# Cointegrated Dynamics for A Generalized Long Memory Process: An Application to Interest Rates*

**Manabu Asai**
Faculty of Economics, Soka University, Japan

**Shelton Peiris**
School of Mathematics and Statistics, University of Sydney, Australia

**Michael McAleer**
Department of Finance, Asia University, Taiwan
Discipline of Business Analytics, University of Sydney Business School, Australia
Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam
The Netherlands
Department of Economic Analysis and ICAE, Complutense University of Madrid, Spain
Institute of Advanced Sciences, Yokohama National University, Japan

**David E. Allen**
School of Mathematics and Statistics, University of Sydney, Australia
Department of Finance, School of Business and Law, Edith Cowan University, Australia
Department of Finance, Asia University, Taiwan

August 2018

**EI2018-32**

*Initial draft of this article was written while the first author was a Visiting Scholar at the University of Sydney. The authors are most grateful to Yoshi Baba for very helpful comments and suggestions. The first author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science, School of Mathematics and Statistics at The University of Sydney, and the Australian Academy of Science. The second author acknowledges the support from the Faculty of Economics at Soka University. The third author is most grateful for the financial support of the Australian Research Council, Ministry of Science and Technology (MOST), Taiwan, and the Japan Society for the Promotion of Science. The fourth author acknowledges the Australian Research Council.

## Abstract

Recent developments in econometric methods enable estimation and testing of general long memory process, which include the general Gegenbauer process. This paper considers the error correction model for a vector general long memory process, which encompasses the vector autoregressive fractionally-integrated moving average and general Gegenbauer process. We modify the tests for unit roots and cointegration, based on the concept of heterogeneous autoregression. The Monte Carlo simulations show that the finite sample properties of the modified tests are satisfactory, while the conventional tests suffer from size distortion. Empirical results for interest rates series for the U.S.A. and Australia indicate that: (1) the modified unit root test detected unit roots for all series, (2) after differencing, all series favour the general Gegenbauer process, (3) the modified test for cointegration found only two cointegrating vectors, and (4) the zero interest rate policy in the U.S.A. has no effect on the cointegrating vector for the two countries.

**Keywords:** Long Memory Processes; Gegenbauer Process; Dickey-Fuller Tests; Cointegration; Differencing; Interest Rates.

**JEL Classification:** C22, C32, C51.

# 1 Introduction

In the early 90's there was extensive analysis conducted on macroeconomic variables as to whether they are integrated of order zero, $I(0)$, or of order one, $I(1)$, using the augmented Dickey-Fuller (ADF) and/or Phillips-Perron tests (see Dolado, Jenkinson, and Sosvilla-Rivero (1990)). Regarding interest rates, Gonzalo and Granger (1995), Johansen (1997), Juselius (1995), and MacKinnon (1996), among others, found that these time series are $I(1)$, and conducted further analysis of cointegration (see also Czudaj and Pruser (2015) for recent results for the U.S.A. and Germany). For the possibility of long memory process with fractional integration of order $d$, $I(d)$ with $|d| < 1$, Baillie (1996) summarized existing empirical results, including the work of Backus and Zin (1993), by stating that some of the initial research found evidence of lo n g memory in the spread and some interest rates in levels. However, Baillie (1996, p.50) claims "The estimation of various ARFIMA (autoregressive fractionally integrated moving-average) models to bond series is relatively inconclusive."

As an extension of fractional integration, Gray, Zhang, and Woodward (1989) developed the Gegenbauer process, based on Gegenbauer polynomials. While the spectral density of the ARFIMA process is unbounded at the origin, the Gegenbauer process has a peak at a frequency away from the origin, which is referred to as the Gegenbauer frequency. As suggested in Woodward, Cheng, and Gray (1998), a general Gegenbauer (GG) process has multiple (unbounded) peaks (see Figure 1 as an example). Caporale and Gil-Alana (2011) and Ferrara and Guégan (2001) show empirical results that inflation rate data fit the GG process better than does the ARFIMA process, Ramachandran and Beaumont (2001) and Smallwood and Norrbin (2008) obtained equivalent results for interest rates.

As Ramachandran and Beaumont (2001) and Smallwood and Norrbin (2008) estimated one factor GG process, without checking the sample spectral densities of interest rates, so there are opportunities to improve upon these results. We can extend their work by estimating the GG process, using the techniques developed in Chan and Tsai (2012), Hidalgo and Soulier (2004), McElroy and Holan (2012, 2016), and Tsai, Rachinger, and Lin (2015). In addition, the previous literature considers tests of unit roots against long memory processes, but we may also consider that the differenced series may have long memory properties such as a stationary multifactor GG process. In other words, we consider integrated GG (IGG) processes, such that the differenced series are stationary GG processes with multiple Gegenbauer frequencies, which may be different from zero. If the sample spectral density of a GG process has no peak at the origin, as in Figure 1(b), there is a non-negligible possibility that the ADF test shows that they are $I(0)$, corresponding to the well-known empirical results.

Consider the concept of (fractional) cointegration. A time series with a spectrum which is finite but non-zero at all frequencies is defined as $I(0)$. If a series, $x_t$, needs differencing $d$ ($d > 0$) times to become $I(0)$, it is called integrated of order $d$, denoted $x_t \sim I(d)$, where $d$ may take fractional values. For an $m \times 1$ vector process $\boldsymbol{x}_t$, let each element of $\boldsymbol{x}_t$ be $I(d)$. If there exists a vector $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}'\boldsymbol{x}_t \sim I(d-b)$ ($b > 0$), where $b$ can take fractional values, then Granger (1986) called $x_t$ cointegrated of order $(d, b)$. For analyzing IGG processes, we can relax the concept by allowing $I(0)$ to have unbounded peaks at the frequencies on $(0, \pi)$, which excludes $\{0, \pi\}$, under the condition that the process is stationary. In other words, we discuss integration and cointegration with respect to the unbounded peak at the origin of the spectral densities, as implicitly assumed by Gonzalo and Granger (1995). Note that we can still consider fractional cointegration under this weak concept.

In this paper, we consider a general error correction model (ECM) for a vector long memory process, which encompasses the vector IGG processes, and suggest modified tests for unit roots and cointegration for general long memory processes. For the latter purpose, we need to consider approximations of the finite order (V)AR model, in the sense of Said and Dickey (1984). For computational convenience, we suggest using the heterogeneous autoregressive (HAR) model of Corsi (2009) to approximate long memory components. The Monte Carlo experiments show that the conventional tests for unit roots over-reject the null hypothesis of no unit roots, and the conventional tests for cointegration suffer from size distortion if the underlying process follows the IGG model. The Monte Carlo simulations indicate that the modified tests are satisfactory.

The organization of the remainder of the paper is as follows. Section 2 introduces the ECM for vector long memory processes and the (vector) IGG process. Section 3 suggests modified tests for unit roots and cointegration, and conducts Monte Carlo experiments for investigating their finite sample properties. Section 4 shows the empirical results for monthly data of interest rates of the U.S.A. and Australia, and finds that the cointegrated dynamics of the two countries is not affected by the recent zero interest rate policy. Section 5 gives some concluding remarks.

# 2 Generalized Long Memory and Vector Error Correction Model

## 2.1 Error Correction Model for Vector Long Memory Process

Let $\boldsymbol{y}_t$ be a $m \times 1$ vector of $I(1)$ time series. Assume that the rank of cointegration is $r$, that is, there exists a matrix $\alpha$ $(m \times r)$ of rank $r$, such that $\alpha' \boldsymbol{y}_t$ is $I(0)$. In other words, $r$ is the number of stationary series derived by linear combinations of the elements of $\boldsymbol{y}_t$. The vector $\boldsymbol{y}_t$ has an ECM representation:

$$\Delta \boldsymbol{y}_t = \gamma \alpha' \boldsymbol{y}_{t-1} + \sum_{i=1}^{\infty} B_i \Delta \boldsymbol{y}_{t-i} + \boldsymbol{c} + \boldsymbol{\varepsilon}_t, \tag{1}$$

where $\Delta = 1 - L$ with lag operator $L$, $\gamma$ is an $m \times r$ matrix, $B_i$ are $m \times m$ matrices of parameters, and $\boldsymbol{c}$ is an $m \times 1$ vector. The ECM (1) is different from that used in Johansen and Juselius (1990), as it includes the infinite-order vector autoregressive terms. As discussed in Gonzalo and Granger (1995), we can use the ECM (1) for the analysis of an $I(1)$ process with long memory.

Gonzalo and Granger (1995) suggested decomposing the permanent and transitory components as:

$$\boldsymbol{y}_t = \boldsymbol{y}_t^P + \boldsymbol{y}_t^T, \quad \boldsymbol{y}_t^P = \alpha_\perp (\gamma_\perp' \alpha_\perp)^{-1} \gamma_\perp \boldsymbol{y}_t, \quad \boldsymbol{y}_t^T = \gamma (\alpha' \gamma)^{-1} \alpha \boldsymbol{y}_t,$$

under the assumption that $\boldsymbol{y}_t$ can be approximated by the VAR model of small order. Here, $\alpha_\perp$ is $m \times (m-r)$ matrix, which satisfies $\alpha_\perp' \alpha = O$, while $\gamma_\perp$ $(m \times (m-r))$ is defined to obtain $\gamma_\perp' \gamma = O$. Note that $\alpha \boldsymbol{y}_t$ is an $r \times 1$ vector, while $\gamma_\perp \boldsymbol{y}_t$ is an $(r-m) \times 1$ vector. Multiplying $\gamma_\perp$ from the left of ECM (1), we notice that $\Delta(\gamma_\perp \boldsymbol{y}_t)$ can be expressed as a stationary infinite-order AR model, which includes long memory processes. In their empirical analysis, Gonzalo and Granger (1995) estimated $\alpha$ and $\gamma_\perp$ for VAR(3), as it is difficult to accompany higher-order lags under the sample size for monthly economic time series.

