

Estimating and Forecasting Generalized Fractional Long Memory Stochastic Volatility Models *

EI2016-27

Shelton Peiris

School of Mathematics and Statistics
University of Sydney, Australia

Manabu Asai

Faculty of Economics
Soka University, Japan

Michael McAleer

Department of Quantitative Finance
National Tsing Hua University, Taiwan
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Department of Quantitative Economics
Complutense University of Madrid, Spain

June 2016

*The authors are most grateful to Yoshi Baba for very helpful comments and suggestions. The first author acknowledges the support from the Faculty of Economics at Soka University during his visit. The second author acknowledges the financial support of Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science, Zengin Foundation for Studies on Economics and Finance, and Australian Academy of Science. The third author is most grateful for the financial support of the Australian Research Council, National Science Council, Ministry of Science and Technology, Taiwan, and the Japan Society for the Promotion of Science.

Abstract

In recent years fractionally differenced processes have received a great deal of attention due to its flexibility in financial applications with long memory. This paper considers a class of models generated by Gegenbauer polynomials, incorporating the long memory in stochastic volatility (SV) components in order to develop the General Long Memory SV (GLMSV) model. We examine the statistical properties of the new model, suggest using the spectral likelihood estimation for long memory processes, and investigate the finite sample properties via Monte Carlo experiments. We apply the model to three exchange rate return series. Overall, the results of the out-of-sample forecasts show the adequacy of the new GLMSV model.

Keywords: Stochastic volatility, GARCH models, Gegenbauer Polynomial, Long Memory, Spectral Likelihood, Estimation, Forecasting.

JEL Classification: C18, C21, C58.

1 Introduction

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

$$\phi(B)Y_t = \theta(B)\epsilon_t, \tag{1.1}$$

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

This standard case of constant variance innovations has been considered in many traditional time series analyses and applications. However, in recent years, there has been a great deal of development with time dependent instantaneous innovation variances (or volatility). Two popular classes have been developed in modeling financial volatility. One is the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) family, pioneered by Engle (1982), while the other emphasizes Stochastic Volatility (SV) models, using the ideas of Clark (1973) and Taylor (1982, 1986) (see the survey papers of McAleer (2005) and Shephard (2005) for further details). Note that the so-called ‘realized volatility’ can be considered as noise plus the realized value of the latent volatility in SV models (see Barndorff-Nielsen and Shephard (2002), Bollerslev and Zhou (2002), and Asai et al. (2012) for further details).

As the conditional volatility displays long memory or long range dependencies in many financial applications, Baillie et al. (1996) and Bollerslev and Mikkelsen (1996) developed the Fractionally Integrated GARCH (FIGARCH) and Fractionally Integrated Exponential GARCH (FIEGARCH) models, respectively. In the light of this evidence, Breidt et al. (1998) developed the long memory SV (LMSV) model, in which log-volatility follows the ARFIMA(p, d, q) (or FARIMA(p, d, q)) process. Empirical evidence from Breidt et al (1998), Andersen et al. (2001, 2003), Pong et al. (2004), Koopman, Jungbacker, and Hol (2005), and Asai et al. (2012) indicate that estimates of d lie between zero and one.

Motivated by these extensions and applications, Arteche (2004) developed the generalized LMSV model, using the Gegenbauer process. The Gegenbauer process is a type of long memory process, developed by Gray et al. (1989). Incorporating the Gegenbauer process in volatility modeling enables a more flexible class of process for the conditional/stochastic

variance that is capable of explaining and representing the observed temporal dependencies in financial market volatility. Arteche (2004) suggested the semi-parametric estimation technique for its long memory parameter.

The main purpose of this paper is to extend the work of Arteche (2004) by considering short memory components and spectral likelihood estimation for general long memory stochastic volatility models.

The organization of the paper is as follows. Section 2 briefly reviews stochastic volatility models, while Section 3 introduces the Gegenbauer ARMA process. Section 4 develops the new generalized LMSV model, and develops its statistical properties. Section 5 suggests estimation via spectral likelihood (SL), which is equivalent to the quasi-maximum likelihood (QML) estimator, and examines the finite sample properties of the SL estimator. Section 5 also explains the method for estimating and forecasting volatility. Section 6 presents empirical results using the exchange rate returns of Japanese Yen (YEN), Euro (EUR), and British Pound (GBP) relative to the US dollar (USD). Section 7 provides concluding remarks.

2 Review of Stochastic Volatility (SV) Models

An alternative to the modeling of the popular GARCH and related conditional volatility models is a class of models such that the variance follows a certain latent stochastic process. Suppose that a discrete time series $\{Y_t\}$ is given by $Y_t = \sigma_t \xi_t$, where $\xi_t \sim IID(0, 1)$ and the volatility process satisfies:

$$\sigma_t = \exp(X_t/2). \tag{2.1}$$

Two popular cases related to (2.1) have been analysed in the literature:

- $\{X_t\}$ follows a stationary and invertible ARMA(p, q) process given by:

$$\phi(L)X_t = C + \theta(L)v_t, \tag{2.2}$$

where v_t is white noise with zero mean and variance σ_v^2 , C is a constant, L is the lag operator, and the roots of $\phi(L)$ (AR(p) polynomial) and $\theta(L)$ (MA(q) polynomial) lie outside the unit circle to ensure stationarity and invertibility of $\{X_t\}$.

- $\{X_t\}$ follows a stationary and invertible ARFIMA(p,d,q) process given by:

$$\phi(L)(1-L)^d X_t = C + \theta(L)v_t, \quad (2.3)$$

where, in addition to the conditions in (2.2), the parameter $d \in (-0.5, 0.5)$ to ensure stationarity and invertibility of $\{X_t\}$.

Particular attention has been paid to the class in (2.3) when $0 < d < 0.5$ to model long memory in SV. In this case, (2.1) and (2.3) describe a family of LMSV. This paper introduces a general family of long memory models with SV. In order to develop the theory, we first consider Gegenbauer polynomials and Gegenbauer ARMA (GARMA).

3 Gegenbauer ARMA (GARMA) Model

Suppose that a time series $\{X_t\}$ is generated by:

$$\phi(L)(1 - 2\eta L + L^2)^d X_t = \theta(L)v_t, \quad (3.1)$$

where the polynomials $\phi(L)$, $\theta(L)$ and noise $\{v_t\}$ are as defined in (2.2), and $|\eta| \leq 1$ and $|d| \leq 1$ are real parameters.

