA-GARCH model, we can use an information criterion, such as AIC and BIC. Furthermore, we can use the conventional  $t$  test for the parameters except for  $u$ . For testing the null hypothesis regarding  $u$ , we can use the approach of Chung [33] to obtain percentiles via simulations.

#### 4.2. Estimation of GARMA-SV Model

Suppose that  $\mathbf{X} = \{X_1, \dots, X_n\}$  are generated from the GARMA  $(p, d, q; u)$ -SV model (2), (8) with the SV structure

$$h_t = \kappa + \rho(h_{t-1} - \kappa) + \xi_t, \quad |\rho| < 1, \quad (18)$$

where  $h_t = \ln(\sigma_t^2)$ ,  $\{\xi_t\} \sim NID(0, \sigma_{\xi}^2)$  and  $\xi_t, z_t$  are independent.

From (18), we have  $E(h_t) = \kappa$  and  $Var(h_t) = \frac{\sigma_{\xi}^2}{1-\rho^2}$ . Further,  $Var(\epsilon_t) = \exp\left(\frac{\sigma_{\xi}^2}{2(1-\rho^2)}\right)$ . Since  $\{h_t\}$  is a latent process, the evaluation of the likelihood function requires integrating it with respect to  $(h_1, \dots, h_n)$ .

As mentioned earlier, the evaluation of the likelihood function involves high-dimensional integration, which is difficult to calculate. Nevertheless, among others, Danielsson and Richard [34], Shephard and Pitt [35], Durbin and Koopman [36,37] and Liesenfeld and Richard [38] suggested evaluating high-dimensional integrals using simulation methods and then maximizing the corresponding likelihood function. In this case, we use the Monte Carlo likelihood (MCL) estimator of Durbin and Koopman [37]. In their recent paper, Bos et al. [24] extended the MCL estimator for the estimation of parameters of ARFIMA-SV parameters. This approach creates a set of realized values for  $\mathbf{h} = (h_1, h_2, \dots, h_n)'$  by 'importance sampling'. Conditional on  $\mathbf{h}$  and using the prediction error decomposition, it is clear that the density is given by:

$$\log(\mathbf{X}|\mathbf{h}, d, \boldsymbol{\phi}, \boldsymbol{\theta}, \kappa, \rho, \sigma_{\xi}^2) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \log(f_t) - \frac{1}{2} \sum_{t=1}^n \frac{a_t^2}{f_t}, \quad (19)$$

where  $a_t$  is the one-step ahead prediction error,  $f_t$  is its variance and  $\mathbf{h}$  is from importance sampling.

From (19), we evaluate the simulated likelihood function based on the true density and the importance density using the results of Durbin and Koopman [37]. We also extend the work of Bos et al. [24], by replacing the autocovariance functions of the ARFIMA by those of the GARMA and use the results of McElroy and Holan [39] to obtain the ML estimator. As a practical issue, we use the grid search procedure for different values of  $u$  over the range  $[-1, 1]$  for minimizing the likelihood function.

Noting that  $\partial h_t / \partial u = 0$ ,  $\partial h_t / \partial \gamma = \mathbf{0}$ , we can consider that the information matrix of  $(u, \gamma', \delta_2')'$  has a block diagonal structure similar to the GARMA-GARCH case, where  $\delta_2 = (\kappa, \sigma_{\xi}^2, \rho)'$ . If the MCL approximates the true likelihood accurately, we can use the conventional  $t$  test for the parameters except for  $u$ . For testing the null hypothesis regarding  $u$ , we can use the approach of Chung [33] to



obtain percentiles via simulations, based on Proposition 1. For this purpose, we need to replace  $K$  in Proposition 1 with  $K^* = E[\exp(-2h_t)]$ , by noting  $\partial h_t / \partial u = 0$  unlike the GARMA-GARCH process.

In the next section, we will investigate the finite sample properties of the MCL estimator. In selecting the order of the GARMA-SV model, we can use the MCL to calculate the information criterion, such as AIC and BIC.

It is worth mentioning a semi-parametric estimation procedure for the long-memory parameter under heteroskedasticity. See, for example, [40]. For the GARMA-SV model, we extend this in two directions: one is to replace the conditional heteroskedasticity by SV, and the other is to extend the ARFIMA to the GARMA. For the latter case regarding the GARMA process, Hidalgo and Soulier [41] developed a log-periodogram regression estimator extending the work of Robinson and Henry [40].

Now, look at a simulation study in order to illustrate certain properties of this GARMA-SV class.

## 5. Simulation Results

In this section, we show two kinds of simulation results for the GARMA-SV model. One is an illustrative example to show the pattern of a GARMA-SV process, and the other is the Monte Carlo results for the finite sample properties of the ML estimator for the GARMA-SV.

### 5.1. An Illustrative Example

We generate a one-shot series to show as an illustrative example. Simulate 500 values from GARMA(0,  $d$ , 0;  $u$ )-SV from (4) and (8) with SV innovations as in (15). We have selected the parameter values  $(u, d) = (0.7, 0.3)$  for the following two cases:

- Case (i) with constant volatility:  $(\kappa, \rho, \sigma_\xi) = (0, 0, 0)$ ,
- Case (ii) with stochastic volatility:  $(\kappa, \rho, \sigma_\xi) = (0, 0.98, 0.2)$ .

Take the time series of 300 observations (discarding the first 200) for further analysis. It is clear that the Gegenbauer frequency is  $\omega_G = 0.2532\pi = 0.7954$ .

For the simulated GARMA(0,  $d$ , 0;  $u$ ), Figure 1 shows the time series plot, the sample periodogram and the autocorrelation functions (ACF) of series. In addition, 95% confidence intervals for the squared series are presented. The periodogram indicates that the peak is at  $\omega = 0.2733\pi$ , which is close to the actual value. The ACFs of the raw series decay hyperbolically and periodically, confirming the features of the Gegenbauer process. On the other hand, the ACFs of the squared series are insignificant for higher lags.