In order to apply the above framework, we consider a vector process of general long memory which encompasses the vector ARFIMA process, by introducing the so-called Gegenbauer long memory.

### 2.2 Univariate and Multivariate Generalized Gegenbauer Processes

Let $y_t$ follow a univariate generalized Gegenbauer (GG) process, defined by:

$$\phi(L)P^*(L)(y_t-\mu)=\theta(L)\varepsilon_t,\quad \varepsilon_t\sim \mathrm{iid}(0,\sigma^2), \tag{2}$$

$$P^*(L)=(1-L)^a(1+L)^b\prod_{l=1}^{k}(1-2\cos(\omega_l)L+L^2)^{d_l}, \tag{3}$$

where $\varepsilon_t$ is independently and identically distributed, $L$ is the lag operator, $\phi(L)=1-\phi_1L-\ldots-\phi_pL^p$, and $\theta(L)=1+\theta_1L+\ldots+\theta_qL^q$. Equation (2) is also known as $k$-factor generalized exponential model (see McElroy and Holan (2012)), and it reduces to the multi-factor Gegenbauer process suggested by Woodward, Cheng, and Gray (1998), by setting $a=b=0$. The process is based on the Gegenbauer polynomials defined by $(1-2\cos(\omega_l)z+z^2)^{-d_l}=\sum_{j=0}^{\infty}c_jz^j$. The structure of $(1-2\cos(\omega_l)L+L^2)^{d_l}$ encompasses $(1-L)^a$ and $(1+L)^b$. We obtain $(1-2\cos(\omega_l)L+L^2)^{d_l}$ as $(1-L)^a$ when $\omega_l=0$ with $a=2d_l$, and we have $(1+L)^b$ when $\omega_l=\pi$ with $b=2d_l$, but we start from the above specification for practical convenience.

As discussed in Bordignon, Caporin, and Lisi (2007) and McElroy and Holan (2012), the specification covers various long memory models. For instance, noting that $(1-L)^{12}=(1-L)(1+L)(1+L+L^2)(1-L+L^2)(1+\sqrt{3}L+L^2)(1-\sqrt{3}L+L^2)$, we have a seasonal long memory process for monthly data, $\phi(L)(1-L^{12})^d(y_t-\mu)=\theta(L)\varepsilon_t$ (see Porter-Hudak (1990)), by setting $a=b=d_l=d$, $k=4$, $\omega_1=\pi/6$, $\omega_2=\pi/3$, $\omega_3=2\pi/3$, and $\omega_4=5\pi/6$. As economic variables typically have unit roots, we will focus on the structure:

$$P^*(L)=P(L)\Delta,\quad \Delta=1-L,\quad P(L)=(1-L)^{d_0}\prod_{l=1}^{k}(1-2\cos(\omega_l)L+L^2)^{d_l}, \tag{4}$$

by setting $a=d_0+1$ and $b=0$, without loss of generality.

We assume the roots of $\phi(z)$ and $\theta(z)$ lie outside the unit circle to ensure stationarity and invertibility of $P(L)\Delta y_t$, respectively. For the stationarity of the long memory structure of $P(L)$, we assume $|d|<1/2$, $|d_l|<1/2$, and $0<\omega_l<\pi$ (see Woodward, Cheng, and Gray (1998), and McElroy and Holan (2012)). By the parameter restrictions, $\Delta y_t$ is stationary, while $y_t$ is nonstationary. For notational convenience, we will refer to the processes $\Delta y_t$ and $y_t$ as the GG and integrated GG (IGG) processes, respectively

The power spectrum of the GG process, obtained by differencing the IGG models (2) and (4), is given by:

$$f(\lambda) = \frac{\sigma^2}{2\pi} g(\lambda)[2\sin(\lambda/2)]^{-2d_0} \prod_{l=1}^{k} [2(\cos\lambda - \cos\omega_l)]^{-2d_l}, \quad -\pi < \lambda < \pi, \tag{5}$$

where $g(\lambda) = \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}$ corresponds to the autoregressive moving-average (ARMA) part. The power spectrum shows the long memory feature characterized by the unbounded spectrum at the Gegenbauer frequencies, $\omega_l$ $(l = 1, \ldots, k)$. For the IGG process, the power spectral density is unavailable as it does not have a finite variance. We may define a pseudo spectrum by multiplying $[2(1 - \cos\omega)]^{-1}$ by $f(\lambda)$. Figure 1 illustrates the (pseudo-)power spectrum of $y_t$ and $\Delta y_t$ with $P(L) = (1 - 2\cos(2\pi/5)L + L^2)^{0.3}(1 - 2\cos(2\pi/3)L + L^2)^{0.2}$ and $g(\lambda) = 1$. Figure 1 implies that using the first differences affects the spectrum near zero and the Gegenbauer frequencies.

We can identify the number $k$, and estimate the location parameter, $\omega_l$, and long memory parameter, $d_l$, by the technique of Hidalgo and Soulier (2004), which will be explained in Appendix A.1. Since $\omega_l$ is a frequency, $(1-2\cos(\omega_l)L+L^2)^{d_l}$ produces periodic long-memory with cycles every $2\pi/\omega_l$, and $2\pi/\omega_l$ is expected to be an integer for economic time series (see empirical analysis of Bordignon, Caporin, and Lisi (2007), Caporale and Gil-Alana (2011), and Peiris and Asai (2016), for instance). Peiris and Asai (2016) estimated $\omega_l$ for monthly data of inflation rates, and found that the estimates were 2, 3, 4, and 6 months cycles.

When we have identified the location parameters with cycles of integer numbers, we can estimate the GG model for $\Delta \boldsymbol{y}_t$ with the Whittle likelihood (WL) estimator. Given $\omega_l$, it is straightforward to show asymptotic normality of the WL estimator based on the results of Chan and Tsai (2012) and Tsai, Rachinger, and Lin (2015).

As a straightforward extension of the univariate case (see, for example, Wu and Peiris (2018)), we define a vector IGG (VIGG) process for an $m \times 1$ vector $\boldsymbol{y}_t$ by:

$$\Phi(L)\boldsymbol{P}^*(L)(\boldsymbol{y}_t - \boldsymbol{\mu}) = \Theta(L)\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \text{iid}(0, \Sigma), \tag{6}$$

where $\boldsymbol{P}^* = \text{diag}\{P_1^*(L), \ldots, P_m^*(L)\}$, with $P_i^*(L) = P_i(L)\Delta$ defined as in equation (4), $\Phi(L) = 1 - \Phi_1 L - \ldots - \Phi_p L^p$, and $\Theta(L) = 1 + \Theta_1 L + \ldots + \Theta_q L^q$. We assume that $\Phi(L)$ and $\Theta(L)$ satisfy the stationary and invertibility conditions as in the conventional VAR model. We allow each $P_i(L)$ to have different $k$, $d$, $d_l$'s and $\omega_l$'s, depending on $i$ ($i = 1, \ldots, m$) satisfying the stationary condition for (4).

# 3 Modified Tests for Unit Root and Cointegration

## 3.1 Approximated ECM

As discussed in Said and Dickey (1984), we can approximate an ARMA process with an AR($p$) model, where $p$ increases more slowly than the sample size, and we can use the augmented Dickey-Fuller test with the same limiting distribution. In the same manner, the VAR($p$) process is used for the analysis of cointegration and the ECM. For a short memory vector process, the VAR($p$) is a suitable approximate model. However, we need to consider a different type of approximation for long memory vector processes.

We approximate the ECM (1) using the heterogeneous autoregressive (HAR) model that was suggested by Corsi (2009). Define the $h$ period average of past information of $\Delta \boldsymbol{y}_t$ as:

$$\left(\Delta \boldsymbol{y}_{t-1}\right)_h = \frac{1}{h}\left(\Delta \boldsymbol{y}_{t-1} + \Delta \boldsymbol{y}_{t-2} + \cdots + \Delta \boldsymbol{y}_{t-h}\right),$$

and consider

$$\Delta \boldsymbol{y}_t = \gamma \alpha' \boldsymbol{y}_{t-1} + \sum_{j=1}^{\infty} C_j \left(\Delta \boldsymbol{y}_{t-1}\right)_{h_j} + \boldsymbol{c} + \boldsymbol{\varepsilon}_t,$$

where $C_j$ are $m \times m$ matrices of parameters, and $\{h_j : j = 1, 2, \ldots\}$ is a given sequence of positive integers increasing to $\infty$. The model has the representation (1), which is obtained by setting, $A_j = \sum_{k=j}^{\infty} h_k^{-1} C_k$, $B_1 = A_1$, $B_{h_j+i} = A_{j+1}$ ($i = 1, 2, \ldots, h_{j+1} - h_j$) for $j = 1, 2, \ldots$, with $h_0 = 0$ (see Hwang and Shin (2013, 2014) and Lee (2014), for instance).

Now consider the following two kinds of approximations, defined by:

$$\Delta \boldsymbol{y}_t = \gamma \alpha' \boldsymbol{y}_{t-1} + \sum_{i=1}^{p-1} B_i \Delta \boldsymbol{y}_{t-i} + \boldsymbol{c} + \boldsymbol{\varepsilon}_t^{**}, \tag{7}$$

$$\Delta \boldsymbol{y}_t = \gamma \alpha' \boldsymbol{y}_{t-1} + \sum_{j=1}^{p-1} C_j \left(\Delta \boldsymbol{y}_{t-1}\right)_{h_j} + \boldsymbol{c} + \boldsymbol{\varepsilon}_t^{*}, \tag{8}$$

where the former process is the VAR($p$) for $\boldsymbol{y}_t$, while the latter is the heterogeneous VAR (HVAR) model of order $p$. The number of parameters is the same in the two specifications, but the latter model can capture longer range dependence as compared with the former. Regarding the HVAR(6) process for monthly data, we may set $h_1 = 1$, $h_2 = 6$, $h_3 = 12$, $h_4 = 24$, $h_5 = 48$ to obtain the average effects of the past one month, 6 months, 1 year, 2 years, and 4 years, respectively. In this case, the HVAR(6) can be considered as VAR(49) with appropriate parameter restrictions.