This family in (3.1) is known as the Gegenbauer ARMA of order($p, d, q; \eta$) or GARMA($p, d, q; \eta$) and enjoys the following properties:

- The power spectrum:

$$f_X(\omega) = [4(\cos \omega - \eta)^2]^{-d} g(\omega), \quad -\pi < \omega < \pi, \quad (3.2)$$

where $g(\omega) = \frac{|\theta(e^{-i\omega})|^2 \sigma_v^2}{|\phi(e^{-i\omega})|^2 2\pi}$ corresponds to the ARMA part.

- The process in (3.1) is stationary and explains long memory when $|\eta| < 1$ and $0 < d < 1/2$, or $|\eta| = 1$ and $0 < d < 1/4$, with the stationary condition on $\phi(L)$.

From (3.2), it is clear that the long memory features are characterized by an unbounded spectrum at the Gegenbauer frequency $\omega = \omega_g = \cos^{-1}(\eta)$ when $|\eta| < 1$, and at $\omega = 0$ when $\eta = 1$, in addition to the hyperbolic decay of the autocorrelation function (acf).

For later reference, we consider a special case, namely, the class of GARMA(0, d , 0; η) given by:

$$(1 - 2\eta L + L^2)^d X_t = v_t. \quad (3.3)$$

Under the AR regularity conditions:

(a1) $|\eta| < 1$ and $d < 1/2$; or

(a2) $|\eta| = 1$ and $d < 1/4$,

the Wold representation of (3.3) is given as:

$$X_t = \psi(L)v_t = \sum_{j=0}^{\infty} \psi_j v_{t-j}, \quad (3.4)$$

where $\psi(L) = (1 - 2\eta L + L^2)^{-d} = \sum_{j=0}^{\infty} \psi_j L^j$, with $\psi_0 = 1$, and the Gegenbauer coefficients ψ_j in terms of the Gamma functions, $\Gamma(\cdot)$, have the explicit representation:

$$\psi_j = \sum_{k=0}^{[j/2]} \frac{(-1)^k (2\eta)^{j-2k} \Gamma(d-k+j)}{k!(j-2k)!\Gamma(d)}, \quad j \geq 0 \quad (3.5)$$

such that $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ (see Erdélyi et al., 1953, 10.9 for details). The coefficients ψ_j , $j \geq 2$, are recursively related by:

$$\psi_j = 2\eta \left(\frac{d-1+j}{j} \right) \psi_{j-1} - \left(\frac{2d-2+j}{j} \right) \psi_{j-2},$$

with initial values $\psi_0 = 1$ and $\psi_1 = 2d\eta$. These coefficients, ψ_j , reduce to the corresponding standard long memory (or binomial) coefficients when $\eta = 1$, such that $\psi_j = \frac{\Gamma(2d+j)}{\Gamma(j+1)\Gamma(2d)}$.

Under the MA regularity conditions:

(b1) $|\eta| < 1$ and $d > -1/2$; or

(b2) $|\eta| = 1$ and $d > -1/4$,

(3.3) admits an invertible solution, such that:

$$v_t = (1 - 2\eta L + L^2)^d X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad (3.6)$$

where the coefficients, π_j , are obtained from (3.5), by replacing d with $-d$.

In the general case (3.1), the corresponding stationary and invertible solutions can be obtained from:

$$X_t = \psi(L) * \psi'(L)v_t,$$

and

$$v_t = [\psi'(L)]^{-1} * (1 - 2\eta L + L^2)^d X_t,$$

respectively, where $\psi'(L) = [\phi(L)]^{-1}\theta(L)$ (see Dissanayake et al. (2016) for further details). In recent papers, Shitan and Peiris (2008, 2013) have considered an alternative family of generalized fractional processes given by:

$$\phi(L)(1 - \alpha L)^d X_t = \theta(L)v_t.$$

As an extension, Section 4 develops a new family of generalized long memory volatility models using Gegenbauer polynomials.

4 Generalized Long Memory SV (GLMSV) Models

This section considers the generalized long memory SV (GLMSV) model, defined by:

$$Y_t = \sigma_t \xi_t, \quad \xi_t \sim IID(0, 1), \quad \sigma_t = \exp(X_t/2), \quad (4.1)$$

$$\phi(L)(1 - 2\eta L + L^2)^d (X_t - \mu) = \theta(L)v_t, \quad (4.2)$$

where $\{\xi_t\}$ is independent of $\{X_t\}$ for all t . In the model, log-volatility follows the GARMA($p, d, q; \eta$) process. From the spectrum of (4.2), it is clear that the log volatility process, $\{X_t\}$, has generalized long memory when $|\eta| < 1$ and $0 < d < 0.5$, with a spectral peak at Gegenbauer frequency $\omega_g = \cos^{-1}(\eta)$. As distinct from Arteche (2004), we incorporated the short memory components, $\phi(L)$ and $\theta(L)$, and excluded seasonal long memory to avoid over-parameterization in long range dependencies.

4.1 Properties of GLMSV

Suppose that $\{v_t\}$ in (3.1) is Gaussian and let $\gamma(k)$ be the autocovariance function (ACVF) of $\{X_t\}$ given by $\gamma(k) = Cov(X_t, X_{t+k})$. It follows from the properties of the lognormal distribution that:

- $E(Y_t) = 0$ and $Var(Y_t) = \exp[\gamma(0)/2]$,
- $\gamma_Y(k) = Cov(Y_t, Y_{t+k}) = 0$ for all $k \neq 0$,
- $\{Y_t\}$ is a martingale difference.

Let $U_t = \log(Y_t^2)$. Then the observation equation satisfies the linear state space model, $U_t = \log(\sigma_t^2) + \log(\xi_t^2)$, and reduces to:

$$U_t = c + X_t + \epsilon_t, \quad (4.3)$$

where $c = \mu + E[\log(\xi_t^2)]$ and $\epsilon_t = \log(\xi_t^2) - E[\log(\xi_t^2)]$ is an *iid* process independent of $\{X_t\}$. Note that, if ξ_t is standard normal, then $\xi_t^2 \sim \chi_1^2$, which gives $E[\log(\xi_t^2)] = -1.2704$ and $Var[\log(\xi_t^2)] = \frac{\pi^2}{2} \approx 4.93$.

It follows from (4.3) that the corresponding spectra are related by:

$$f_U(\omega) = f_X(\omega) + \frac{\sigma_\epsilon^2}{2\pi}, \quad -\pi < \omega < \pi, \quad (4.4)$$

where $f_X(\omega) = g(\omega)[4(\cos \omega - \eta)^2]^{-d}$, $-\pi < \omega < \pi$, and $\sigma_\epsilon^2 = Var(\epsilon_t)$.