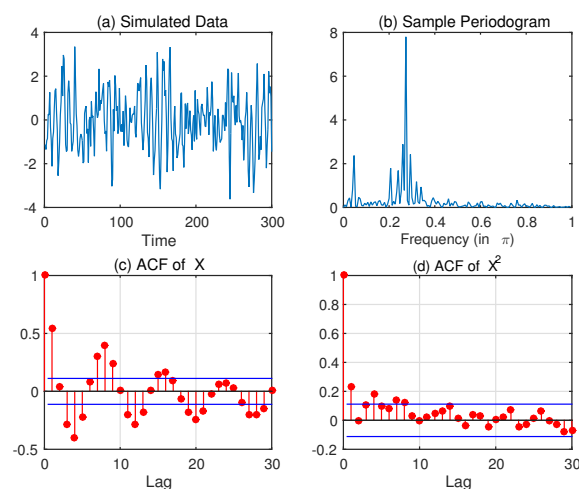
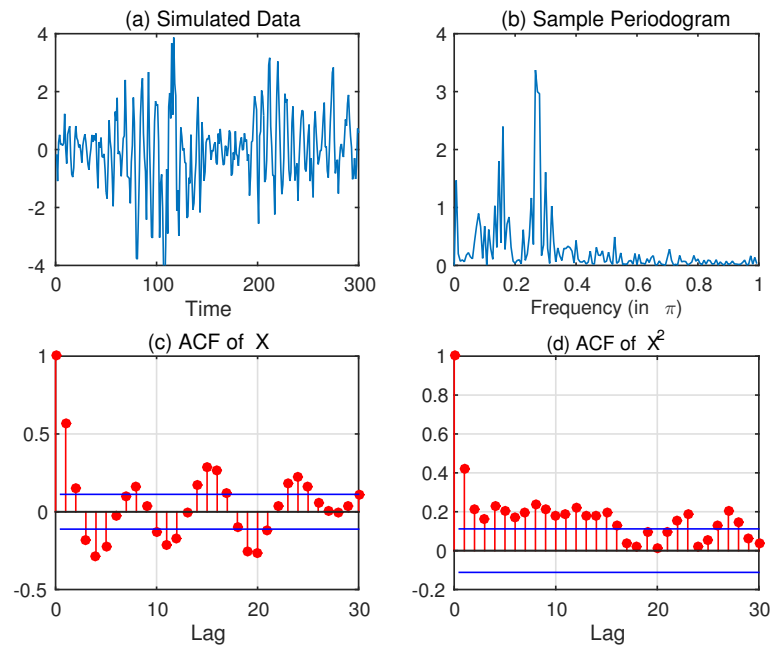


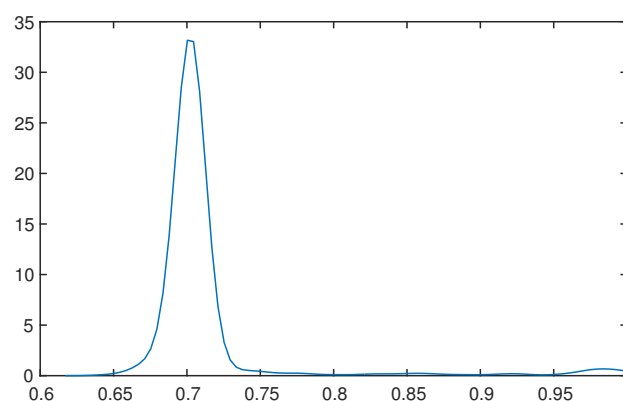
Figure 1. Simulated GARMA(0,  $d$ , 0;  $u$ ) process.

Figure 2 presents the plots for a simulated GARMA(0,  $d$ , 0;  $u$ )-SV process. The time series plot indicates the pattern of the time-dependent volatility. Figure 2b shows three major peaks at frequencies of  $0.0467\pi$ ,  $0.1600\pi$  and  $0.2733\pi$ . Clearly, the last peak is close to the true Gegenbauer frequency. Figure 2c shows that the ACF of the raw series decays hyperbolically and periodically (similar to the result in Figure 1c of the Gegenbauer process). The ACF in Figure 2d shows a geometric decay, which is different from Figure 1d, which shows constant volatility.



**Figure 2.** Simulated GARMA(0,  $d$ , 0;  $u$ )-stochastic volatility (SV) process.

As shown in Figure 1b, the periodogram of the GARMA process has one major peak at the Gegenbauer frequency, which is further from the origin. These plots show that there is a possibility of having multiple peaks under the time-dependent volatility model. The next section gives an empirical study based on a monthly consumer price index as an example.



**Figure 3.** Estimated density of the MCL estimator of  $u$ .

## 5.2. Monte Carlo Experiments

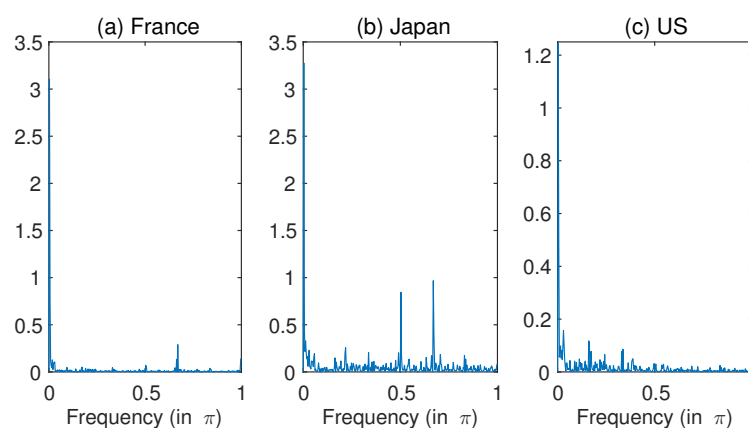
In order to examine the finite sample properties of the MCL estimator for the GARMA-SV model, we conduct a simulation study. For the ARFIMA-SV model, Bos et al. [24] conducted Monte Carlo experiments for the MCL estimator for the ARFIMA  $(1, d, 0)$ -SV model, and found (i) downward bias in  $d$ , (ii) upward bias in  $\phi$ , (iii) downward bias in  $\rho$  and (iv) upward bias in  $\sigma_{\xi}^2$ , with the sample size  $T = 500$ . We consider the GARMA  $(1, d, 0; u)$ -SV model in order to check the effects of the additional parameter  $u$ . As explained in the previous section, we use the grid-search for estimating  $u$  over the interval  $[-1, 1]$ .

We specify the parameters for the GARMA part as  $(\mu, d, u, \phi) = (0.1, 0.4, 0.7, 0.384)$ , implying that the value of the first-order autocorrelation function is 0.6. For the parameters for the SV part, we use the values  $(\kappa, \rho, \sigma_{\xi}^2) = (-7.36, 0.95, 0.26)$ , which are used in the simulation experiments of Jacquier et al. [42]. We consider a sample size of  $T = 500$  with 2000 replications.

Table 1 shows the sample means, standard deviations and root mean squared errors (RMSE) of the MCL estimators. For the parameter  $u$ , the bias and RMSE are smaller than those for  $d$  and  $\phi_1$ , implying a faster convergence rate for the MCL estimator of  $u$ , as expected from Theorem 1 of Chung [33], for the simple GARMA process and from Proposition 1 of the current paper for the GARMA-GARCH model. Figure 3 shows the estimated density of the  $\hat{u}$ , indicating that 94% of the estimates are located between (0.65, 0.75) and that 2.8% of them are greater than 0.95. For the remaining parameters, Table 1 agrees with the simulation results of Bos et al. [24], except for  $\phi_1$ . The MCL estimator for  $\phi_1$  gives a downward bias, which may be explained by the fact that not only  $(d, \phi_1)$ , but also  $u$ , affect the autocorrelation structure of the GARMA part.