By the HVAR approximation, we can interpret the P-T decomposition of Gonzalo and Granger (1995) differently, as discussed in Proposition 1 of Gonzalo and Granger (1995). By approximating the long memory process, $\boldsymbol{y}_t^T$ now captures long-range dependencies, which is inappropriate to be called the transitory component. We consider that the decomposition is for the stationary and nonstationary components of the vector long memory process, $\boldsymbol{y}_t = \boldsymbol{y}_t^{(s)} + \boldsymbol{y}_t^{(n)}$, where

$$\boldsymbol{y}_t^{(s)} = \gamma(\alpha'\gamma)^{-1}\alpha\boldsymbol{y}_t \quad \text{and} \quad \boldsymbol{y}_t^{(n)} = \alpha_\perp(\gamma_\perp'\alpha_\perp)^{-1}\gamma_\perp\boldsymbol{y}_t, \tag{9}$$

are the stationary and nonstationary components, respectively.

In the following, we modify the tests for unit roots and cointegration, based on the above approach.

### 3.2 Modified Tests for Unit Roots

We develop the modified version of the augmented Dickey-Fuller (ADF) test. Consider the following two kinds of augmented regressions:

$$\Delta y_t = \alpha y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + c + \varepsilon_t^{**}, \tag{10}$$

$$\Delta y_t = \alpha y_{t-1} + \sum_{j=1}^{p-1} \gamma_j \left(\Delta y_{t-1}\right)_{h_j} + c + \varepsilon_t^{*}. \tag{11}$$

The former is used for the ADF test, while the latter is for its modified version. The ADF test examines the negative of the parameter $\alpha$ based on its regression $t$ ratio. For the conventional ADF test with (10), Cheung and Lai (1995) provide the critical values based on their response surface methodology, as:

$$\begin{aligned} CV_{0.10} &= -2.566 - 1.319/T - 15.086/T^2 + 0.667(p-1)/T - 0.650(p-1)^2/T^2, \\ CV_{0.05} &= -2.857 - 2.675/T - 23.558/T^2 + 0.748(p-1)/T - 1.077(p-1)^2/T^2, \end{aligned} \tag{12}$$

where $CV_{0.10}$ and $CV_{0.05}$ are the critical values for the 10 and 5 percent significance levels, respectively. For the modified version, we replace $p$ with $\min(h_{p-1}+1, \frac{5}{18}T)$, for which the upper bound is based on the experimental design of Cheung and Lai (1995).

### 3.3 Modified Tests for Cointegration and Estimation of Cointegrating Vectors

The conventional tests for cointegration are based on (7), while the modified test uses (8). The tests developed by Johansen (1988, 1991) examine the rank of the matrix $\Pi = \gamma\alpha'$, that is $r$. As shown in Johansen (1988, 1991), we can obtain the ML estimator of $\alpha$ for (7) by using the least squares residuals $\boldsymbol{r}_{0t}$ and $\boldsymbol{r}_{1t}$, obtained by regressing $\Delta\boldsymbol{y}_t$ and $\boldsymbol{y}_{t-1}$ on $(1, \Delta\boldsymbol{y}_{t-1}, \ldots, \Delta\boldsymbol{y}_{t-p})$, respectively. Then we obtain the concentrated model:

$$\boldsymbol{r}_{0t} = \gamma\alpha'\boldsymbol{r}_{1t} + \text{error},$$

which is known as the reduced rank regression. Johansen (1988, 1991) translated the minimization problem into a generalized eigenvalue problem:

$$|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}| = 0, \tag{13}$$

where $S_{ij}$ are the residual product matrices, $S_{ij} = T^{-1}\sum_{t=1}^{T}\boldsymbol{r}_{it}\boldsymbol{r}'_{jt}$, $(i, j = 0, 1)$, and suggested test statistics based on the $r$ largest eigenvalues, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq \cdots \lambda_m \geq 0$.

The 'trace' test statistic is the likelihood ratio (LR) test for the null hypothesis of at most $l$ cointegrating vectors $r \leq l$ $(l = m-1, \ldots, 0)$, which is given by:

$$Q_{\text{trace}} = -T\sum_{i=l+1}^{m}\ln(1-\lambda_i),$$

with $\lambda_{l+1}, \ldots, \lambda_m$ begin the $m-l$ smallest eigenvalues of $S_{10}S_{11}^{-1}S_{01}$ with respect to $S_{11}$. The 'maximal eigenvalue' LR test statistic examines the null hypothesis of $l$ cointegrating vectors against the alternative of $l+1$ cointegrating vectors $(l = m-1, \ldots, 0)$, which is given by:

$$Q_{\max} = -T\ln(1-\lambda_{l+1}).$$

For the modified tests, we need to replace $(1, \Delta\boldsymbol{y}_{t-1}, \ldots, \Delta\boldsymbol{y}_{t-p})$ with $\big(1, (\Delta\boldsymbol{y}_{t-1})_{h_1}, \ldots, (\Delta\boldsymbol{y}_{t-1})_{h_{p-1}}\big)$, in the above procedures.

Osterwald-Lenum (1992) gives the asymptotic critical values of the two LR tests, which are obtained by simulation. Our specification corresponds to Table 1.1* in Osterwald-Lenum (1992). Although Cheung and Lai (1993) provide finite sample corrections based on the response surface methodology, they use the asymptotic critical values of Johansen and Juselius (1990). We use the critical values given in Osterwald-Lenum (1992) rather than the finite sample corrections of Cheung and Lai (1993).

By the modified approach, we obtain the estimate of $\alpha$ as $\hat{\alpha} = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_r]$, where $\boldsymbol{v}_i$ is the eigenvector corresponding to the eigenvalue $\lambda_i$ $(i = 1, \ldots, r)$, which satisfies the normalization $V'S_{11}V = I$, with $V = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_m]$. Simultaneously, we obtain $\hat{\gamma} = S_{01}\hat{\alpha}$. In the decomposition (9), we need to estimate $\gamma_\perp$ and $\alpha_\perp$. Gonzalo and Granger (1995) proved that the ML estimator is obtained by the following procedure. First, solve the equation:

$$|\lambda S_{00} - S_{01}S_{11}^{-1}S_{10}| = 0,$$

which yields eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ and eigenvectors $M = [\boldsymbol{m}_1 \cdots \boldsymbol{m}_m]$, normalized such that $M'S_{00}M = I$. Using these eigenvalues and eigenvectors, we obtain the ML estimator, $\hat{\gamma}_\perp = [\boldsymbol{m}_{r+1} \cdots \boldsymbol{m}]$, and $\hat{\alpha}_\perp = S_{10}\hat{\gamma}_\perp$.

Based on (8), we can modify the test on the cointegrating vector as suggested by Johansen (1991). The null hypothesis of the test is:

$$H_0 : \alpha = G\varphi,$$

where $G$ is an $m \times s$ matrix, and $\varphi$ is an $s \times r$ vector $(r \leq s \leq m)$. The alternative hypothesis is that $\alpha$ consists of $r$ cointegrating vectors without any restrictions. We first solve:

$$|\lambda G'S_{11}G - G'S_{10}S_{00}^{-1}S_{01}G| = 0,$$

to obtain $\lambda_{g,1} \geq \ldots \geq \lambda_{g,s}$. Under the null hypothesis $\alpha = G\varphi$, the test statistic:

$$Q_g = T\sum_{i=1}^{r} \log\left((1 - \lambda_{g,i})/(1 - \lambda_i)\right), \tag{14}$$

has an asymptotic $\chi^2$ distribution with degrees of freedom given by $(m - s)r$.

### 3.4 Monte Carlo Experiments

We conduct Monte Carlo experiments for the modified tests for unit roots and cointegration, as defined above. For the modified ADF test, we consider three kinds of IGG process defined by (2) and (4), with $\mu = 0.1$, $\sigma = 1$, $d_0 = 0$, and:

$$(d_1, d_2, \omega_1, \omega_2) = \begin{cases} (0.4, 0.4, 0.4\pi, 0.8\pi) & \text{for DGP1} \\ (0.35, 0.3, 0.4\pi, 0.8\pi) & \text{for DGP2} \\ (0.4, 0.4, 0.45\pi, 0.9\pi) & \text{for DGP3.} \end{cases} \tag{15}$$

Compared with DGP1, DGP2 considers lower persistence in the long memory parameters, and DGP3 uses smaller values for the location parameters. We consider sample sizes $T = \{250, 500\}$, with $R = 2,000$ replications. We generated the data of the GG process, $\{\Delta y_t\}$, as explained in Appendix A.2, and accumulate it to obtain the IGG process, $\{y_t\}$, with an initial value of $y_{-50} = 0$. Figure 2 shows the autocorrelation functions and coefficients of the AR($\infty$) representation for the GG process. As the autocorrelation function decays slowly, the AR coefficients quickly approach values that are close to zero as the lag length increases.

Table 1 shows the rejection frequencies of the modified unit root tests, accompanied by the results of the ADF test with $p = 6$. We use nominal sizes of 5% and 10%, with the critical values given by (12). As the level series have a unit root, the rejection frequencies are expected to be close to the nominal size. On the other hand, the rejection frequencies for the first difference series should be close to one. Table 1 indicates that the modified ADF tests perform satisfactorily, while the conventional tests suffer from size distortion.