From the results in Granger and Morris (1976) for the sum of an MA process and noise, we can write:

$$U_t = c + \sum_{j=0}^{\infty} \tilde{\psi}_j v_{t-j} + \epsilon_t = c + \sum_{j=0}^{\infty} \kappa_j e_{t-j}, \quad (4.5)$$

where $\{e_t\}$ is a white noise process, and $\tilde{\psi}_j$ is the j th coefficient of the polynomial $\tilde{\psi}(z) = (1 - 2\eta z + z^2)^{-d} \phi(z)^{-1} \theta(z)$, with $\tilde{\psi}_0 = 1$. Hence, we obtain the MA(∞) representation of U_t . The distribution of e_t can be obtained by the convolution of the distributions of X_t and ϵ_t , where $\{e_t\}$ is serially uncorrelated, but is not an independent process.

Clearly, (4.4) implies that the log squared returns of $\{Y_t\}$ have long memory, with the same memory parameter d as in the volatility process $\{X_t\}$. In particular, when $\eta = 1$ and $0 < d < 1/4$, GLMSV reduces to the standard LMSV. These spectral properties can be used to identify the GLMSV and LMSV processes in practice.

4.2 Identification of GLMSV and LMSV

The following lemma on spectral densities can be used to identify LMSV and/or GLMSV.

Lemma: $f_U(\omega) \sim f_X(\omega)$ as $\omega \rightarrow \omega_g = \arccos(\eta)$.

Proof: Let $f^*(\omega) = [f_X(\omega)]^{-1} \frac{\sigma_v^2}{2\pi}$. Then from (4.4) we have:

$$f_U(\omega) = f_X(\omega)[1 + f^*(\omega)]. \quad (4.6)$$

Clearly, $f^*(\omega)$ is bounded from above and bounded away from zero when $0 < d < 0.5$, and $f^*(\omega) \rightarrow 0$ as $\omega \rightarrow \omega_g = \arccos(\eta)$. Hence, the lemma holds. ■

The lemma shows that the spectrum of $\{U_t\}$ behaves like that of $\{X_t\}$ near the Gegenbauer frequency, ω_g . We illustrate this for three important cases by taking $\phi(L) = \theta(L) = 1$ for simplicity.

Illustrations

- Standard LMSV when $\eta = 1$:

The sdf of $\{U_t\}$ is given by:

$$f_U(\omega) \sim [2(\sin(\omega/2))]^{-4d} \frac{\sigma_v^2}{2\pi}, \quad -\pi < \omega < \pi, \quad (4.7)$$

and is unbounded as $\omega \rightarrow 0$ when $0 < d < 1/2$.

The following diagram illustrates $f = f_U(\omega)$, $d = 0.4$, $\sigma_v^2 = 2$:

- GLMSV when $|\eta| < 1$:

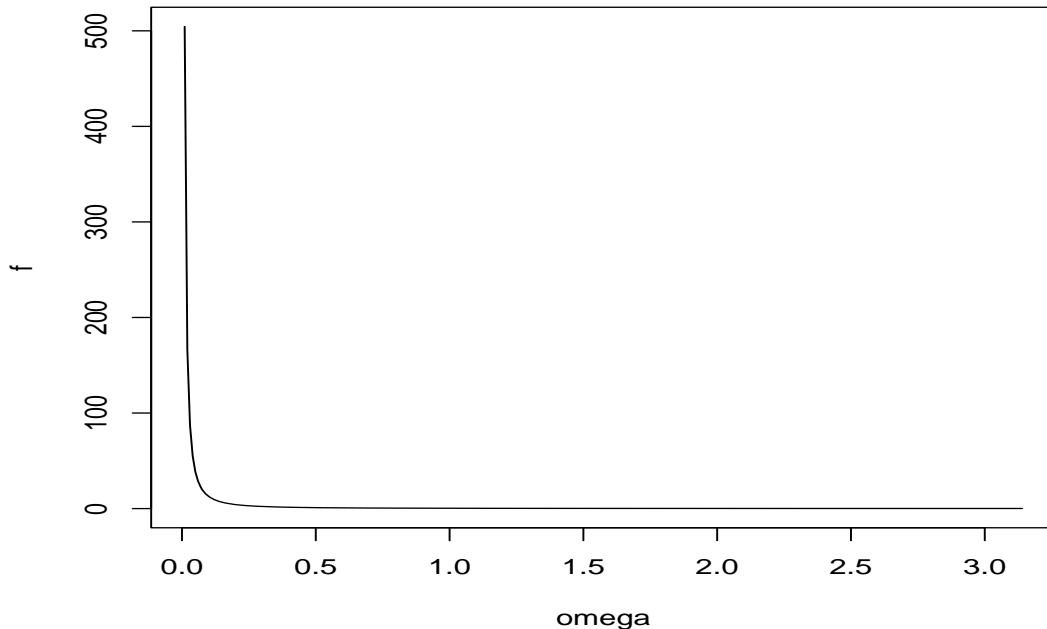
The sdf of $\{U_t\}$ is given by:

$$f_U(\omega) \sim [4(\cos \omega - \eta)^2]^{-d} \frac{\sigma_v^2}{2\pi}, \quad -\pi < \omega < \pi, \quad (4.8)$$

and is unbounded as $\omega \rightarrow \cos^{-1}(\eta)$ (the Gegenbauer frequency, which is away from the origin) for $|\eta| < 1$ and $0 < d < 1/2$.