**Table 1.** Finite sample performance of the Monte Carlo likelihood (MCL) estimator for  $T = 500$ .

Parameters	True	Mean	SD	RMSE
$u$	0.700	0.7137	0.0530	0.0547
$d$	0.400	0.2317	0.0798	0.1862
$\phi_1$	0.384	0.0320	0.0688	0.3587
$\mu$	0.100	0.0997	0.0077	0.0077
$\kappa$	−7.360	−6.9968	0.4034	0.5425
$\rho$	0.950	0.9243	0.0412	0.0485
$\sigma_{\xi}^2$	0.260	0.2984	0.0795	0.0882



**Figure 4.** Sample periodograms of inflation rates. The inflation rates are seasonally adjusted via monthly dummy variables.

## 6. Applications of GARMA-SV Models

We use the monthly consumer price index (CPI)  $P_t$  in three countries: France, Japan and the United States (U.S.) to illustrate the modelling described in this paper. The sample period starts

from 1960M11 to 2015M11 in all three countries. The data source is the IMF's International Financial Statistics published on the IMF web page. The CPIs are normalized at 100 for the year 2010. From the price index series, we calculate the monthly inflation rates,  $\Pi_t = 100 \times \log(P_t/P_{t-1})$  giving  $T = 660$  observations. By removing the seasonal effects, regress the inflation rates on a series of monthly dummies  $D_t$  as  $\Pi_t = D_t\delta + e_t$ . Define  $X_t = \bar{\Pi} + \hat{e}_t$ , where  $\bar{\Pi}$  is the average of inflation rates and  $\hat{e}_t$  is the residuals through the method of OLS. Table 2 shows the corresponding descriptive statistics. For seasonally-adjusted inflation rates, Figure 4 shows the estimated periodograms. As in Figure 4, the highest peak is at  $\omega = 0.0033\pi$  for all three countries. Apart from the pole, there are at least two small peaks. The second highest peak is at  $\omega = 0.6697\pi$  for France and Japan, while the peak in the second mass is at  $\omega = 0.1606\pi$  for the U.S. Although it is difficult to distinguish, the third highest peak for France is at  $\omega = 0.9970\pi$ . Based on the ACF and spectral properties, we examine the reasons for these multiple peaks. The first candidate is time-dependent volatility as discussed in the previous section. The second one is the general GARMA process with multi-factors, and the third is a nested model of the multi-factor GARMA and the SV.

**Table 2.** Descriptive statistics of inflation rates.

Variable	Average	SD	Min	Max
France				
$\Pi_t$	0.3507	0.3940	−1.0072	1.9147
$X_t$	0.3507	0.3853	−0.9171	2.0047
Japan				
$\Pi_t$	0.2580	0.6766	−1.1940	4.2182
$X_t$	0.2580	0.6147	−1.0172	4.1818
US				
$\Pi_t$	0.3154	0.3542	−1.9339	1.7924
$X_t$	0.3154	0.3365	−1.7434	1.7659

Note:  $\Pi_t$  denotes the original inflation rate, while  $X_t$  denotes the seasonally-adjusted series.

In this case, we estimate the three-factor GARMA-SV model given by:

$$(1 - \phi L)(1 - 2u_1B + B^2)^{d_1}(1 - 2u_2B + B^2)^{d_2}(1 - 2u_3B + B^2)^{d_3}X_t = \epsilon_t,$$

based on (8) and (18) and using the MCL estimation technique. By the sample periodogram in Figure 4, we expect that one of the three Gegenbauer frequencies is zero, corresponding to the highest peak at the point close to the origin. We do not report the higher orders of  $(p, q)$ , as the estimates were insignificant (using the mean subtracted series following the recommendation of Chung [28]). The estimates of  $d_i$  are located in  $(0.05, 0.2)$  and are significant at five percent, except for  $d_2$  for the U.S. series. For all three countries, the null hypothesis of  $u_1 = 1$  cannot be rejected. A positive value of  $u_i$  indicates that the Gegenbauer frequency is close to the origin, while a negative value implies that it is close to  $\pi$ . Hence, the Gegenbauer frequency for U.S. data is close to the origin, while the estimates of  $(u_2, u_3)$  of the other two countries are close to  $\pi$ . It is interesting to note that the estimates of  $\rho$  are positive and significant, indicating the appropriateness of accommodating the time-dependent volatility structure in this multi-factor GARMA model. The empirical results support this three-factor GARMA-SV model for France and Japan, while the U.S. series favours the two-factor GARMA-SV model.

**Table 3.** QML estimates of the multi-factor GARMA-SV model.

Parameter	France		Japan		U.S.	
$\kappa$	−3.4302	(0.1799)	−12.2791	(0.7865)	−4.0358	(0.2580)
$\sigma_{\xi}^2$	0.3210	(0.1489)	0.0670	(0.0243)	0.1336	(0.0637)
$\rho$	0.8837	(0.0924)	0.9969	(0.0033)	0.9723	(0.0233)
$\sigma_{\varepsilon}$	0.2524	(0.0068)	0.5200	(0.0139)	0.2485	(0.0067)
$\phi$	0.1585	(0.2106)	−0.0882	(0.1318)	−0.0703	(0.2031)
$d_1$	0.1971	(0.0683)	0.1769	(0.0245)	0.1696	(0.0244)
$d_2$	0.1701	(0.0332)	0.1516	(0.0320)	0.0639	(0.0337)
$d_3$	0.0719	(0.0230)	0.1109	(0.0287)	0.0870	(0.0338)
$u_1$	1.0000		0.9999		1.0000	
	[0.9997,1.0000]		[0.9995,1.0000]		[0.9996,1.0000]	
$u_2$	−0.5082		−0.5083		0.8754	
	[−0.5295,−0.4869]		[−0.5322,−0.4844]		[0.8435,0.9072]	
$u_3$	−0.9996		−0.0095		0.4917	
	[−1.0013,−0.9979]		[−0.0474,−0.0284]		[0.4496,0.5339]	

Note: Standard errors are in parenthesis. For  $(u_1, u_2, u_3)$ , 95% confidence intervals are in brackets.

Since the estimates of  $\rho$  are close to one for Japan and the U.S. series, we consider a GARMA model for the SV process,  $\{h_t\}$ . Using the MCL estimates from Table 3, we obtain the residuals to satisfy

$$\hat{\varepsilon}_t = (1 - 2\hat{u}_3B + B^2)^{\hat{d}_3}(1 - 2\hat{u}_2B + B^2)^{\hat{d}_2}(1 - 2\hat{u}_1B + B^2)^{\hat{d}_1}(1 - \hat{\phi}L)X_t.$$

Now, we consider the interpretation of the parameter  $u_i$ . First of all, the behaviour of a periodic long-memory process is different from a periodic short memory process (for example, a seasonal ARMA). As shown by Chung [28,33], the  $j$ -th autocovariance function for the GARMA process is approximated by  $Kj^{2d-1}\cos(\omega_g j)$ , where  $K$  is a constant. Hence, the operator  $(1 - u_iB + B^2)^{d_i}$  produces a periodic long-memory with cycles of every  $2\pi/\arccos(\omega_{g,i})$ . For France, the values for  $i = 2, 3$  are 2.986 and 2.018, respectively. These results indicate that there are periodic long-range dependences with respect to every two and three months. The values of Japan are 2.986 and 3.976, implying the periodic long-memory with respect to every three and four months. For the U.S., the estimate for  $u_3$  is 5.946, producing the periodic long-memory with respect to every six months. As  $u_1$  is equal to one in three countries, there is no periodic long-memory for the first factor.