For the modified cointegration tests, the Monte Carlo experimental design is similar to Gregory, Nason, and Watt (1996). We consider the following data generating process given by a system of two variables and one cointegrating vector:

$$\begin{aligned} y_{1t} &= \mu_1 + b y_{2t} + z_t, \quad z_t = \rho z_{t-1} + u_{1t}, \quad u_{1t} \sim N(0, \sigma_u^2), \\ y_{2t} &= \mu_2 + [P^*(L)]^{-1} \varepsilon_{2t}, \quad \varepsilon_{2t} \sim N(0, \sigma_2^2), \end{aligned}$$

where $P^*(L) = (1 - L)P(L)$ as in (4), which has the error correction form:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \rho - 1 \\ 0 \end{bmatrix} z_{t-1}^* + \sum_{i=1}^{\infty} \begin{pmatrix} 0 & b\psi_i \\ 0 & \psi_i \end{pmatrix} \begin{bmatrix} \Delta y_{1,t-i} \\ \Delta y_{2,t-i} \end{bmatrix} + \begin{bmatrix} (1-\rho)\mu_1 \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{16}$$

where $z_t^* = y_{1t} - by_{2t}$, $\psi_i$'s are defined by $P(z) = 1 - \sum_{i=1}^{\infty} \psi_i z^i$, and $\varepsilon_{1t} = u_{1t} + b\varepsilon_{2t}$.

The derivation of the ECM (16) is straightforward, as it corresponds to equation (1). By the specification, $\rho = 1$ implies no cointegration relationship, while there is one cointegrating vector if $|\rho| < 1$. We generate $y_{2t}$ as the two kinds of IGG processes with parameters in (15). For the remaining parameters, we set $\mu_1 = 0.5$, $b = 1$, and $\rho = \{0.5, 1\}$. Based on the specification, we can consider six kinds of DGPs; For $l = 1, 2, 3$,

- DGP$l$co: Generate $y_{2t}$ as the IGG process with the parameters of DGP$l$, as in the ADF case, and consider one cointegrating vector for $(y_{1t}, y_{2t})$, by setting $\rho = 0.5$.

- DGP$l$nc: The same as DGP$l$co, except $\rho = 1$ for considering the no cointegrating relationship.

Table 2 reports the rejection frequencies for the conventional and modified versions of Johansen's trace and maximal eigenvalue tests. For DGPs without co integration, we expect the rejection rates to be close to the nominal 5%. While the conventional tests suffer from size distortion, the modified tests perform better. Regarding the DGPs with one cointegrating vector, Table 2 indicates that the conventional and modified tests have sufficient power. It is worth noting that the corrected critical values of Cheung and Lai (1993) were also considered, but the results are omitted as the tests tended to over-accept the null hypothesis in all cases.

The Monte Carlo results suggest that the conventional ADF tests tend to over-reject the null hypothesis of unit roots, and the modified tests for unit roots and cointegration outperform the conventional tests.

# 4 Empirical Analysis

## 4.1 Data and Preliminary Results

The purpose of this section is to examine the cointegration relationship based on the new approximate ECM model for interest rates in Australia and the U.S.A. For this purpose, we consider three kinds of interest rates with different maturities in each country, namely, short term, medium term, and long term interest rates. In Australia, the short term rate is the yield of the 3-month BABs/NCDs (Bank Accepted Bills/Negotiable Certificates of Deposit) (R3M AU), while the medium term and long-term interest rates are government bonds with maturities of 3 years (R3Y AU) and 10 years (R10Y AU), respectively.

In the U.S.A., the short term rate is the three-month Treasury bills (R3M US), the medium-term rate refers to the 3-year constant maturity government bonds (R3Y US), and the long term rate refers to the 10-year constant-maturity bonds (R10Y US). The data consist of 306 observations from 1992:7 to 2017:12, obtained from the Reserve Bank of Australia and the Federal Reserve Bank of the U.S.A. Table 1 presents the descriptive statistics of the three interest rates. Figure 3 shows the time series plots of the data. They decrease rapidly after September 2008, caused by the Global Financial Crisis (GFC). Figure 3(b) shows R3M (US) is close to zero for the period December 2008 to December 2015 due to the zero-interest rate policy. Figure 3 indicates that the trajectories are very similar in Australia. We define the vector of the raw data (in the above order) as $\boldsymbol{y}_t^\dagger$. In order to remove seasonal effects, we regress each element of $\boldsymbol{y}_t^\dagger$ on a series of monthly dummies $D_t$ as $y_{it}^\dagger = D_t\underline{\delta} + e_{it}$ $(i = 1, \ldots, 6)$. Define $y_{it} = \bar{y}_i^\dagger + \hat{e}_{it}$, where $\bar{y}_i^\dagger$ is the average interest rate and $\hat{e}_{it}$ is the residual obtained from OLS.

We use the first 48 observations as the pre-sample values, and reserve the last 12 observations for forecasting analysis, setting $T = 246$. First, we check the general fractional integration of the level and differenced series. Figures 4 and 5 show the sample spectral densities of $\boldsymbol{y}_t$ and $\Delta\boldsymbol{y}_t$, respectively. Figure 4 indicates that all level series have an unbounded peak at the origin, while Figure 5 implies that there may exist several peaks apart from the origin. The results imply that it is useful to consider the IGG model to examine the data.

We test the null hypothesis of a unit root by the ADF test and its modified version, as suggested in the previous section. We selected the lag length by the Akaike Information Criterion (AIC), $p_{aic}$. As suggested in Pantula et al. (1994) for improving finite sample performance, we also use $p_{aic2} = \min(p_{aic} + 2, p_{\max})$, with $p_{\max} = 6$. Table 4 reports the results of the $\tau$ tests based on the AR model with lag lengths, $p_{aic}$ and $p_{aic2}$, and the HAR model, with critical values calculated by (12). While we cannot reject the null hypothesis of a unit root for the level series, the test for the difference series indicates that all three series have no unit roots. The results support the findings of the sample spectral densities.

Second, we test the general integration for $\Delta y_{it}$ by the approach of Hidalgo and Soulier (2004) (see the Appendix A.1 for details). Table 5 presents the semiparametric estimates of the location parameters. Table 5 indicates that $k = 1$ is selected for ΔR3M (AU) and (US), while $k = 3$ is chosen for ΔR3Y (US). For the remaining 3 differenced series, the first candidate of the location is $\omega_1 = 0.2286\pi$. Although the corresponding modified GPH estimator is insignificant, it is worth examining the location parameter more closely. Based on the results, we may choose $\omega_1 = 2\pi \times \{1/49, 1/9, 1/9, 1/82, 1/61, 1/9\}$ for $\Delta \boldsymbol{y}_t$, respectively. For ΔR3Y (US), we select $\omega_2 = \pi/6$ and $\omega_3 = 2\pi/3$.

Third, with the estimated location parameters, we estimate the GG model with different orders $(p, q)$ of the autoregressive and moving average terms in order to select the optimal order via AIC. We use the WL estimator, as discussed Section 2, and set the maximum orders of $p$ and $q$ as $p_{\max} = q_{\max} = 6$. Table 6 shows the selected orders $(p, q)$, and corresponding estimates of the long memory parameters. AIC selected relatively small $p$ and $q$ values for all series. The estimates of $d_1$ are significant at 5% percent level in all series. For ΔR3Y (US), $d_2$ is significant, while $d_3$ is insignificant. We observe the changes in the significance of the long memory parameters, $d_l$, caused by: (1) the efficiency gained by the parametric WL estimation as compared with the semiparametric GPH estimation, and (2) the effects of the ARMA parameters. There may exist some relationship between ΔR3Y (AU) and ΔR10Y (US), as these two have the same location parameter, $\omega_1$, with similar estimates of the long memory parameter, $d_1$.

Tables 5 and 6 show that IGG process is appropriate for modelling these six series. We need not consider the (general) fractional cointegration, as considered in Ramachandran and Beaumont (2001) and Smallwood and Norrbin (2008), as our data show the differences in the general long memory patterns for the six series.

## 4.2 Empirical Results for Cointegration

We examine the cointegrating relationship by the modified Johansen tests, and check the model adequacy by the forecasts, and extract the common generalized long memory in the system.

Table 7 shows the results for the modified test for cointegration, and also gives the asymptotic critical values. The trace and maximal eigenvalue tests favour the existence of two co-integrating vectors. Table 8 presents the eigenvalues, $\lambda_i$, and corresponding eigenvectors. As explained in Section 3, we obtain the ML estimates of $\alpha$ and $\gamma_\perp$ as $\hat{\alpha} = [\boldsymbol{v}_1\ \boldsymbol{v}_2]$ and $\hat{\gamma}_\perp = [\boldsymbol{m}_3 \cdots \boldsymbol{m}_6]$.

In order to check the adequacy of the model, we obtain one-step-ahead forecasts for the last 12 observations, which were reserved for this purpose. We fix the sample size for the rolling window. Figure 6 shows the one-step-ahead forecasts with 95% forecasts intervals, accompanied by the outcomes. For the comparison of the outcomes, the seasonal effects, removed by the dummy variables, are recovered. Figure 6 indicates that the outcomes fall into the 95% forecasts intervals in all cases, supporting the adequacy of the model specification.

As discussed in the previous section, we can decompose $\boldsymbol{y}_t$ into the stationary and nonstationary components. The stationary component is based on $\alpha' \boldsymbol{y}_t = (\boldsymbol{v}_1' \boldsymbol{y}_t, \boldsymbol{v}_2' \boldsymbol{y}_t)'$, while the nonstationary component uses $\gamma_\perp' \boldsymbol{y}_t = (\boldsymbol{m}_3' \boldsymbol{y}_t, \boldsymbol{m}_4' \boldsymbol{y}_t, \boldsymbol{m}_5' \boldsymbol{y}_t, \boldsymbol{m}_6' \boldsymbol{y}_t)'$. Based on the preliminary results, $\alpha' \boldsymbol{y}_t$ and $\gamma_\perp' \Delta \boldsymbol{y}_t$ are stationary long memory processes. Figure 7 presents the stationary and nonstationary components for all series.

We consider several hypotheses for the cointegrating vector using the test statistic (14). The first type considers the cointegration of one country. Gonzalo and Granger (1995) and MacKinnon (1996) consider the relationship among short, medium, and long term interest rates of one country, and we examine the hypothesis. The second type is interest rate parity, which means that the difference in the interest rates of two countries is $I(0)$, examined by Czudaj (2015), Johansen (1997), and Jusellius (1995). The third type is concerned with the zero-interest rate policy of the U.S.A. from 2008 to 2015. Owing to the policy, the U.S.A. short term interest rate moves quite differently, and we examine the effect.