The second diagram illustrates $f2 = f_U(\omega)$, $d = 0.4$, $\eta = 0.8$, $\sigma_v^2 = 2$:

SDF of LMSV



5 Estimation and Forecasting

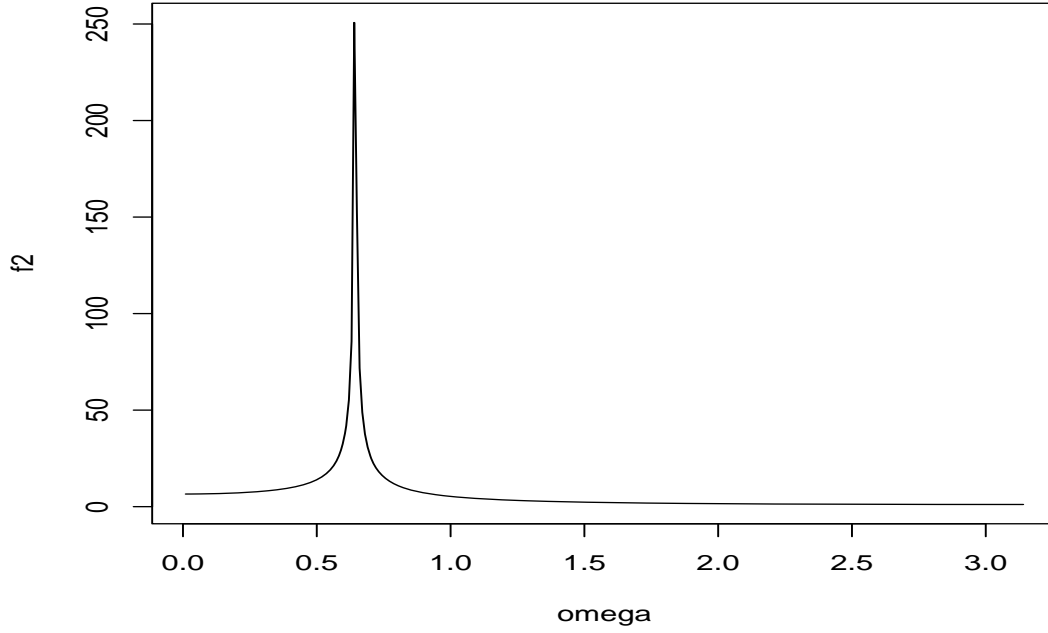
5.1 Spectral-Likelihood Estimator

Though the process $\{v_t\}$ is non-Gaussian, a reasonable estimation procedure is to maximize the quasi-likelihood, or the likelihood computed as if $\{v_t\}$ was Gaussian. For the LMSV models, the approaches of So (1999, 2002) and Doornik and Ooms (2003) enable us to compute the quasi-likelihood exactly, using the autocovariance functions up to order n . For the GLMSV model, it is not easy to calculate the exact autocovariances, but it is possible to obtain their approximate values with the use of the algorithm of McElroy and Holan (2012). Hence, the effectiveness of the QML estimation of this type depends on the accuracy of the approximation of the autocovariance functions. Rather than the approximate approach, we suggest a spectral domain estimator, which was used in estimating the LMSV model by Breidt et al. (1998).

The spectral-likelihood (SL) estimator is obtained by minimizing:

$$\mathcal{L}(\lambda) = \frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \left[\log(f_U(\omega_j)) + \frac{I_n(\omega_j)}{f_U(\omega_j)} \right], \quad (5.1)$$

SDF of GLMSV



where $\lambda = (d, \eta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_v, \sigma_\epsilon)'$ is the vector of unknown parameters, $[\cdot]$ denotes the integer part, $\omega_j = 2\pi j/n$ is the j th Fourier frequency, and

$$I_n(\omega_j) = \frac{1}{2\pi n} \left| \sum_{t=1}^n x_t \exp(-i\omega_j t) \right|^2, \quad j = 1, \dots, [n/2].$$

If we know the value of η *a priori*, we should omit the observation which corresponds to $\omega = \arccos(\eta)$. In a general framework, Hosoya (1997) showed that the SL estimator, $\hat{\lambda}$, is consistent, and:

$$\sqrt{T}(\hat{\lambda} - \lambda_0) \xrightarrow{d} N(0, W^{-1}U(W^*)^{-1}),$$

where λ_0 is the true value,

$$W = \frac{\partial R(\lambda)}{\partial \lambda'},$$

$$R(\lambda) = \frac{\partial}{\partial \lambda} \int_{-\pi}^{\pi} \log f_U(\omega; \lambda) d\omega - \int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \lambda} \log f_U(\omega; \lambda) \right] d\omega,$$

$$\begin{aligned} U &= 4\pi \int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \lambda} \log f_U(\omega; \lambda) \right] \left[\frac{\partial}{\partial \lambda'} \log f_U(\omega; \lambda) \right] d\omega \\ &\quad + (2\pi)^3 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \lambda} \log f_U(\omega_1; \lambda) \right] \left[\frac{\partial}{\partial \lambda'} \log f_U(\omega_2; \lambda) \right] \\ &\quad \quad \quad \times Q^e(\omega_1, \omega_2, -\omega_2) d\omega_1 d\omega_2, \end{aligned}$$

and $Q^e(\omega_1, \omega_2, \omega_3)$ is the fourth-order cumulant spectral density of e_t , defined by (4.5). Furthermore, the SL estimator has the same limiting distribution as the QML estimator in the time domain. In practice, the second term of U can be estimated by the approach of Taniguchi (1982) (see Taniguchi and Kakizawa (2000, Chapter 5) and Zaffaroni (2009) for the general justification of the SL estimator). Zaffaroni (2009) shows the consistency and asymptotic normality of the SL estimator for conditional and stochastic volatility models with both short and long range dependencies.

Following Gray et al. (1989) and Chung (1996a,b), we use the grid search procedure for different values of η over the range $[-1, 1]$ for minimizing (5.1).

5.2 Finite Sample Properties

We conducted Monte Carlo experiments for investigating the finite sample properties of the SL estimator. The parameter values for X_t are specified as:

$$(\mu, \sigma, \phi, d, \eta) = \begin{cases} (0, 0.199, 0.98, 0, 1) & \text{for AR(1)} \\ (0, 0.572, 0.30, 0.2, 1) & \text{for ARFIMA(1,2d,0)} \\ (0, 0.520, 0.30, 0.4, 0.7) & \text{for GARMA(1,d,0), Case 1} \\ (0, 0.675, 0.70, 0.3, 0.3) & \text{for GARMA(1,d,0), Case 2.} \end{cases}$$

In the parameter settings, all the variances of X_t are equal to one. Note that the standard deviation of ϵ_t is $\sigma_\epsilon = \pi/\sqrt{2} = 2.221$, which is greater than twice the standard deviation of X_t . We consider sample sizes $n = \{1024, 2048\}$, with $R = 2000$ replications. For the AR