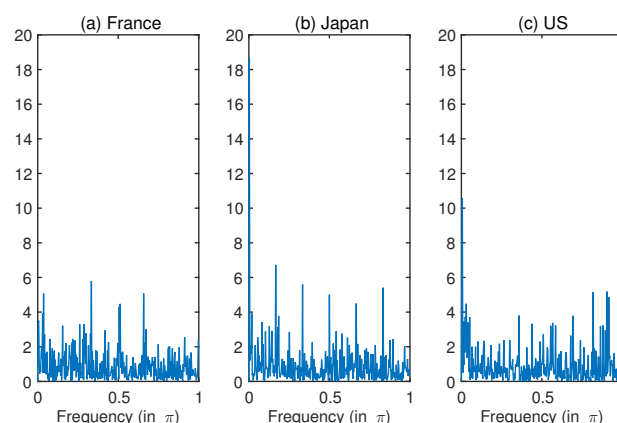
**Figure 5.** Sample periodograms of the log of squared residuals.

Figure 5 shows the periodograms for the log of the squared residuals,  $\log \hat{\varepsilon}_t^2$ , which can be considered as the proxy of log-volatility. Figure 5 shows that there is a distinguished peak at zero frequency for the case of Japan and the U.S., implying that a short memory model is adequate

for France, while long-memory models are appropriate for the remaining two countries. Furthermore, Japan also has five other periodic peaks, implying a multiple periodic long-memory. For estimating the GARMA-SV model when  $h_t$  follows a GARMA model, the MCL technique requires a smoother simulation to be applied to the second GARMA process, which is an extension of the work of de Jong and Shephard [43]. Another task is to reduce the computational time under the multiple grid-search method for finding optimal values of  $u_i$ 's in mean and volatility. We need to wait for further research for these problems.

In this section, we found that in addition to non-periodical long-range dependence, the empirical results indicate the existence of the periodic long memory under the time-dependent volatility.

## 7. Conclusions

This paper considers the class of GARMA-GARCH and GARMA-SV models in detail. We have established the existence and uniqueness of second order solutions and investigated the asymptotic results of the approximated ML estimator for the GARMA-GARCH model. We have explained the Monte Carlo likelihood (MCL) technique for estimating the GARMA-SV model. Using a set of simulated data, we have presented an illustrative example with possible multiple peaks in the GARMA-SV model. We also conducted a Monte Carlo experiment to investigate finite sample properties of the MCL estimator. In the empirical analysis, we have estimated a three-factor GARMA-SV model for monthly inflation rates of France, Japan and the United States and found that the data favour multi-factor GARMA-SV models showing that periodic long-memory exists.

There are many possible extensions of this new approach. In a future paper, we will develop models allowing both the mean and variance to follow Gegenbauer processes, as discussed in the previous section. For this purpose, we may extend the MCL estimation method by considering the simulation smoother for the GARMA process. By extending the work of Robinson and Henry [40], we plan to develop an estimation method for the Gegenbauer frequency and the long-memory parameter in the presence of long-memory in volatility.

**Acknowledgments:** The authors are most grateful to the editor, Kerry Patterson, and five anonymous reviewers for very helpful comments and suggestions. The second author acknowledges the financial supports of Zengin Foundation for Studies on Economics and Finance and the Japan Society for the Promotion of Science (JSPS Grant Number JP16K03603). The paper was written while the first author was visiting at the Faculty of Economics, Soka University.

**Author Contributions:** Both authors equally contributed to the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Proofs of Propositions

### Appendix A.1. Preliminary Results

To obtain the ML estimator  $\hat{\lambda}$ , we need to find the first-order derivatives and the information matrix. For each  $t$ , we obtain

$$\begin{aligned}\frac{\partial l_t}{\partial u} &= \frac{1}{2\sigma_t^2} \left( \frac{\epsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial u} - \frac{\epsilon_t}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial u}, \\ \frac{\partial l_t}{\partial \gamma} &= \frac{1}{2\sigma_t^2} \left( \frac{\epsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial \gamma} - \frac{\epsilon_t}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial \gamma}, \\ \frac{\partial l_t}{\partial \delta} &= \frac{1}{2\sigma_t^2} \left( \frac{\epsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial \delta},\end{aligned}\tag{A1}$$

where

$$\begin{aligned}
 \frac{\partial \epsilon_t}{\partial u} &= [1 - 2uB + B^2]^{-1}(-2dB)\epsilon_t = -2d \sum_{j=1}^{\infty} a_j \epsilon_{t-j}, \\
 \frac{\partial \epsilon_t}{\partial d} &= \log(1 - 2uB + B^2) \cdot \epsilon_t = -2 \sum_{j=1}^{\infty} \frac{\cos(j\omega_g)}{j} B^j \epsilon_t, \\
 \frac{\partial \epsilon_t}{\partial \phi_j} &= -[\phi(L)]^{-1} \epsilon_{t-j}, \\
 \frac{\partial \epsilon_t}{\partial \theta_j} &= -[\theta(L)]^{-1} \epsilon_{t-j}, \\
 \frac{\partial \sigma_t^2}{\partial \delta} &= \mathbf{b}_t + \sum_{i=1}^s \beta_i \frac{\partial \sigma_{t-i}^2}{\partial \delta}, \\
 \frac{\partial \sigma_t^2}{\partial \gamma} &= 2 \sum_{i=1}^r \alpha_i \epsilon_{t-i} \frac{\partial \epsilon_{t-i}}{\partial \gamma} + \sum_{i=1}^s \beta_i \frac{\partial \sigma_{t-i}^2}{\partial \gamma}, \\
 \frac{\partial \sigma_t^2}{\partial u} &= 2 \sum_{i=1}^r \alpha_i \epsilon_{t-i} \frac{\partial \epsilon_{t-i}}{\partial u} + \sum_{i=1}^s \beta_i \frac{\partial \sigma_{t-i}^2}{\partial u},
 \end{aligned} \tag{A2}$$

with

$$a_j = \begin{cases} j & \text{if } u = 1, \\ \sin(j\omega_g) / \sin(\omega_g) & \text{if } |u| < 1, \\ (-1)^{j-1} j & \text{if } u = -1, \end{cases}$$

and  $\mathbf{b}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-r}^2, \sigma_{t-1}^2, \dots, \sigma_{t-s}^2)'$ . Note that only  $n$  observations are available. However,  $\epsilon_t$ ,  $\sigma_t^2$ ,  $\partial \epsilon_t / \partial u$ ,  $\partial \epsilon_t / \partial \gamma$ ,  $\partial \sigma_t^2 / \partial u$ ,  $\partial \sigma_t^2 / \partial \gamma$  and  $\partial \sigma_t^2 / \partial \delta$  all depends on the theoretically infinite past history of  $\{X_t\}$  or  $\{\epsilon_t\}$ . For simplicity, we assume that the pre-sample values of  $\{X_t\}$  and  $\{\epsilon_t\}$  are zero and choose the pre-sample estimates of  $\sigma_t^2$  and  $\epsilon_t^2$  to be  $\sum_{t=1}^n \epsilon_t^2 / n$ , as in Bollerslev [2], Baillie et al. [16], Ling and Li [21] and Weiss [44]. This will not affect the asymptotic efficiency and other asymptotic properties.