As above, we test four kinds of hypothesis regarding the cointegrating vector under $r = 1$, namely: (1) the cointegrating vector exists only on Australia; (2) only the U.S.A. produces the cointegration relationship; (3) interest rate parity holds for each maturity; and (4) the short term interest rate in the U.S.A. has no effect on the cointegrating vector, and the corresponding restrictions on parameters are given by:

$$G_1 = \begin{pmatrix} I_3 \\ O_{3\times3} \end{pmatrix}, \quad G_2 = \begin{pmatrix} O_{3\times3} \\ I_3 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad G_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Table 9 shows the results of the tests. The first three hypothesis were rejected at the 5% significance level, while the null hypothesis that the short term U.S.A. interest rate has no effect on the cointegrating vector was not rejected.

## 5 Conclusion

In this paper, we considered the error correction model for the vector general long memory process, especially a vector of IGG processes, which we can estimate and test using recently developed econometric methods. For such time series, we suggested an approach to modify the tests for unit roots and cointegration. The Monte Carlo results showed that the sizes and powers of the modified tests are satisfactory, while the conventional ADF and Johansen tests suffer from size distortion under the IGG process. Using the short, medium, and long term interest rates in the Australia and U.S.A., the dataset favors the IGG process rather than the ARFIMA and/or GG processes examined in the literature. We also found there exists two cointegrating vectors, and that the short term interest rate in the U.S.A. has no effect on cointegration. Although the U.S. zero interest rate policy affected the short term interest rate, the medium and long term rates and Australian interest rates produced the cointegration relationship.

The results given in the paper can be extended by re-examining the cointegration relationships for a variety of macroeconomic variables. This topic is left for future research.

# References


Backus, D. and S. Zin (1993), "Long Memory Ination Uncertainty: Evidence from the Term Structure of Interest Rates", *Journal of Money, Credit and Banking*, **25**, 681–700.

Baillie, R.T. (1996), "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics*, **73**, 5-59.

Bordignon, S., M. Caporin, and F. Lisi (2007), "Generalised Long-Memory GARCH Models for Intra-Daily Volatility", *Computational Statistics & Data Analysis*, **51**, 5900–5912.

Caporale, G.M. and L.A. Gil-Alana (2011), "Multi-Factor Gegenbauer Processes and European Ination Rates", *Journal of Economic Integration*, **26**, 386–409.

Chan, K.S. and H. Tsai (2012), "Inference of Seasonal Long-Memory Aggregate Time Series", *Bernoulli*, **18**, 1448–1464.

Cheung, Y.-W. and K.S. Lai (1993), "Finite-Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration", *Oxford Bulletin of Economics and Statistics*, **55**, 313–328.

Cheung, Y.-W. and K.S. Lai (1995), "Lag Order and Critical Values of the Augmented DickeyFuller Test" *Journal of Business & Economic Statistics*, **13**, 277–280,

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

Czudaj, R. and J. Pruser (2015), "International Parity Relationships between Germany and the USA Revisited: Evidence from the Post-DM Period", *Applied Economics*, **47**, 2745–2767.

Dolado, J.J., T. Jenkinson, and S. Sosvilla-Rivero (1990), "Cointegration and Unit Roots", *Journal of Economic Surveys*, **4**, 249–273.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica*, **55**, 251–276.

Ferrara, L. and D. Guégan (2001), "Forecasting with $k$-Factor Gegenbauer Processes: Theory and Applications", *Journal of Forecasting*, **20**, 581–601.

Geweke, J. and S. Porter-Hudak (1983), "The Estimation and Application of Long-Memory Time Series Models", *Journal of Time Series Analysis*, **4**, 87–104.

Gonzalo, J. and C.W.J. Granger (1995), "Estimation of Common Long-Memory Components in Cointegrated System", *Journal of Business & Economic Statistics*, **13**, 27–35.

Granger, C.W.J. (1986), "Developments in the Study of Cointegrated Economic Variables," *Oxford Bulletin of Economics and Statistics*, **48**, 213-228.

Gray, H.L., N. Zhang, and W.A. Woodward (1989), "On Generalized Fractional Processes", *Journal of Time Series Analysis*, **10**, 233–257.

Gregory, A.W., J.M. Nason, and D.G. Watt (1996), "Testing for Structural Breaks in Cointegrated Relationships", *Journal of Econometrics*, **71**, 321–341.

Hidalgo, J. and P. Soulier (2004), "Estimation of the location and exponent of the spectral singularity of a long memory process", *Journal of Time Series Analysis*, **25**, 55–81.

Hwang, E., and D.W. Shin (2013), "A CUSUM test for a long memory heterogeneous autoregressive model", *Economics Letters*, **121**, 379–383.

Hwang, E., and D.W. Shin (2014), “Infinite-order long memory heterogeneous autoregressive models”, *Computational Statistics & Data Analysis*, **76**, 339-358.

Johansen, S. (1988), “Statistical Analysis of Cointegration Vectors”, *Journal of Economic Dynamics and Control*, **12**, 231–254.

Johansen, S. (1991), “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models”, *Econometrica*, **59**, 1551–1580.

Johansen, S. (1997), “Likelihood Analysis of the I(2) Model”, *Scandinavian Journal of Statistics*, **24**, 433–462.

Juselius, K. (1995), “Do Purchasing Power Parity and Uncovered Interest Rate Parity Hold in the Long Run? An Example of Likelihood Inference in A Multivariate Time-Series Model”, *Journal of Econometrics*, **69**, 211–240.

Lee, O. (2014), “The functional central limit theorem and structural change test for the HAR($\infty$) model”, *Economics Letters*, **124**, 370–373.

MacKinnon, J.G. (1996), “Numerical Distribution Functions for Unit Root and Cointegration Tests”, *Journal of Applied Econometrics*, **11**, 601–618.

McElroy, T. S. and S. H. Holan (2012), “On the Computation of Autocovariances for Generalized Gegenbauer Processes”, *Statistica Sinica*, **22**, 1661–1687.

McElroy, T.S. and S.H. Holan (2016), “Computation of the Autocovariances for Time Series with Multiple Long-Range Persistencies”, *Computational Statistics & Data Analysis*, **101**, 44–56.

Osterwald-Lenum, M. (1992), “A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics”, *Oxford Bulletin of Economics and Statistics*, **54**, 461–472.

Pantula, S.G., G. Gonzalez-Farias, and W.A. Fuller (1994), “A Comparison of Unit-Root Test Criteria”, *Journal of Business & Economic Statistics*, **12**, 449–459

Peiris, S. and M. Asai (2016), “Generalized Fractional Processes with Long Memory and Time Dependent Volatility Revisited”, *Econometrics*, **4(3)**, 1-21.

Porter-Hudak, S. (1990), “An Application of the Seasonal Fractionally Differenced Model to the Monetary Aggregates”, *Journal of the American Statistical Association*, **85**, 338–344.

Ramachandran, R. and P.M. Beaumont (2001), “Robust Estimation of GARMA Model Parameters with an Application to Cointegration among Interest Rates of Industrialized Countries”, *Computational Economics*, **17**, 179–201.

Said, E.S. and D.A. Dickey (1984), “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order”, *Biometrika*, **71**, 599–607.

Smallwood, A. and S.C. Norrbin (2008), “An Encompassing Test of Real Interest Rate Equalization”, *Review of International Economics*, **16**, 114–126.

Tsai, H., H. Rachinger, and E.M.H. Lin (2015), “Inference of Seasonal Long-Memory Time Series with Measurement Error”, *Scandinavian Journal of Statistics*, **42**, 137–154.

Woodward, W.A., Q.C. Cheng, and H.L. Gray (1998), “A $k$-Factor GARMA Long Memory Model”, *Journal of Time Series Analysis*, **19**, 485–504.

Wu, H. and S. Peiris (2018), “An Introduction to Vector Gegenbauer Processes with Long Memory”, to appear in *Stat*.

# Appendix A

## A.1 Identification and Estimation of Gegenbauer Frequencies

We discuss the identification of $k$ and give a short explanation of semiparametric estimation of the location frequency parameters. First, we explain the semiparametric technique of Hidalgo and Soulier (2004) for estimating the location frequency parameters of $(\omega_1, \ldots, \omega_k)$ for the univariate GG process. We assume that $k$ is known until we discuss the identification of parameters. Let $I_T(\boldsymbol{z}, \lambda)$ be the periodogram defined by:

$$I_T(\boldsymbol{z}, \lambda) = (2\pi T)^{-1} \left| \sum_{t=1}^{T} \boldsymbol{z}_t e^{it\lambda} \right|^2, \quad -\pi < \lambda \leq \pi,$$

for $z_1, \ldots, z_T$. Define $\tilde{n} = \lfloor (T-1)/2 \rfloor$ and let $\lambda_j = 2j\pi/T$ $(-\tilde{n} \leq j \leq \tilde{n})$ be the Fourier frequencies.