Table 1: Finite Sample Performance of the SL Estimator of GLMSV

DGP	Parameters					
	μ	σ_ϵ	σ	ϕ	d	η
AR(1)						
True	0	2.221	0.199	0.98	0	1
$n = 1024$	0.0091 (0.2967) [0.2967]	2.1201 (0.1901) [0.2152]	0.4863 (0.4455) [0.5297]	0.9693 (0.0305) [0.0323]	-0.1086 (0.3847) [0.3993]	0.8972 (0.0830) [0.1320]
$n = 2048$	0.0011 (0.2170) [0.2168]	2.1666 (0.1137) [0.1261]	0.4047 (0.3361) [0.3938]	0.9753 (0.0119) [0.0128]	-0.0869 (0.3590) [0.3691]	0.8986 (0.0797) [0.1289]
ARFIMA(1,2d,0)						
True	0	2.221	0.572	0.30	0.2	1
$n = 1024$	0.0034 (0.4011) [0.4007]	2.0796 (0.3942) [0.4185]	0.5907 (0.6551) [0.6547]	0.8551 (0.2127) [0.5944]	0.1004 (0.3359) [0.3500]	0.8901 (0.0824) [0.1373]
$n = 2048$	0.0063 (0.3394) [0.3391]	2.2014 (0.1618) [0.1629]	0.4303 (0.4617) [0.4826]	0.8939 (0.1983) [0.6261]	0.1568 (0.2928) [0.2957]	0.9003 (0.0870) [0.1322]
GARMA(1,d,0), Case 1						
True	0	2.221	0.520	0.30	0.4	0.7
$n = 1024$	0.0027 (0.0684) [0.0684]	1.9444 (0.5856) [0.6473]	0.9212 (0.5432) [0.6749]	0.0988 (0.3501) [0.4035]	0.3301 (0.1072) [0.1280]	0.6984 (0.0307) [0.0307]
$n = 2048$	-0.0017 (0.0564) [0.0564]	2.0965 (0.3353) [0.3575]	0.7608 (0.4068) [0.4724]	0.1693 (0.3143) [0.3401]	0.3572 (0.0797) [0.0904]	0.7005 (0.0052) [0.0053]
GARMA(1,d,0), Case 2						
True	0	2.221	0.675	0.70	0.3	0.3
$n = 1024$	0.0037 (0.0871) [0.0871]	2.0348 (0.5725) [0.6016]	0.8180 (0.5207) [0.5395]	0.6441 (0.2018) [0.2092]	0.2668 (0.1481) [0.1516]	0.3016 (0.0937) [0.0936]
$n = 2048$	-0.0014 (0.0696) [0.0696]	2.1928 (0.2076) [0.2094]	0.7022 (0.2269) [0.2283]	0.6847 (0.1099) [0.1109]	0.2905 (0.0747) [0.0752]	0.3006 (0.0459) [0.0458]

Note: Entries show the means of the SL estimates. Standard errors are in parentheses, and root mean squared errors are in brackets.

and ARFIMA models, the structure $(1 - 2\eta L + L^2)^d$ implies that the estimate of η can take any value when the estimate of d is close to zero.

Table 1 shows the finite sample performances of the SL estimator for the GARMA model. The bias for the estimator of d is negligible for both $n = 1024$ and $n = 2048$. The bias for η is negligible when $d > 0$, and it is meaningless if $d = 0$. As noted before, when the true value of d is zero, the estimates of η can take any values. The results for $d = 0$ show that the estimates of η are close to 0.7, and the RMSE has no major change with respect to the sample size. The bias for the estimates of μ and ϕ are negligible. The estimator of σ_ϵ has a downward bias, while that of σ is biased upward. The result may come from the difference in the sizes of the parameters. The biases for σ_ϵ and σ become small as the sample size increases. For all the parameters, except for the meaningless case of η , the bias, standard deviation, and RMSE decrease as the sample size increases. Next we support the above findings using real data.

5.3 Estimating and Forecasting Volatility

We introduce an algorithm of Harvey (1998) regarding signal extraction and forecasting of long memory plus noise processes. Define $\mathbf{U} = (U_1, \dots, U_n)'$, $\mathbf{X}^* = (X_1 - \mu, \dots, X_n - \mu)'$, and $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)'$, in order to obtain:

$$\mathbf{U} = (c + \mu)\mathbf{1}_n + \mathbf{X}^* + \boldsymbol{\epsilon},$$

where $\mathbf{1}_n$ is an $n \times 1$ vector of ones. Then, the minimum mean square linear estimator of \mathbf{X}^* is given by:

$$\tilde{\mathbf{X}}^* = (I_n - \sigma^2 \mathbf{V}^{-1}) \{\mathbf{U} - (c + \mu)\mathbf{1}_n\},$$

where $\mathbf{V} = \mathbf{V}_X + \sigma^2 I_n$, and \mathbf{V}_X denotes the covariance matrix of \mathbf{X}^* . As noted in Subsection 5.1, \mathbf{V}_X can be approximated by the algorithm of McElroy and Holan (2012) (see the Appendix for details). Harvey (1998) recommends using the volatility estimate:

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_Y^2 \exp\left(\tilde{X}_t^*\right),$$

where $\tilde{\sigma}_Y^2 = n^{-1} \sum_{t=1}^n \tilde{Y}_t^2$, and $\tilde{Y}_t = Y_t \exp(-0.5\tilde{X}_t^*)$ are the heteroskedasticity-corrected observations.

Table 2: Descriptive Statistics for Exchange Rate Returns

Data	Mean	Std. Dev.	Skewness	Kurtosis
YEN/USD	0.0028	0.6617	-0.3225	8.1747
EUR/USD	-0.0045	0.6383	0.1717	5.9683
GBP/USD	-0.0060	0.6163	-0.3377	9.6188

For predicting the observations on U_t for $t = n + 1, \dots, n + h$, denote \mathbf{U}_h as the $h \times 1$ vector of predicted values. Then the corresponding MMSLEs are given by:

$$\tilde{\mathbf{U}}_h = (\mu + c)\mathbf{1}_h + \mathbf{R}\mathbf{V}^{-1}\{\mathbf{U} - (\mu + c)\mathbf{1}_n\}.$$

Using $\tilde{\mathbf{X}}_h = \tilde{\mathbf{U}}_h - (\mu + c)\mathbf{1}_h$, the predictions of σ_{n+j}^2 ($j = 1, \dots, h$) are given by exponentiating the elements of $\tilde{\mathbf{X}}_h$, and multiplying by $\tilde{\sigma}_Y^2$.

6 Empirical Analysis

6.1 Data and Preliminary Results

The empirical analysis focuses on estimating and forecasting the GLMSV model for three sets of exchange rate data, namely YEN/USD, EUR/USD, and GBP/USD. The sample period is from October 4, 2005 to November 25, 2015, giving 2549 observations. We calculated the returns series, $R_t = \log P_t - \log P_{t-1}$, where P_t is the closing price on day t . We use the first $n = 2048$ returns for estimating the GLMSV models, and the remaining 500 series for forecasting. The estimation period includes the global financial crisis. Table 2 presents descriptive statistics for the whole sample. As our interest is on volatility, we use the mean subtracted returns, $Y_t = R_t - \bar{R}$.