We obtain the second derivatives, by directly differentiating  $\partial l_t / \partial \gamma$ ,  $\partial l_t / \partial u$  and  $\partial l_t / \partial \delta$ . For using later, we only give:

$$\frac{\partial^2 l_t}{\partial u^2} = -\frac{1}{\sigma_t^2} \left( \frac{\partial \epsilon_t}{\partial u} \right)^2 - \frac{\epsilon_t^2}{2\sigma_t^6} \left( \frac{\partial \sigma_t^2}{\partial u} \right)^2 + \left( \frac{\epsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial}{\partial u} \left[ \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial u} \right] + \frac{2\epsilon_t}{\sigma_t^4} \frac{\partial \epsilon_t}{\partial u} \frac{\partial \sigma_t^2}{\partial u} - \frac{\epsilon_t}{\sigma_t^2} \frac{\partial^2 \epsilon_t}{\partial u^2} \tag{A3}$$

Let  $\mathcal{F}_t$  be the past information of  $\epsilon_t$  available up to time  $t$ . Then, we obtain the conditional expectation of the second derivatives:

$$\begin{aligned}
 E \left( \frac{\partial^2 l_t}{\partial u^2} \middle| \mathcal{F}_{t-1} \right) &= -\frac{1}{\sigma_t^2} \left( \frac{\partial \epsilon_t}{\partial u} \right)^2 - \frac{1}{2\sigma_t^4} \left( \frac{\partial \sigma_t^2}{\partial u} \right)^2, \\
 E \left( \frac{\partial^2 l_t}{\partial \gamma \partial \gamma'} \middle| \mathcal{F}_{t-1} \right) &= -\frac{1}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial \gamma} \frac{\partial \epsilon_t}{\partial \gamma'} - \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'}, \\
 E \left( \frac{\partial^2 l_t}{\partial \delta \partial \delta'} \middle| \mathcal{F}_{t-1} \right) &= -\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'},
 \end{aligned} \tag{A4}$$



and those of the cross partial derivatives,

$$\begin{aligned} E \left( \frac{\partial^2 l_t}{\partial \gamma \partial u} \middle| \mathcal{F}_{t-1} \right) &= -\frac{1}{\sigma_t^2} \frac{\partial \epsilon_t}{\partial \gamma} \frac{\partial \epsilon_t}{\partial u} - \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial u}, \\ E \left( \frac{\partial^2 l_t}{\partial \delta \partial u} \middle| \mathcal{F}_{t-1} \right) &= -\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial u}, \\ E \left( \frac{\partial^2 l_t}{\partial \gamma \partial \delta'} \middle| \mathcal{F}_{t-1} \right) &= -\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \delta'}. \end{aligned}$$

Below, we will deal with  $u$  and  $(\gamma', \delta')'$  separately, since their speeds of convergence are different, as shown by Chung [33].

Noting that

$$E \left[ \left( \frac{\partial \epsilon_t}{\partial d} \right)^2 \right] = 2 \left( \frac{\pi^2}{3} - \pi \omega_g + \omega_g^2 \right) E(\epsilon_t^2) < \infty,$$

we apply the proof of Theorem 3.1 of Ling and Li (1997) for  $(\gamma', \delta')'$  to obtain the asymptotic properties of the information matrix:

$$-\frac{1}{n} \sum_{t=1}^n \begin{pmatrix} \partial^2 l_t / \partial \gamma \partial \gamma' & \partial^2 l_t / \partial \gamma \partial \delta' \\ \partial^2 l_t / \partial \delta \partial \gamma' & \partial^2 l_t / \partial \delta \partial \delta' \end{pmatrix} \xrightarrow{\text{a.s.}} \begin{pmatrix} \Omega_\gamma & O \\ O & \Omega_\delta \end{pmatrix}, \quad (\text{A5})$$

as  $n \rightarrow \infty$ , and  $\Omega_\gamma$  and  $\Omega_\delta$  are positive definite matrices, where  $\Omega_\gamma$  and  $\Omega_\delta$  are defined in (17). We can show that the information matrix is a block diagonal, in the following way. We can express each element of  $\partial \sigma_t^2 / \partial \delta$  as the infinite sum  $\epsilon_{t-i}^2$  ( $i = 1, 2, \dots$ ), while each element of  $\partial \sigma_t^2 / \partial \gamma$  has a representation of the infinite sum of  $\epsilon_{t-i} \epsilon_{t-i-j}$  ( $j = 1, 2, \dots$ ). Thus, we have  $E[(\partial \sigma_t^2 / \partial \gamma)(\partial \sigma_t^2 / \partial \delta)] = O$  if  $E(\epsilon_{t-i}^3) = 0$ , which is satisfied by the normality assumption of  $\epsilon_t$ . Noting the difference of time of  $\sigma_t^2$  and  $(\partial \sigma_t^2 / \partial \gamma)(\partial \sigma_t^2 / \partial \delta)$ , we can show that  $-n^{-1} \sum_{t=1}^n (\partial^2 l_t / \partial \gamma \partial \delta') \xrightarrow{\text{a.s.}} O$ .

As in Ling and Li [21], we can also show that  $-n^{-1} \sum_{t=1}^n (\partial^2 l_t / \partial \delta \partial u) \xrightarrow{\text{a.s.}} O$ . The block diagonal structure of the information matrix with respect to  $(u, \gamma')'$  and  $\delta$  implies that the distribution of  $(\hat{u}, \hat{\gamma}')'$  is asymptotically independent of that of  $\hat{\delta}$ . Hence, we focus on the first derivatives and the information matrix regarding  $(u, \gamma')'$  for deriving the asymptotic properties of  $\hat{u}$ , in the next subsection.