For purposes of introducing the approach of Hidalgo and Soulier (2004), we consider a simple case of a univariate process which produces $I_T(\lambda)$, under the assumptions $d = 0$, $\omega_1 \neq 0$, $\omega_2 \neq 0$, $d_1 \geq d_2$, and $k = 2$. Then we can estimate $\omega_1$ and $\omega_2$ consistently as:

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

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

$$\frac{1}{n} + \frac{n}{T} \to 0 \text{ as } T \to \infty,$$

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

Hidalgo and Soulier (2004) modified the GPH estimator of Geweke and Porter-Hudak (1983), which was originally suggested to estimate the long memory parameter, $d$, using a log-periodogram regression, in order to estimate $d_l$ at the Gegenbauer frequency $\omega_l$. In order to identify the number of location frequencies, $k$, we follow the approach of Hidalgo and Soulier (2004), based on their modified GPH estimator for $d_1, \ldots, d_k$, which is defined by:

$$\hat{d}_l = \sum_{\substack{1 \leq |j| \leq n \\ 0 < \hat{\omega}_l + \lambda_j \leq \pi}} \xi_k \ln \left\{ I_T(\hat{\omega}_l + \lambda_j) \right\}, \tag{17}$$

where $\xi_k = s_n^{-2}(\zeta(\lambda_j) - \bar{\zeta}_n)$, $\zeta(\lambda) = -\ln(|1 - e^{i\lambda}|)$, $\bar{\zeta}_n = n^{-1}\sum_{j=1}^{n}\zeta(\lambda_j)$, and $s_n^2 = \sum_{j=1}^{n}(\zeta(\lambda_j) - \bar{\zeta}_n)^2$. Hidalgo and Soulier (2004) show that $n^{1/2}(\hat{d}_l - d_l)$ converges weakly to $N(0, \pi^2/12)$, under the assumption of a Gaussian process. For the case $\omega_l = 0, \pi$, Hidalgo and Soulier (2004) also show that the limiting distribution is $N(0, \pi^2/6)$.

The procedure of Hidalgo and Soulier (2004) consists of the following steps: (i) Find the largest periodogram ordinate; (ii) if the corresponding estimate of $d_l$ is significant, add the respective Gegenbauer filter to the model, otherwise terminate the procedure; (iii) exclude the neighborhood of the last pole from the periodogram, and repeat the procedure from (i) onward. In the empirical analysis, we apply the method just described for identifying $k$ and estimating $(\omega_1, \ldots, \omega_k)$ for the GG model.

### A.2 Generation of GG Processes

We generate the GG process by the modified Durbin's algorithm based on the theoretical covariance function for the whole sample, applying the discussion of Doonik and Oomes (2003). In the following, we first explain the calculation of the coefficients of the MA($\infty$) representation of the general Gegenbauer process in equation (2) in order to show the calculation of the autocovariance functions.

Even for the simple Gegenbauer process with ARMA parameters, it is not easy to obtain explicit formulas for the coefficients for the MA($\infty$) representation, and the autocovariances, that are valid for all lags. Recently, McElroy and Holan (2012, 2016) developed a computationally efficient method for calculating these values. The spectral density of the general Gegenbauer process is given by (5). For convenience, we define $\kappa(z)$ so that $g(\lambda) = |\kappa(e^{-i\lambda})|^2$. Then, $\kappa(z)$ takes the form $\kappa(z) = \prod_l (1 - \zeta_l z)^{p_l}$ for (possibly complex) reciprocal roots, $\zeta_l$, of the moving average and autoregressive polynomials, where $p_l$ is one if $l$ corresponds to a moving average root, and minus one if $l$ corresponds to an autoregressive root. We set $\alpha = \max\{d_l\}$ for notational convenience.

Define:

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

$$\beta_j = \frac{2}{j}\left\{d_0 + 2\sum_{l=1}^{k} d_l \cos(\omega_l j)\right\} + g_j,$$

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

McElroy and Holan (2012) showed that the MA($\infty$) representation of (2) is given by:

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

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

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

where

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

and $F(a, b; c; z)$ is the hypergeometric function evaluated at $z$. In the above, $K_l$ is a component depending on the structure of (5), and our GG processes have

$$K_l = 2\cos(lw_1)\left[2\left\{1-\cos(2\omega_1)\right\}\right]^{-d_1}\left[2\left\{1-\cos(\omega_1-\omega_2)\right\}\right]^{-d_2}$$
$$\times \left[2\left\{1-\cos(\omega_1+\omega_2)\right\}\right]^{-d_2}\left[1+\phi^2-2\phi\cos(\omega_1)\right]^{-1},$$

with $k = 2$, $d_0 = 0$, $d_1 \geq d_2$. Note that $\gamma_{-l} = \gamma_l$. McElroy and Holan (2012) recommend using the cutoff value $J \geq 2,000$, and we set $J = 20,000$ in this paper.

Table 1: Rejection Frequencies of the Modified Unit Root Tests

| | | Modified ADF | | | | ADF with $p = 6$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Level | | First Diff. | | Level | | First Diff. | |
| DGP | $T$ | 5% | 10% | 5% | 10% | 5% | 10% | 5% | 10% |
| DGP1 | 250 | 0.0910 | 0.1480 | 1.0000 | 1.0000 | 0.1380 | 0.2300 | 1.0000 | 1.0000 |
| | 500 | 0.0515 | 0.1010 | 1.0000 | 1.0000 | 0.1565 | 0.2265 | 1.0000 | 1.0000 |
| DGP2 | 250 | 0.0845 | 0.1430 | 0.9995 | 1.0000 | 0.1335 | 0.2155 | 1.0000 | 1.0000 |
| | 500 | 0.0460 | 0.0940 | 1.0000 | 1.0000 | 0.1325 | 0.2160 | 1.0000 | 1.0000 |
| DGP3 | 250 | 0.0670 | 0.1025 | 0.9455 | 0.9725 | 0.1470 | 0.2355 | 1.0000 | 1.0000 |
| | 500 | 0.0460 | 0.0805 | 1.0000 | 1.0000 | 0.1580 | 0.2415 | 1.0000 | 1.0000 |

Note: DGPs are the IGG process with parameters (15). We use nominal sizes of 5% and 10 %. The critical value for the test is given by (12).

Table 2: Rejection Frequencies of the Modified Tests for Cointegration

| | | Modified Johansen's Test | | | | Johansen's Test with $p = 6$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Trace | | $\lambda_{max}$ | | Trace | | $\lambda_{max}$ | |
| DGP | $T$ | $l = 1$ | $l = 0$ | $l = 1$ | $l = 0$ | $l = 1$ | $l = 0$ | $l = 1$ | $l = 0$ |
| DGP1nc | 250 | 0.0060 | 0.0885 | 0.0060 | 0.1100 | 0.0000 | 0.0100 | 0.0000 | 0.0180 |
| | 500 | 0.0005 | 0.0435 | 0.0005 | 0.0605 | 0.0000 | 0.0065 | 0.0000 | 0.0155 |
| DGP2nc | 250 | 0.0015 | 0.0790 | 0.0015 | 0.0995 | 0.0000 | 0.0075 | 0.0000 | 0.0170 |
| | 500 | 0.0005 | 0.0255 | 0.0005 | 0.0435 | 0.0000 | 0.0050 | 0.0000 | 0.0090 |
| DGP3nc | 250 | 0.0015 | 0.0760 | 0.0015 | 0.1005 | 0.0005 | 0.0100 | 0.0005 | 0.0160 |
| | 500 | 0.0000 | 0.0305 | 0.0000 | 0.0485 | 0.0005 | 0.0060 | 0.0005 | 0.0115 |
| DGP1co | 250 | 0.0090 | 0.8855 | 0.0090 | 0.9195 | 0.0060 | 0.9770 | 0.0060 | 0.9895 |
| | 500 | 0.0035 | 0.9995 | 0.0035 | 1.0000 | 0.0085 | 1.0000 | 0.0085 | 1.0000 |
| DGP2co | 250 | 0.0075 | 0.8995 | 0.0075 | 0.9315 | 0.0060 | 0.9755 | 0.0060 | 0.9915 |
| | 500 | 0.0070 | 1.0000 | 0.0070 | 1.0000 | 0.0060 | 1.0000 | 0.0060 | 1.0000 |
| DGP3co | 250 | 0.0125 | 0.9000 | 0.0120 | 0.9285 | 0.0050 | 0.9715 | 0.0050 | 0.9885 |
| | 500 | 0.0050 | 0.9985 | 0.0050 | 0.9985 | 0.0020 | 1.0000 | 0.0020 | 1.0000 |
| | CV | 8.18 | 17.95 | 8.18 | 14.90 | 9.24 | 19.96 | 9.24 | 15.67 |

Note: For $l = 1, 2, 3$, DGP$l$co (DGP$l$nc) indicates that $y_{2t}$ is generated by the IGG process, and that there is one (no) cointegrating vector for $(y_{1t}, y_{2t})$. The null hypothesis of the trace test is the number of cointegrating vectors, $r$, is at most $l$, that is, $r \leq l$. The $\lambda_{\max}$ statistic is used for testing the null hypothesis $r = l$ against $r = l + 1$. We use nominal size of 5%, and the critical values (CV) are given in the last row.

Table 3: Descriptive Statistics of Interest Rates

| Data | Average | Std. Dev. | Skewness | Kurtosis |
|---|---|---|---|---|
| R3M (AU) | 4.8473 | 1.6396 | −0.2164 | 2.4348 |
| R3Y (AU) | 5.0207 | 1.9070 | 0.0541 | 2.7498 |
| R10Y (AU) | 5.5608 | 1.8801 | 0.3202 | 2.9763 |
| R3M (US) | 2.4452 | 2.1085 | 0.2012 | 1.4266 |
| R3Y (US) | 3.3086 | 2.1133 | 0.1263 | 1.6091 |
| R10Y (US) | 4.3195 | 1.6608 | 0.0479 | 1.9234 |

Table 4: Unit Root Tests

| Data | Level Series | | | Differenced Series | | |
|---|---|---|---|---|---|---|
| | $\tau_{aic}$ | $\tau_{aic2}$ | $\tau_{har}$ | $\tau_{aic}$ | $\tau_{aic2}$ | $\tau_{har}$ |
| R3M (AU) | −2.3812 (4) | −2.2494 (6) | −0.3254 | −5.0125*(6) | −5.0125*(6) | −4.3863* |
| R3Y (AU) | −2.2491 (3) | −1.9525 (5) | −1.0113 | −6.6059*(6) | −6.6059*(6) | −6.0596* |
| R10Y (AU) | −2.4047 (3) | −2.2830 (5) | −1.7548 | −12.375*(1) | −7.8805*(3) | −7.5677* |
| R3M (US) | −1.7078 (4) | −1.7114 (6) | −0.8629 | −5.2338*(3) | −4.9020*(5) | −4.5267* |
| R3Y (US) | −1.4875 (3) | −1.6292 (5) | −1.1747 | −12.044*(1) | −7.5180*(3) | −6.4625* |
| R10Y (US) | −1.6396 (6) | −1.6396 (6) | −1.1720 | −8.3009*(3) | −7.6667*(5) | −8.3408* |

Note: The entries show the values of $\tau$ statistic for (10) and (11). The parentheses give the selected value of $p$ in (10) by AIC ($p_{aic}$) and $p_{aic2} = \min(p_{aic} + 2, 6)$. '*' indicates the significance at the 5% level. The critical value for the test is given by (12).