As a preliminary analysis, we estimated the new generalized fractionally integrated EGARCH (GIEGARCH) model, defined by:

$$Y_t = \sqrt{h_t}\xi_t, \quad \xi_t \sim IID(0, 1),$$

$$\phi(L)(1 - 2\eta L + L^2)^d(\log h_t - \mu) = \theta(L)\zeta(\xi_{t-1}),$$

where $g(\xi_t)$ is the generalized return, and $\phi(L)$ and $\theta(L)$ are defined in Section 3. Following Hansen, Huang, and Shek (2012), we consider the second-order Hermite polynomial for the

Table 3: QML Estimates of FIEGARCH and GIEGARCH for Daily Currency Returns

Parameters	YEN/USD		EUR/USD		GBP/USD	
	FIEGARCH	GIEGARCH	FIEGARCH	GIEGARCH	FIEGARCH	GIEGARCH
μ	-0.7736 (0.0474)	-0.8589 (0.0427)	-0.7916 (0.0715)	-0.9170 (0.0776)	-0.8771 (0.0809)	-1.0338 (0.0760)
ϕ	-0.1084 (0.0278)	0.9749 (0.0034)	-0.2401 (0.0509)	0.9854 (0.0014)	-0.2006 (0.0379)	0.9881 (0.0019)
γ_1	-1.1991 (0.3571)	-0.0415 (0.0064)	-0.2439 (0.1325)	-0.0119 (0.0036)	-0.7873 (0.2147)	-0.0229 (0.0042)
γ_2	0.7254 (0.2696)	0.0321 (0.0043)	0.5071 (0.1496)	0.0140 (0.0023)	0.4779 (0.1727)	0.0108 (0.0027)
d	0.1491 (0.0345)	0.3350 (0.0750)	0.2368 (0.0365)	0.4988 (0.0854)	0.2495 (0.0431)	0.4996 (0.0624)
η	1	0.3892 (0.0026)	1	0.8583 (0.0014)	1	0.8570 (0.0006)
ω_g	0	1.1710	0	0.5388	0	0.5414

Note: FIE and GIE denote the FIEGARCH and GIEGARCH models, respectively. Standard errors are in parentheses. The Gegenbauer frequency is given by $\omega_g = \arccos(\eta)$.

error term, as:

$$\zeta(\xi_t) = \gamma_1 \xi_t + \gamma_2 (\xi_t^2 - E(\xi_t^2)).$$

Assuming that ξ_t has finite fourth moment, it is straightforward to show $E[g(\xi_t)] = 0$ and $V[g(\xi_t)] < \infty$. When $\eta = 1$, the new GIEGARCH($p, d, q; \eta$) model reduces to the class of the FIEGARCH($p, 2d, q$) model of Bollerslev and Mikkelsen (1996). Following Bollerslev and Mikkelsen (1996), we truncate the MA(∞) representation of the GARMA process of log-volatility as:

$$\log h_t = \mu + \sum_{j=0}^J \tilde{\psi}_j \zeta(\xi_{t-1-j}),$$

where $\tilde{\psi}_j$ is the j th coefficient of the polynomial $\tilde{\psi}(z) = (1 - 2\eta z + z^2)^{-d} \phi(z)^{-1} \theta(z)$, with $\tilde{\psi}_0 = 1$. We calculate the value of $\tilde{\psi}_j$ by the approximating technique of McElroy and Holan (2012) up to $J = 1000$ (see the Appendix).

Table 3 shows the QML estimates of the FIEGARCH(1,2d,0) and GIEGARCH(1,d,0; η) models. For the FIEGARCH model, the estimates of d indicate that the conditional log-

Table 4: Estimates of GLMSV for Daily Currency Returns

Parameters	YEN/USD	EUR/USD	GBP/USD
μ	-1.2366 (0.0579)	-1.2030 (0.0544)	-1.3069 (0.0538)
σ_ϵ	2.5173 (0.0414)	2.3482 (0.0384)	2.2844 (0.0368)
σ	0.0868 (0.0350)	0.1621 (0.0378)	0.0974 (0.0311)
ϕ	0.9872 (0.0066)	0.9939 (0.0042)	0.9980 (0.0039)
d	0.3173 (0.1475)	0.4702 (0.1029)	0.4987 (0.1869)
η	0.8032 (0.0001)	0.9597 (0.0003)	0.8400 (0.0009)
ω_g	0.6381	0.2849	0.5735

Note: Standard errors are in parentheses. The Gegenbauer frequency is given by $\omega_g = \arccos(\eta)$.

volatility, $\ln_h t$, has long range dependence. The estimates of γ_1 are negative, while those of γ_2 are positive. The estimates of ϕ are located in the interval $(-0.25, -0.1)$. Except for the estimates of γ_1 for the EUR/USD return, all parameter estimates are significant at the five percent level. These estimates are similar to the values obtained in the literature.

The estimates of d in the GIEGARCH model are about twice of those for the FIEGARCH model. The estimates of η are positive, and the estimates of ϕ are close to one. The estimates of γ_1 are negative, while those of γ_2 are positive. All parameter estimates are significant at the five percent level. As the estimates of η are significantly different from one, the estimates of the Gegenbauer frequency, $\omega_g = \arccos(\eta)$, are different from zero.

6.2 Estimates and Forecasts for the GLMSV Model

In the following, we show the empirical results for the GLMSV models as compared with those of the GIEGARCH model.

Table 4 gives the SL estimates of the GLMSV model. The estimates of d and ϕ are close to the values of the GIEGARCH model. Compared with the GIEGARCH model, the estimates of η are higher. The estimates of μ are different from those of the GIEGARCH

model, and the differences may arise from the statistical flexibility of the class of SV models compared with their conditional heteroskedasticity counterparts. All estimates are significant at the five percent level. As the estimates of η are significantly different from one, the estimates of the Gegenbauer frequency $\omega_g = \arccos(\eta)$ are different from zero.

As explained previously, we use the last 500 observations for the forecasting analysis, based on the approach in the previous section. For this purpose, we calculated the Value-at-Risk (VaR) thresholds, assuming normality of ξ_t . Combined with the one-day-ahead forecasts of log-volatility, we computed the 1 and 5 percent VaR thresholds as $-2.326\hat{\sigma}_{n+1}^2$ and $-1.645\hat{\sigma}_{n+1}^2$, respectively, fixing the sample size as $n = 2048$.