For deriving the asymptotic properties of the information matrix regarding  $\hat{u}$ , we present the results of Chung [33].

$$\begin{aligned} I_u &= E \left[ \left( \frac{\partial \epsilon_t}{\partial u} \right)^2 \right] = E(\epsilon_t^2) \times 4d^2 \sum_{j=1}^{\infty} a_j^2, \\ I_{ud} &= \left[ \left( \frac{\partial \epsilon_t}{\partial u} \right) \left( \frac{\partial \epsilon_t}{\partial d} \right) \right] = E(\epsilon_t^2) \times 4d \sum_{j=1}^{\infty} \frac{\cos(j\omega_g) a_j}{j}, \end{aligned}$$

and the results indicate that when  $|u| < 1$

$$I_u = E(\epsilon_t^2) \times \frac{4d^2}{\sin^2(\omega_g)} \sum_{j=1}^{\infty} \sin^2(j\omega_g) = \infty, \quad I_{ud} = E(\epsilon_t^2) \times \frac{d(\pi - 2\omega_g)}{\sin(\omega_g)} < \infty,$$

and when  $u = \pm 1$ :

$$I_u = E(\epsilon_t^2) \times 4d^2 \sum_{j=1}^{\infty} j^2 = \infty, \quad I_{ud} = E(\epsilon_t^2) \times \sum_{j=1}^{\infty} (\pm 1)^{j-1} \cos(j\omega_g) < \infty.$$

As discussed in [21], we also show that every element of  $I_{\gamma u} = E[(\partial \epsilon_t / \partial \gamma)(\partial \epsilon_t / \partial u)]$  is finite. Using a similar approach as in Theorem 3.1 of [21], we have  $E[\partial^2 L(\lambda) / \partial \gamma \partial u] = O(1)$ . By the above results for  $I_u$ , we show that  $E[\partial^2 L(\lambda) / \partial u^2]$  does not exist, indicating that the usual asymptotic theory based on  $O_p(n^{-1/2})$  convergence will not work as discussed in [33].

#### Appendix A.2. Proof of Proposition 1

Since the information matrix is block diagonal, we focus on the parts related to  $(u, \gamma')'$  in order to prove the proposition.

We consider the following Taylor series expansion of the first conditions for the maximization of the approximate likelihood function around the true value of  $\lambda_0$

$$\begin{pmatrix} \frac{1}{c_n} \frac{\partial L(\lambda_0)}{\partial u} \\ \frac{1}{\sqrt{n}} \frac{\partial L(\lambda_0)}{\partial \gamma} \end{pmatrix} + \begin{pmatrix} \frac{1}{c_n^2} \frac{\partial^2 L(\lambda_0)}{\partial u^2} & \frac{1}{c_n \sqrt{n}} \frac{\partial^2 L(\lambda_0)}{\partial u \partial \gamma'} \\ \frac{1}{c_n \sqrt{n}} \frac{\partial^2 L(\lambda_0)}{\partial \gamma \partial u} & \frac{1}{n} \frac{\partial^2 L(\lambda_0)}{\partial \gamma \partial \gamma'} \end{pmatrix} \begin{pmatrix} c_n(\hat{u} - u_0) \\ \sqrt{n}(\hat{\gamma} - \gamma_0) \end{pmatrix} = o_p(1),$$

where

$$c_n = \begin{cases} n & \text{when } |u| < 1, \\ n^2 & \text{when } |u| = 1, \end{cases}$$

are the rates suggested by Chung [33].

If  $c_n$  is appropriate, we have  $c_n^{-1} n^{-1/2} \partial^2 L(\lambda_0) / \partial \gamma \partial u = o_p(1)$ , immediately. If we can show

$$\frac{1}{c_n} \sum_{t=1}^n \frac{\partial l_t(\lambda_0)}{\partial u} = O_p(1), \quad \frac{1}{c_n^2} \sum_{t=1}^n \frac{\partial^2 l_t(\lambda_0)}{\partial u^2} = O_p(1), \quad (\text{A6})$$

then we obtain

$$c_n(\hat{u} - u) = - \left[ \frac{1}{c_n^2} \sum_{t=1}^n \frac{\partial^2 l_t(\lambda_0)}{\partial u^2} \right]^{-1} \left[ \frac{1}{c_n} \sum_{t=1}^n \frac{\partial l_t(\lambda_0)}{\partial u} \right] + o_p(1), \quad (\text{A7})$$

since  $\hat{u}$  is asymptotically independent of  $\hat{\gamma}$ . In the remainder of the proof, we derive the limiting distributions of the two series in (A6) and then that of  $c_n(\hat{u} - u)$  in (A7). For this purpose, we use the theory of nonstationary ARMA process with GARCH errors, derived by Ling and Li [45].

By the conditions of Lemma 4, we can write

$$[\theta(z)]^{-1} = \sum_{k=0}^{\infty} \varphi_{\epsilon}(k) z^k, \quad \alpha(z)[\beta(z)]^{-1} = \sum_{k=1}^{\infty} \varphi_{\sigma}(k) z^k.$$

Assuming pre-sample values,  $\epsilon_0 = \epsilon_1 = \epsilon_2 = \dots = 0$  and combining with (A1), we have

$$\frac{\partial \sigma_t^2}{\partial u} = \sum_{k=1}^{\infty} \varphi_{\sigma}(k) \epsilon_{t-k} \frac{\partial \epsilon_{t-k}}{\partial u} = \sum_{k=1}^{t-1} \varphi_{\sigma}(k) \epsilon_{t-k} \frac{\partial \epsilon_{t-k}}{\partial u}.$$

Let

$$\zeta_t = (1 - 2uB + B^2)^{-1} \theta(B) \epsilon_t.$$

Then, from (A2), we have

$$\frac{\partial \epsilon_t}{\partial u} = -2d[\theta(B)]^{-1} \zeta_{t-1} = -2d \sum_{k=0}^{\infty} \varphi_{\epsilon}(k) \zeta_{t-k-1} = -2d \sum_{k=0}^{t-1} \varphi_{\epsilon}(k) \zeta_{t-k-1}.$$

Denote  $\zeta_t = (\zeta_t, \zeta_{t-1})'$  and define

$$\begin{aligned}\zeta_t &= \begin{pmatrix} \zeta_{1t} \\ \zeta_{2t} \end{pmatrix} = \frac{\epsilon_t}{\sigma_t^2} \sum_{k=0}^{t-1} \varphi_\epsilon(k) \zeta_{t-k-1} - \frac{1}{\sigma_t^2} \left( \frac{1}{\sigma_t^2} - 1 \right) \sum_{j=1}^{t-1} \sum_{k=0}^{t-1} \varphi_\sigma(j) \varphi_\epsilon(k) \epsilon_{t-j} \zeta_{t-j-k-1}, \\ \Xi_t &= \begin{pmatrix} \Xi_{11,t} & \Xi_{12,t} \\ \Xi_{21,t} & \Xi_{22,t} \end{pmatrix} \\ &= \frac{1}{\sigma_t^2} \sum_{j=0}^{t-1} \sum_{k=0}^{t-1} \varphi_\epsilon(j) \varphi_\epsilon(k) \zeta_{t-j-1} \zeta'_{t-k-1} \\ &\quad + \frac{2\epsilon_t^2}{\sigma_t^6} \sum_{j_1, j_2=1}^{t-1} \sum_{k_1, k_2=0}^{t-1} \varphi_\epsilon(j_1) \varphi_\epsilon(j_2) \varphi_\epsilon(k_1) \varphi_\epsilon(k_2) \epsilon_{t-j_1} \epsilon_{t-j_2} \zeta_{t-j_1-k_1-1} \zeta'_{t-j_2-k_2-1}.\end{aligned}$$

From (A1)–(A3), it is easy to verify

$$\frac{\partial l_t(\lambda_0)}{\partial u} = -2d\zeta_{1t}, \quad \frac{\partial^2 l_t(\lambda_0)}{\partial u^2} = 4d^2\Xi_{11,t} + o_p(1) \quad (\text{A8})$$

Below, we apply the theorems of [45] to derive their limiting distributions. For this purpose, we separately consider three cases:  $|u| < 1$ ,  $u = 1$  and  $u = -1$ .