Table 5: Semiparametric Estimates of the Location Parameters

| | ΔR3M (AU) | | | ΔR3Y (AU) | | | ΔR10Y (AU) | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $\hat{\omega}_l$ | Months | $P$-value | $\hat{\omega}_l$ | Months | $P$-value | $\hat{\omega}_l$ | Months | $P$-value |
| 1 | 0.0408 | 49.000 | 0.0000* | 0.2286 | 8.7500 | 0.0551 | 0.2286 | 8.7500 | 0.0735 |
| 2 | 0.2939 | 6.9056 | 0.1839 | — | — | — | — | — | — |
| | ΔR3M (US) | | | ΔR3Y (US) | | | ΔR10Y (US) | | |
| $k$ | $\hat{\omega}_l$ | Months | $P$-value | $\hat{\omega}_l$ | Months | $P$-value | $\hat{\omega}_l$ | Months | $P$-value |
| 1 | 0.0245 | 81.667 | 0.0000* | 0.0327 | 61.250 | 0.0000* | 0.2286 | 8.7500 | 0.4585 |
| 2 | 0.1959 | 10.208 | 0.0583 | 0.1796 | 11.1364 | 0.0002* | — | — | — |
| 3 | — | — | — | 0.6776 | 2.9512 | 0.0000* | — | — | — |
| 4 | — | — | — | 0.4245 | 4.7115 | 0.3613 | — | — | — |

Note: The estimates of $\omega_l$ are reported with the unit of $\pi$. 'Months' indicates the period corresponding to $\hat{\omega}_l$. '$P$-value' shows the $P$-value for the modified GPH estimates of $d_l$, and '*' indicates significance at the 5% level.

Table 6: Whittle Likelihood Estimates of Long Memory parameters for GG Models

| | WL Estimates | | | Selected Order | |
|---|---|---|---|---|---|
| Data | $d_1$ | $d_2$ | $d_3$ | $p$ | $q$ |
| $\Delta$R3M (AU) | 0.1763* | — | — | 3 | 5 |
| $\Delta$R3Y (AU) | 0.1771* | — | — | 4 | 1 |
| $\Delta$R10Y (AU) | 0.2986* | — | — | 5 | 5 |
| $\Delta$R3M (US) | 0.2474* | — | — | 2 | 0 |
| $\Delta$R3Y (US) | 0.1286* | 0.1492* | 0.0240 | 3 | 3 |
| $\Delta$R10Y (US) | 0.1737* | — | — | 3 | 5 |

Note: We selected order $(p, q)$ via AIC with $p_{\max} = q_{\max} = 6$. '*' indicates the significance at the 5% level. Based on the estimates of Table 5, we set $\omega_1$ as $2\pi \times \{1/49, 1/9, 1/9, 1/82, 1/61, 1/9\}$, respectively. For $\Delta$R3Y (US), we set $\omega_2 = \pi/6$ and $\omega_3 = 2\pi/3$.

Table 7: Test Statistics for Cointegration

| | Trace Test | | $\lambda_{\max}$ Test | |
|---|---|---|---|---|
| $l$ | $Q_{\text{trace}}$ | $CV_{\infty}^{\text{trace}}$ | $Q_{\max}$ | $CV_{\infty}^{\max}$ |
| 5 | 3.9751 | 8.18 | 3.9751 | 8.18 |
| 4 | 9.0101 | 17.95 | 5.0350 | 14.90 |
| 3 | 18.4252 | 31.52 | 9.4151 | 21.07 |
| 2 | 39.8663 | 48.28 | 21.4411 | 27.14 |
| 1 | 73.5341$*$ | 70.60 | 33.6678$*$ | 33.32 |
| 0 | 159.0148$*$ | 95.18 | 85.4807$*$ | 39.43 |

Note: The null hypothesis of the trace test is number of cointegrating vectors, $r$, is at most $l$, that is, $r \leq l$. The $\lambda_{\max}$ statistic is used for testing the null hypothesis $r = l$ against $r = l + 1$. $CV^{\text{trace}}$ and $CV^{\max}$ are the asymptotic critical values for the trace and maximal eigenvalue tests, respectively. '*' indicates significance at the 5% level.

Table 8: Estimates of the Cointegration Structure

| Eigenvalue | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\lambda_4$ | $\lambda_5$ | $\lambda_6$ |
|---|---|---|---|---|---|---|
| | 0.2945 | 0.1284 | 0.0838 | 0.0377 | 0.0203 | 0.0161 |
| Eigenvector | $\boldsymbol{v}_1$ | $\boldsymbol{v}_2$ | $\boldsymbol{v}_3$ | $\boldsymbol{v}_4$ | $\boldsymbol{v}_5$ | $\boldsymbol{v}_6$ |
| R3M (AU) | 3.2783 | 6.9649 | −2.7028 | −0.4990 | 0.2767 | −0.5146 |
| R3Y (AU) | 1.4275 | −19.2370 | 5.8008 | 0.2237 | −1.6248 | −3.3289 |
| R10Y (AU) | −7.4865 | 15.4970 | −2.0736 | −1.1891 | 0.3324 | 1.2616 |
| R3M (US) | 0.7514 | −0.5414 | 0.8728 | 0.9320 | −6.4749 | 7.7676 |
| R3Y (US) | −3.5770 | 7.7979 | −3.5436 | −3.8896 | 7.1857 | −9.7112 |
| R10Y (US) | 6.1519 | −10.9961 | 1.0071 | 4.7146 | −0.6964 | 5.7297 |
| Eigenvector | $\boldsymbol{m}_1$ | $\boldsymbol{m}_2$ | $\boldsymbol{m}_3$ | $\boldsymbol{m}_4$ | $\boldsymbol{m}_5$ | $\boldsymbol{m}_6$ |
| R3M (AU) | −3.9282 | −8.4948 | −2.3993 | 0.7298 | 0.7372 | 1.0877 |
| R3Y (AU) | 0.6403 | 5.7389 | 10.4030 | −3.8397 | 3.8601 | 4.5805 |
| R10Y (AU) | 9.6147 | −4.8819 | −10.4533 | 6.4945 | −3.9529 | −4.3240 |
| R3M (US) | −2.7804 | −1.2362 | 3.9238 | 7.4595 | −0.1367 | −4.9479 |
| R3Y (US) | 3.8550 | 0.0596 | −7.0936 | −8.0315 | 5.6673 | −5.0566 |
| R10Y (US) | −11.1614 | 2.7439 | 5.0395 | 4.4722 | −1.1098 | 7.8298 |

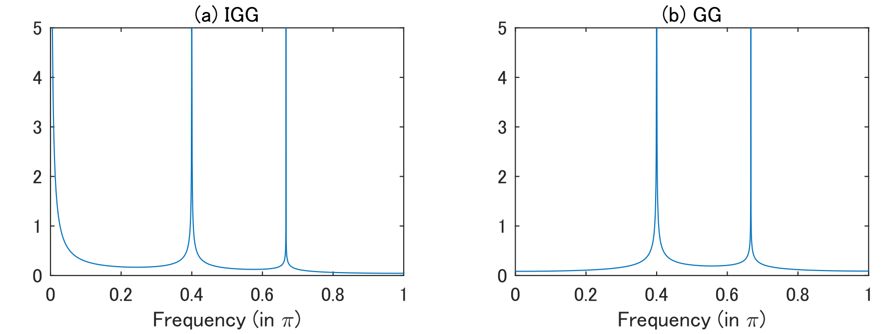
Note: $\hat{\alpha} = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_r]$ and $\hat{\gamma}_\perp = [\boldsymbol{m}_{7-r} \cdots \boldsymbol{m}_6]$, where $r$ is the number of cointegrating vectors.

Table 9: Tests for Parameter Restrictions on the Cointegrating Vector

| Restriction | $G_1$ | $G_2$ | $G_3$ | $G_4$ |
|---|---|---|---|---|
| Test Stat. | 4.04433 | 85.0521 | 47.5583 | 0.6015 |
| $P$-value | [0.0000] | [0.0000] | [0.0000] | [0.7403] |

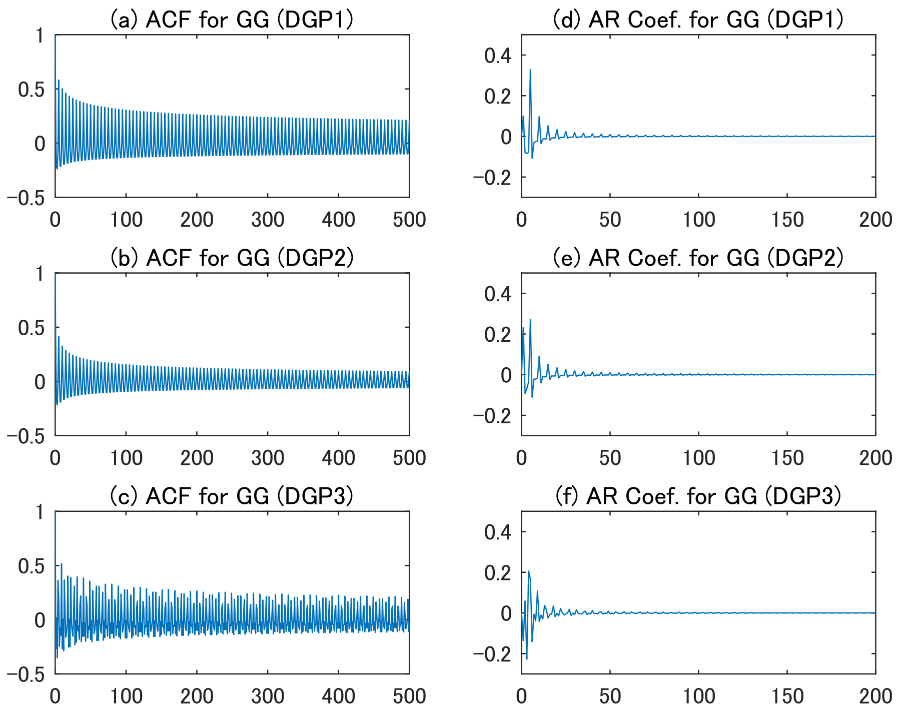
Note: $P$-values are given in brackets.