In order to assess the estimated VaR thresholds, the unconditional coverage and independence tests developed by Christoffersen (1998) are widely used. A drawback of the Christoffersen (1998) test for independence is that it tests against a particular alternative of first-order dependence. The duration-based approach in Christoffersen and Pelletier (2004) allows for testing against more general forms of dependence, but still requires a specific alternative. Recently, Candelon et al. (2011) developed a more robust procedure which does not need a specific distributional assumption for the durations under the alternative. Consider the ‘hit sequence’ of VaR violations, which takes a value of one if the loss is greater than the VaR threshold, and the value zero if VaR is not violated. If we could predict the VaR violations, then that information may help to construct a better model. Hence, the hit sequence of violations should be unpredictable, and should follow an independent Bernoulli distribution with parameter p , indicating that the duration of the hit sequence should follow a geometric distribution.

The GMM duration-based test developed by Candelon et al. (2011) works with the J -statistic based on the moments defined by the orthonormal polynomials that are associated with the geometric distribution. The conditional coverage test and independence test based on q orthonormal polynomials have the asymptotic χ_q^2 and χ_{q-1}^2 distributions under their respective null distributions. The unconditional coverage test is given as a special case of the conditional coverage test, with $q = 1$.

Table 5 shows the percentage of VaR violations and test results for the FIEGARCH, GIEGARCH and GLMSV models. For the FIEGARCH model, some of the test statistics are rejected at the five percent significance level. On the other hand, for the GIEGARCH

Table 5: Backtesting VaR Thresholds for One-Step-Ahead Forecasts

(a) YEN/USD					
VaR	PV	UC	IND	CC	
FIEGARCH(1,d,0)					
5%	0.034	0.9752 [0.3234]	8.1759 [0.0853]	76.331*	[0.0000]
1%	0.004	0.5989 [0.4390]	0.9347 [0.9195]	2.0605	[0.8407]
GIEGARCH(1,d,0; η)					
5%	0.036	0.5029 [0.4782]	2.1775 [0.7032]	5.4058	[0.3684]
1%	0.004	1.5034 [0.2201]	1.1999 [0.8781]	3.7080	[0.5922]
GLMSV(1,d,0; η)					
5%	0.040	0.0733 [0.7866]	5.2835 [0.2594]	7.2375	[0.2036]
1%	0.010	1.2558 [0.2625]	1.6877 [0.7930]	1.6877	[0.8905]

(b) EUR/USD					
VaR	PV	UC	IND	CC	
FIEGARCH(1,d,0)					
5%	0.058	2.0866 [0.1486]	1.3895 [0.8460]	3.0081	[0.6987]
1%	0.010	0.3457 [0.5566]	1.7610 [0.7796]	1.7610	[0.8811]
GIEGARCH(1,d,0; η)					
5%	0.054	0.3174 [0.5732]	0.0726 [0.9994]	0.3424	[0.9968]
1%	0.008	0.0097 [0.9214]	0.2397 [0.9934]	0.1625	[0.9995]
GLMSV(1,d,0; η)					
5%	0.044	0.6177 [0.4319]	3.5391 [0.4720]	6.6769	[0.2458]
1%	0.018	2.0001 [0.1573]	0.9180 [0.9220]	2.8541	[0.7225]

(c) GBP/USD					
VaR	PV	UC	IND	CC	
FIEGARCH(1,d,0)					
5%	0.048	0.0193 [0.8894]	18.555*	23.801*	[0.0002]
1%	0.004	0.4137 [0.5201]	0.8417 [0.9328]	1.7850	[0.8780]
GIEGARCH(1,d,0; η)					
5%	0.050	0.0281 [0.8669]	1.0782 [0.8977]	1.0782	[0.9560]
1%	0.004	2.0368 [0.1535]	1.2781 [0.8651]	4.6977	[0.4539]
GLMSV(1,d,0; η)					
5%	0.052	1.7100 [0.1910]	7.1894 [0.1262]	3.9607	[0.5551]
1%	0.012	2.3767 [0.1232]	2.7581 [0.5991]	3.8607	[0.5697]

Note: PV denotes the percentage of violations, which is the percentage of days when returns are less than the VaR threshold. UC, IND, and CC are the generalized method of moments duration-based tests for unconditional coverage, independence and conditional coverage, developed by Candelon et al. (2011). The number of orthonormal polynomials is set to 5. P values are in brackets.

and GLMSV models, the tests do not reject the null hypothesis at the 5% and 1% VaR thresholds, thereby indicating that the estimated VaR thresholds are satisfactory.

7 Concluding Remarks

In this paper, we proposed a new generalized long memory volatility (GLMSV) model, based on the GARMA($p, d, q; \eta$) process, and examined the statistical properties of the new model. We applied the spectral likelihood (SL) estimation method, for which the asymptotic distribution is the same as that of the QML estimator. Then we conducted Monte Carlo experiments for investigating the finite sample properties of the SL estimator, and found that the finite sample biases are negligible for $n = 2048$.

In addition, we estimated the FIEGARCH, GIEGARCH, and GLMSV models, using three exchange rate returns for YEN/USD, EUR/USD, and GBP/USD. The empirical results supported long memory for log-volatility, and also showed a non-zero Gegenbauer frequency. Furthermore, the new specification of generalized long memory improved the out-of-sample forecasts for the VaR thresholds satisfactorily, which shows that the GLMSV model is a useful addition to the existing models in the literature.

Appendix

We explain the calculation of the coefficients of the $\text{MA}(\infty)$ representation of the $\text{GARMA}(p, d, q; \eta)$ model, and the calculation of the autocovariance functions.

For the GARMA process, it is not easy to obtain explicit formulas for the MA coefficients and the autocovariances that are valid for all lags. Recently, McElroy and Holan (2012) developed a computationally efficient method for calculating these values. Using the Gegenbauer frequency, $\lambda = \omega_g$, the spectral density of X_t can be written as:

$$f(\omega) = \sigma^2 |1 - e^{-i\lambda} e^{-i\omega}|^{-2d} |1 - e^{i\lambda} e^{-i\omega}|^{-2d} g(\omega),$$

where $g(\omega)$ represents the short memory part of the spectrum. For convenience, we define $\kappa(z)$ so that $g(\omega) = |\kappa(e^{-i\omega})|^2$. Then, $\kappa(z)$ takes the form $\kappa(z) = \prod_l (1 - \zeta_l z)^{p_l}$ for (possibly complex) reciprocal roots, ζ_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.