**Case 1:**  $|u| < 1$

By Theorem 4.3 of [45], we obtain

$$\begin{aligned}\frac{1}{n} \sum_{t=1}^n \zeta_t &\xrightarrow{\mathcal{L}} \begin{pmatrix} \zeta_1^* \\ \zeta_2^* \end{pmatrix}, \\ \frac{1}{n^2} \sum_{t=1}^n \Xi_t &\xrightarrow{\mathcal{L}} \frac{K}{4\sin^2(\omega_g)} \left( \int_0^1 \tilde{W}_1^2(r) dr + \int_0^1 W_1^2(r) dr \right) \begin{pmatrix} 1 & \cos(\omega_g) \\ \sin(\omega_g) & 1 \end{pmatrix},\end{aligned}$$

where

$$\begin{aligned}\zeta_1^* &= \frac{1}{2\sin(\omega_g)} \left[ \int_0^1 \tilde{W}_1 dW_2 - \int_0^1 W_1 d\tilde{W}_2 \right], \\ \zeta_2^* &= \frac{1}{2\sin(\omega_g)} \left\{ \cos(\omega_g) \left[ \int_0^1 \tilde{W}_1 dW_2 - \int_0^1 W_1 d\tilde{W}_2 \right] \right. \\ &\quad \left. - \sin(\omega_g) \left[ \int_0^1 \tilde{W}_1 d\tilde{W}_2 + \int_0^1 W_1 dW_2 \right] \right\},\end{aligned}$$

and  $K$ ,  $(\tilde{W}_1(t), \tilde{W}_2(t))$  and  $(W_1(t), W_2(t))$  are defined in Proposition 1. Noting (A8), we obtain (15).

**Case 2:**  $u = 1$

By Theorem 4.1 of [45], we obtain

$$\sum_{t=1}^n M_n \zeta_t \xrightarrow{\mathcal{L}} \begin{pmatrix} \zeta_1^* \\ \zeta_2^* \end{pmatrix}, \quad \sum_{t=1}^n M_n \Xi_t M_n' \xrightarrow{\mathcal{L}} K \begin{pmatrix} \Xi_{11}^* & \Xi_{12}^* \\ \Xi_{21}^* & \Xi_{22}^* \end{pmatrix},$$

where

$$\begin{aligned} M_n &= \begin{pmatrix} 1/n^2 & 0 \\ 1/n & -1/n \end{pmatrix}, \\ \xi_1^* &= \int_0^1 \left\{ \int_0^r W_1(s) ds \right\} dW_2(r), \\ \xi_2^* &= \int_0^1 W_1(r) dW_2(r), \\ \Xi_{11}^* &= \int_0^1 \left\{ \int_0^r W_1(s) ds \right\}^2 dr, \\ \Xi_{12}^* &= \Xi_{21}^* = \int_0^1 \left\{ \int_0^r W_1(s) ds \right\} \left\{ \int_0^r \int_0^s W_1(t) dt ds \right\} dr, \\ \Xi_{22}^* &= \int_0^1 \left\{ \int_0^r \int_0^s W_1(t) dt ds \right\}^2 dr. \end{aligned}$$

Noting (A8), we obtain (16) for the distribution of  $n^2(\hat{u} - 1)$ .

**Case 3:**  $u = -1$

By Theorem 4.2 of [45], we obtain

$$\sum_{t=1}^n M_n \xi_t \xrightarrow{\mathcal{L}} \begin{pmatrix} \xi_1^* \\ \xi_2^* \end{pmatrix}, \quad \sum_{t=1}^n M_n \Xi_t M_n' \xrightarrow{\mathcal{L}} K \begin{pmatrix} \Xi_{11}^* & \Xi_{12}^* \\ \Xi_{21}^* & \Xi_{22}^* \end{pmatrix},$$

where

$$\begin{aligned} M_n &= \begin{pmatrix} 1/n^2 & 0 \\ -1/n & 1/n \end{pmatrix}, \\ \xi_1^* &= - \int_0^1 \left\{ \int_0^r W_1(s) ds \right\} dW_2(r), \\ \xi_2^* &= - \int_0^1 W_1(r) dW_2(r), \end{aligned}$$

and  $\Xi_{11}^*$ ,  $\Xi_{12}^*$ ,  $\Xi_{21}^*$  and  $\Xi_{22}^*$  are the same as in Case 2.

Noting (A8), we obtain (16) for the distribution of  $n^2(\hat{u} + 1)$ .

### Appendix A.3. Proof of Proposition 2

Applying (A5) and the proof of Theorem 3.2 of [21], we can show that the conditions provided by Basawa et al. [46] are satisfied. The only difference is on the third derivative of  $\epsilon_t$  with respect to  $d$ :

$$\frac{\partial^3 \epsilon_t}{\partial d^3} = [\log(1 - 2uB + B^2)]^3 \epsilon_t,$$

which has the finite second moment as:

$$\begin{aligned} E \left[ \left( \frac{\partial^3 \epsilon_t}{\partial d^3} \right)^2 \right] &= 64 E \left[ \left( \sum_{k_1, k_2, k_3=1}^{\infty} \frac{\cos(k_1 \omega_g) \cos(k_2 \omega_g) \cos(k_3 \omega_g)}{k_1 k_2 k_3} \epsilon_{t-k_1-k_2-k_3} \right)^2 \right] \\ &= 64 \sum_{k_1, k_2, k_3=1}^{\infty} \frac{\cos^2(k_1 \omega_g) \cos^2(k_2 \omega_g) \cos^2(k_3 \omega_g)}{k_1^2 k_2^2 k_3^2} E(\epsilon_{t-k_1-k_2-k_3}^2) \\ &\leq 64 \sum_{k_1, k_2, k_3=1}^{\infty} \frac{1}{k_1^2 k_2^2 k_3^2} E(\epsilon_{t-k_1-k_2-k_3}^2) = \left( \frac{2\pi}{3} \right)^3 E(\epsilon_t^2) < \infty. \end{aligned}$$