Figure 1: Power Spectrum of the Generalized Gegenbauer Processes




Note: The pseudo power spectral density is used for the IGG model.

Figure 2: Autocorrelation Function and AR Coefficients of the Generalized Gegenbauer Processes




Note: We consider the GG process for the setting of parameters of each DGP, that is, the differenced series ($\Delta y_t$). Parameters of the GG process are given in (15).

Figure 3: Interest Rates for Australia and U.S.A.

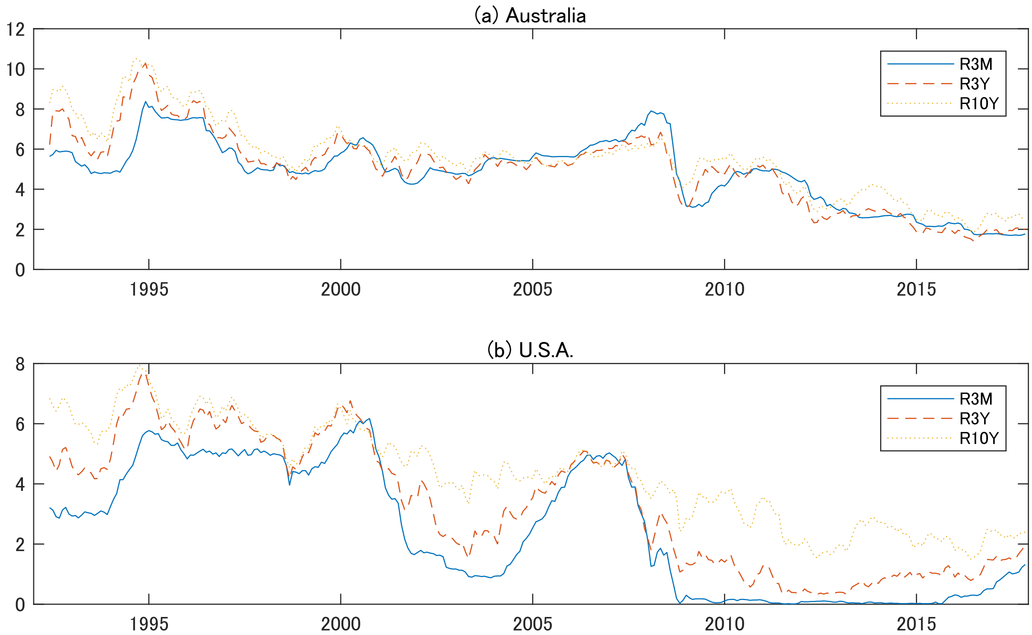
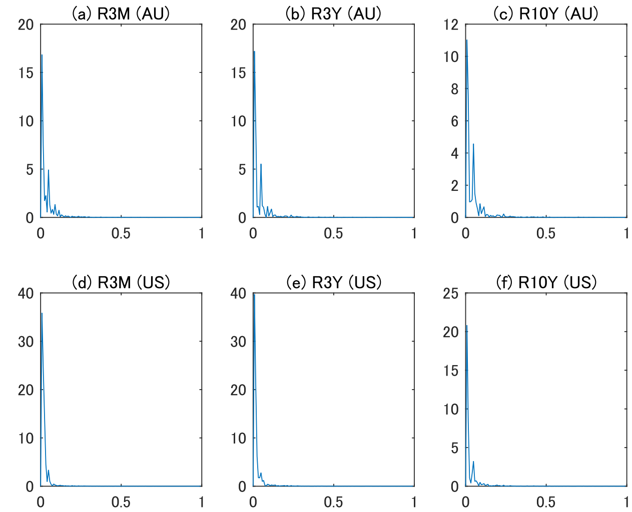
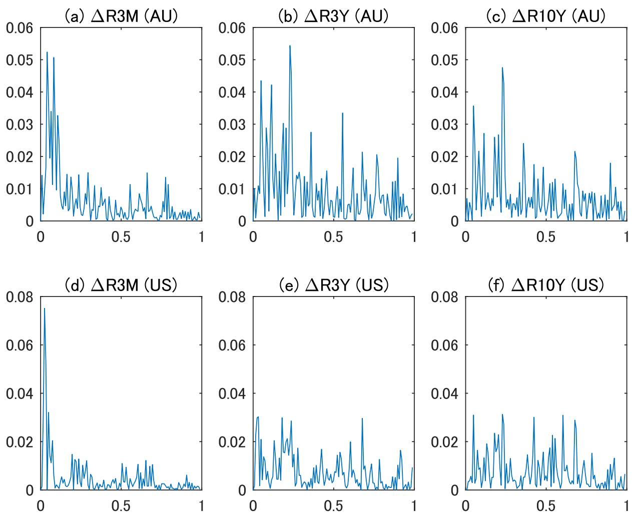


Figure 4: Sample Spectral Density for the Levels Series



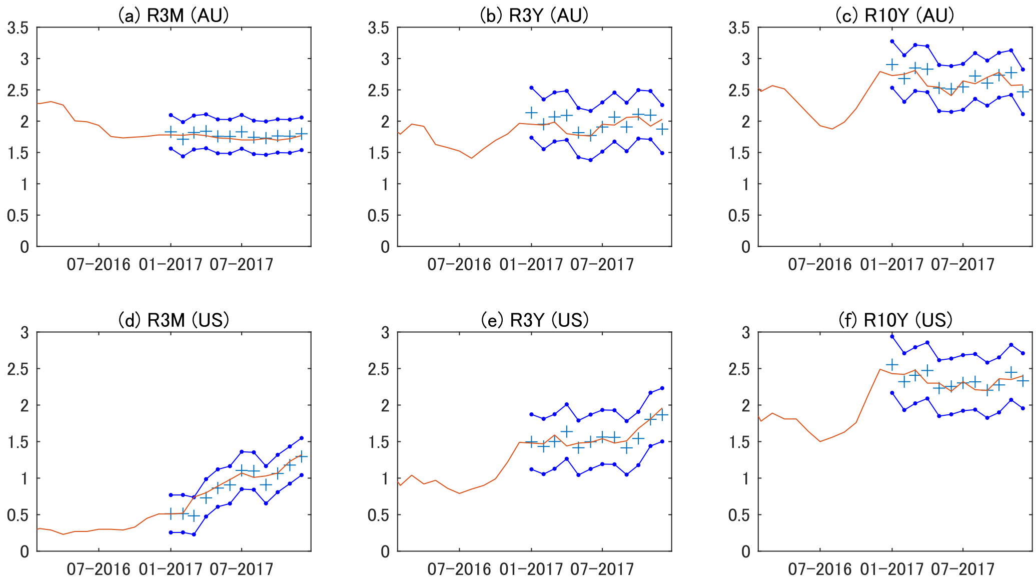

Note: $x$-axis displays the frequency in $\pi$.

Figure 5: Sample Spectral Density for the First Differenced Series



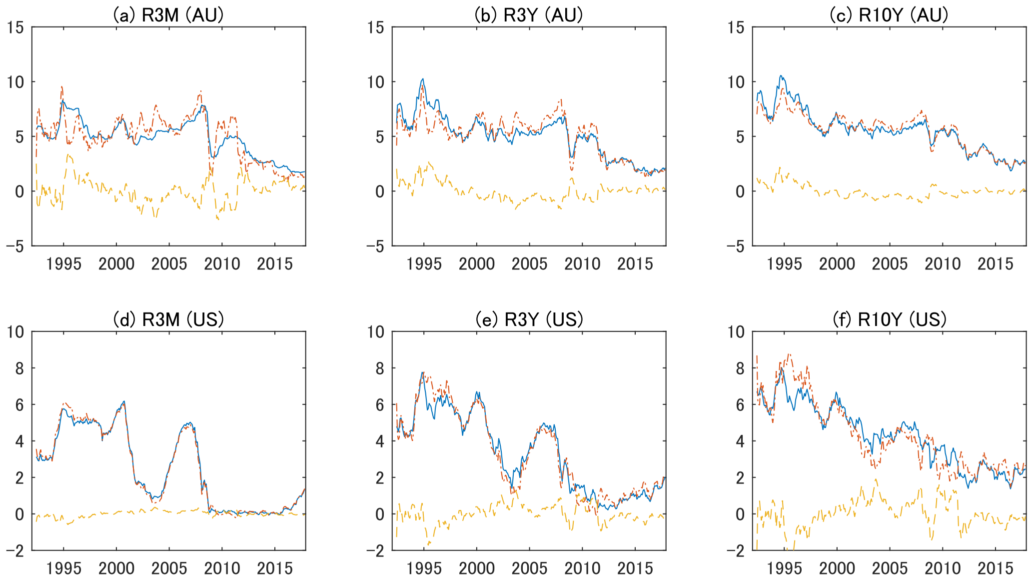

Note: $x$-axis displays the frequency in $\pi$.

Figure 6: One-Step Ahead Forecasts and Outcomes




Note: The solid line shows the outcomes, while '+' indicates the forecast. The lines with dots show the 95% forecasts intervals.

Figure 7: Stationary and Nonstationary Components of the Interest Rates




Note: The solid lines represent the outcomes, the dashed line indicates the stationary components, and the dash with dot line is the nonstationary components.