Define:

$$\begin{aligned} g_j &= -2 \sum_l \frac{p_l \zeta_l^j}{j}, \\ \beta_j &= \frac{4d \cos(j\omega)}{j} + g_j, \\ \tilde{\psi}_j &= \frac{1}{2j} \sum_{m=1}^l m \beta_m \tilde{\psi}_{j-m}, \quad \tilde{\psi}_0 = 1. \end{aligned}$$

McElroy and Holan (2012) showed that the $\text{MA}(\infty)$ representation of (4.2) is given by:

$$X_t = \mu + \sum_{j=0}^{\infty} \tilde{\psi}_j v_{t-j},$$

and the autocovariances of X_t for $h \geq 0$ are given by:

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

where

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

and $F(a, b; c; z)$ is the hypergeometric function evaluated at z . Note that $\gamma_{-h} = \gamma_h$. McElroy and Holan (2012) recommend using the cutoff value $J \geq 2,000$.

References

- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2001). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, **96**, 42-55.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., (2003). Modeling and forecasting realized volatility, *Econometrica*, **71**, 529-626.
- Arteche, J. (2004). Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models, *Journal of Econometrics*, **119**, 131-154.
- Asai, M., McAleer, M., Medeiros, M.C. (2012). Asymmetry and long memory in volatility modeling, *Journal of Financial Econometrics*, **10**, 495-512.
- Baillie R.T., Bollerslev, T., Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **74**, 3-30.
- Barndorff-Nielsen, O.E., Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B*, **64**, 253-280.
- Bollerslev, T., Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility, *Journal of Econometrics*, **109**, 33-65.
- Bollerslev, T., Mikkelsen, H.O. (1996). Modeling and pricing long-memory in stock market volatility, *Journal of Econometrics*, **73**, 151-184.
- Breidt, F.J., Crato, N., de Lima, P. (1998). The detection and estimation of long memory, *Journal of Econometrics*, **83**, 325-348.
- Candelon, B., Colletaz, G., Hurlin, C., Tokpavi, S. (2011), Backtesting value-at-risk: A GMM duration-based approach, *Journal of Financial Econometrics*, **9**, 314-343.
- Christoffersen, P. (1998). Evaluating interval forecasts, *International Economic Review*, **39**, 841-862.
- Christoffersen, P., Pelletier, D. (2004). Backtesting value-at-risk: A duration-based approach, *Journal of Financial Econometrics*, **2**, 84-108.
- Chung, C.F. (1996a). Estimating a generalized long memory process, *Journal of Econometrics*, **73**, 237-259.
- Chung, C.F. (1996b). A generalized fractionally integrated autoregressive moving-average process, *Journal of Time Series Analysis*, **17**, 111-140.
- Clark, P.K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica*, **41**, 135-156.
- Deo, R.S., Hurvich, C.M. (2001). On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models, *Econometric Theory*, **17**, 686-710.
- Dissanayake, G., Peiris, S., Proietti, T. (2016). State space modeling of Gegenbauer processes with long memory, *Computational Statistics & Data Analysis*, **100**, 115-130.

- Doornik J.A. Ooms, M. (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models, *Computational Statistics & Data Analysis*, **42**, 333–348.
- Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987-1007.
- Granger, C.W.J., Morris, M. (1976). Time series modeling and interpretation, *Journal of the Royal Statistical Society, Series A*, **139**, 246-257.
- Gray, H.L., Zhang, N., Woodward, W.A. (1989). On generalized fractional processes, *Journal of Time Series Analysis*, **10**, 233-257.
- Hansen, P.R., Z. Huang, and H.H. Shek (2012). Realized GARCH: A Complete Model of Returns and Realized Measures of Volatility, *Journal of Applied Econometrics*, **27**, 877–906.
- Harvey, A. (1998). Long memory in stochastic volatility, in: Knight, J. and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, Oxford: Butterworth-Haineman, 307–320.
- Hosoya, Y. (1997). A limit theory for long-range dependence and statistical inference on related models, *Annals of Statistics*, **25**, 105–137.
- Koopman, S.J., Jungbacker, B., Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements, *Journal of Empirical Finance*, **12**, 445-475.
- McAleer, M. (2005). Automated inference and learning in modeling financial volatility, *Econometric Theory*, **21**, 232-261.
- McElroy, T.S., Holan, S.H. (2012). On the computation of autocovariances for generalized Gegenbauer processes, *Statistica Sinica*, **22**, 1661–1687.
- Pong S., Shackelton, M.B., Taylor, S.J., Xu, X. (2004). Forecasting currency volatility: a comparison of implied volatilities and AR(FI)MA models, *Journal of Banking and Finance*, **28**, 2541-2563.
- Sandmann, G., Koopman, S.J. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood, *Journal of Econometrics*, **87**, 271–301.
- Shephard, N. (2005). General introduction, in N. Shephard (ed.), *Stochastic Volatility*, Oxford University Press, Oxford, pp.1–33.
- Shitan, M., Peiris, M. (2008). Generalized autoregressive (GAR) model: A comparison of maximum likelihood and Whittle estimation procedures using a simulation study, *Communications in Statistics - Theory and Methods*, **37(3)**, 560–570.
- Shitan, M., Peiris, M. (2013). Approximate asymptotic variance-covariance matrix for the Whittle estimators of GAR(1) parameters, *Communications in Statistics - Theory and Methods*, **42(5)**, 756–770.
- So, M.K.P. (1999). Time series with additive noise, *Biometrika*, **86**, 474–482.

- So, M.K.P. (2002). Bayesian analysis of long memory stochastic volatility models, *Sankhyā*, **24**, 1–10.
- Taniguchi, M. (1982). On estimation of the integrals of the fourth order cumulant spectral density, *Biometrika*, **69**, 117–122.
- Taniguchi, M., Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*, New York: Springer-Verlag.
- Taylor, S.J. (1982). Financial returns modeled by the product of two stochastic processes - a study of daily sugar prices 1961-79, in: Anderson, O.D. (ed.), *Time Series Analysis: Theory and Practice*, vol.1, Amsterdam: North-Holland, 203-226.
- Taylor, S.J. (1986). *Modelling Financial Time Series*, Chichester: Wiley.
- Zaffaroni, P. (2009). Whittle estimation of EGARCH and other exponential volatility models, *Journal of Econometrics*, **151**, 190–200.