## References

- Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **1982**, *50*, 987–1008.
- Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *J. Econom.* **1986**, *31*, 307–327.
- Niemi, H. On the effect of a nonstationary noise on ARMA models. *Scand. J. Stat.* **1983**, *10*, 11–17.
- Peiris, M.S. Analysis of multivariate ARMA processes with nonstationary innovations. *Commun. Stat. Theory Methods* **1990**, *19*, 2847–2852.
- Nelson, D.B. Stationarity and persistence in the GARCH (1,1) model. *Econom. Theory* **1990**, *6*, 318–334.
- Bougerol, P.; Picard, N. Stationarity of GARCH processes and of some nonnegative time series. *J. Econom.* **1992**, *52*, 115–127.
- McAleer, M. Automated inference and learning in modeling financial volatility. *Econom. Theory* **2005**, *21*, 232–261.
- Shephard, N. General introduction. In *Stochastic Volatility*; Shephard, N., Ed.; Oxford University Press: Oxford, UK, 2005; pp. 1–33.
- Abdrabbo, N.A.; Priestley, M.B. On the prediction of nonstationary processes. *J. R. Statist. Soc. Ser. B* **1967**, *29*, 570–585.
- Hallin, M. Mixed autoregressive-moving average multivariate processes with time-dependent coefficients. *J. Multivar. Anal.* **1978**, *8*, 567–572.
- Priestley, M.B. *Spectral Analysis and Time Series*; Academic Press: New York, NY, USA, 1981.
- Hallin, M.; Ingenbleek, J.-F. Nonstationary Yule-Walker equations. *Stat. Probab. Lett.* **1983**, *1*, 189–195.
- Peiris, S.; Singh, N. A note on the properties of some nonstationary ARMA processes. *Stoch. Process. Appl.* **1987**, *24*, 151–155.
- Backus, D.; Zin, S. Long memory inflation uncertainty: Evidence from the term structure of interest rates. *J. Money Credit Bank.* **1993**, *25*, 681–700.
- Hassler, U.; Wolters, J. Long memory in inflation rates: International evidence. *J. Bus. Econ. Stat.* **1995**, *13*, 37–45.
- Baillie, R.T.; Chung, C.F.; Tieslau, M.A. Analyzing inflation by the fractionally integrated ARFIMA-GARCH Model. *J. Appl. Econom.* **1996**, *11*, 23–40.
- Caporale, G.M.; Gil-Alana, L.A. Multi-factor Gegenbauer processes and European inflation rates. *J. Econ. Integr.* **2011**, *26*, 386–409.
- Delgado, M.; Robinson, R.M. New methods for the analysis of long memory time series: Application to Spanish inflation. *J. Forecast.* **1994**, *13*, 97–107.
- Baillie R.T.; Bollerslev, T.; Mikkelsen, H.O. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1996**, *74*, 3–30.
- Bollerslev, T.; Mikkelsen, H.O. Modeling and pricing long-memory in stock market volatility. *J. Econom.* **1996**, *73*, 151–184.
- Ling, S.; Li, W.K. On fractionally integrated autoregressive moving average time series with conditional heteroscedasticity. *J. Am. Stat. Assoc.* **1997**, *92*, 1184–1194.
- Breidt, F.J.; Crato, N.; de Lima, P. The detection and estimation of long memory in stochastic volatility. *J. Econom.* **1998**, *83*, 325–348.
- Deo, R.S.; Hurvich, C.M. On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econom. Theory* **2001**, *17*, 686–710.
- Bos, C.; Koopman, S.J.; Ooms, M. Long memory with stochastic variance model: A recursive analysis for US inflation. *Comput. Stat. Data Anal.* **2014**, *76*, 144–157.
- Woodward, W.A.; Cheng, Q.C.; Gray, H.L. A  $k$ -factor GARMA long memory model. *J. Time Ser. Anal.* **1998**, *19*, 485–504.
- Ferrara, L.; Guegan, D. Forecasting with  $k$ -factor Gegenbauer processes: theory and applications. *J. Forecast.* **2001**, *20*, 581–601.
- Gray, H.L.; Zhang, N.; Woodward, W.A. On Generalized fractional processes. *J. Time Ser. Anal.* **1989**, *10*, 233–257.
- Chung, C.F. A generalized fractionally integrated autoregressive moving-average process. *J. Time Ser. Anal.* **1996**, *17*, 111–140.

29. Shitan, M.; Peiris, M. Generalized autoregressive (GAR) model: A comparison of maximum likelihood and whittle estimation procedures using a simulation study. *Commun. Stat. Theory Methods* **2008**, *37*, 560–570.
30. Shitan, M.; Peiris, M. Approximate asymptotic variance-covariance matrix for the Whittle estimators of GAR(1) parameters. *Commun. Stat. Theory Methods* **2013**, *42*, 756–770.
31. Erdelyi, A.; Magnus, W.; Oberhettinger, F.; Tricomi, F.G. *Higher Transcendental Functions*; Bateman Manuscript Project; McGraw and Hill: New York, NY, USA, 1953; Volume 2.
32. Chesney, M.; Scott, L.O. Pricing European currency options: A comparison of the modified black-scholes model and a random variance model. *J. Financ. Quant. Anal.* **1989**, *24*, 267–284.
33. Chung, C.F. Estimating a generalized long memory process. *J. Econom.* **1996**, *73*, 237–259.
34. Danielsson, J.; Richard, J.-F. Quadratic acceleration for simulated maximum likelihood evaluation. *J. Appl. Econom.* **1993**, *8*, 153–173.
35. Shephard, N.; Pitt, M.K. Likelihood analysis of non-gaussian measurement time series. *Biometrika* **1997**, *84*, 653–667.
36. Durbin, J.; Koopman, S.J. Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika* **1997**, *84*, 669–684.
37. Durbin, J.; Koopman, S.J. Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives. *J. R. Stat. Soc. Ser. B* **2000**, *62*, 3–56.
38. Liesenfeld, R.; Richard, J.-F. Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *J. Empir. Finance* **2003**, *10*, 505–531.
39. McElroy, T.S.; Holan, S.H. On the computation of autocovariances for generalized Gegenbauer processes. *Stat. Sinica* **2012**, *22*, 1661–1687.
40. Robinson, P.M.; Henry, M. Long and short memory conditional heteroscedasticity. *J. Time Ser. Anal.* **1999**, *15*, 299–336.
41. Hidalgo, J.; Soulier, P. Estimation of the location and exponent of the spectral singularity of a long memory process. *J. Time Ser. Anal.* **2004**, *25*, 55–81.
42. Jacquier, E.; Polson, N.G.; Rossi, P.E. Bayesian analysis of stochastic volatility models (with discussion). *J. Bus. Econ. Stat.* **1994**, *12*, 371–389.
43. De Jong, P.; Shephard, N. The simulation smoother for time series models. *Biometrika* **1995**, *82*, 339–350.
44. Weiss, A.A. Asymptotic theory for ARCH models: Estimation and testing. *Econom. Theory* **1986**, *2*, 107–131.
45. Ling, S.; Li, W.K. Limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with general conditional heteroscedastic errors. *Ann. Stat.* **1998**, *26*, 84–125.
46. Basawa, I.V.; Feign, P.D.; Heyde, C.C. Asymptotic properties of maximum likelihood estimators for stochastic processes. *Sankhya Ser. A* **1976**, *38*, 259–270.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).