

**2008/47**



An easy test for two stationary long processes being  
uncorrelated via AR approximations

Shin-Huei Wang and Cheng Hsiao

**CORE**

Voie du Roman Pays 34

B-1348 Louvain-la-Neuve, Belgium.

Tel (32 10) 47 43 04

Fax (32 10) 47 43 01

E-mail: [corestat-library@uclouvain.be](mailto:corestat-library@uclouvain.be)

<http://www.uclouvain.be/en-44508.html>

CORE DISCUSSION PAPER  
2008/47

**An easy test for two stationary long processes  
being uncorrelated via AR approximations**

Shin-Huei WANG<sup>1</sup> and Cheng HSIAO<sup>2</sup>

August 2008

**Abstract**

This paper proposes an easy test for two stationary autoregressive fractionally integrated moving average (ARFIMA) processes being uncorrelated via AR approximations. We prove that an ARFIMA process can be approximated well by an autoregressive (AR) model and establish the theoretical foundation of Haugh's (1976) statistics to test two ARFIMA processes being uncorrelated. Using AIC or Mallow's  $C_p$  criterion as a guide, we demonstrate through Monte Carlo studies that a lower order AR( $k$ ) model is sufficient to prewhiten an ARFIMA process and the Haugh test statistics perform very well in finite sample. We illustrate the methodology by investigating the independence between the volatility of two daily nominal dollar exchange rates-Euro and Japanese Yen and find that there exists "strongly simultaneous correlation" between the volatilities of Euro and Yen within 25 days.

**Keywords:** forecasting, long memory process, structural break.

**JEL Classification:** C22, C53

---

<sup>1</sup> CORE, Université catholique de Louvain, Belgium. E-mail: shin-huei.wang@uclouvain.be

<sup>2</sup> Department of Economics, University of Southern California, USA

Wang was supported by the contract "Projet d'Actions de Recherche Concertées" 07/12-002 of the Communauté française de Belgique", granted by the "Académie universitaire Louvain" and Chiang Ching Kuo Foundation for International Scholarly Exchange.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.



## 1. Introduction

This paper proposes an easy to implement test for two long memory processes being uncorrelated. Recently, we have witnessed increasingly large number of studies using the autoregressive fractionally integrated moving average process of order  $p, d, q$ , denoted as ARFIMA  $(p, d, q)$  or  $I(d)$  process, when the integrated process of order  $d$ , or the differencing parameter  $d$  is a fractional number (e.g., Robinson, 1991; Ding and Granger, 1996; Baillie, 1996; Ray and Tsay, 2000; Bollerslev and Wright, 2000). One of the advantages of using the ARFIMA process is that it provides a parsimonious way to describe data series with long-range dependence. However, the use of the ARFIMA process also incurs some problems in some popularly used test statistics. Tsay and Chung (2000) proved that two stationary  $I(d_1)$  and  $I(d_2)$  processes can yield spurious regression so long as  $d_1 + d_2 > 0.5$ ,  $d_1 \in (0, 0.5)$ , and  $d_2 \in (0, 0.5)$ . On the other hand, Tsay (Theorem 4, 2000) showed that the insignificant testing problem first considered by Robinson (1993) could also arise if  $d_1 d_2 < 0$ , because the  $t$  ratio will converge to zero, the actual size of using  $t$ -test for the null hypothesis of no relationship between the two unrelated  $I(d)$  processes will be significantly below the nominal size. Thus, the usual  $t$  statistic could be of no use in empirical applications when the DGP are  $I(d)$  processes.

One possible way to tackle the preceding testing problems is to first estimate the parameters of the ARFIMA model for  $Y_{1t}$  and  $Y_{2t}$ , respectively. Under the assumption that the differencing parameters  $d_1$  and  $d_2$  are known, the null hypothesis that  $Y_{1t}$  and  $Y_{2t}$  are independent of each other can be tested with accuracy by Monte Carlo simulation (Tsay and Chung, 2000). Nevertheless, it is well known that the estimation of the differencing parameter of the ARFIMA process may not be very accurate in finite samples when  $d$  is close to 0.5. Therefore, the simulation method may not be a good way to meet our purpose.

Our test follows the suggestions of Haugh (1976) and Hong (1996) to construct the residual cross-correlation functions from prewhitened time series. We shall first justify the use of Haugh's statistics when the DGP are the ARFIMA  $(p, d, q)$  processes. We then show that  $y_t$  can be approximated well by an AR( $k$ ) model. This finding itself contains some interesting empirical implications. First, the AR approximation is easy to implement because we do not need to estimate the differencing parameters of the data series. Second, through a simulation study, Ray (1993) showed that an AR approximation could be useful for long-range forecasting of a long-memory process. Her simulation results are explained by our theoretical analysis. Third, the good performance of the indirect estimation method for the  $I(d)$  process proposed by Martin and Wilkins (1999) and that of the efficient method of moments (EMM) employed by Gallant, Hsieh and Tauchen (1997) are easily understood, because they both employed an AR model as the auxiliary model. Fourth, the issues of spurious regression between two stationary long memory processes considered by Tsay and Chung (2000) will not arise.

We also consider the determination of the order of an AR( $k$ ) model fitted to a long memory process. We propose to use the AIC or the Mallows'  $C_p$  to select a finite order AR( $k$ ) model. Our Monte Carlo experiments show that an AR( $k$ ) approximation can yield the mean of the residual variance estimate very close to the true variance of  $e_t$  and the lag length needed to approximate the ARFIMA process is quite small. Our experiment clearly demonstrates that the good finite sample performance of Haugh's statistics when the DGP are the ARFIMA processes. The size distortion of the Haugh's statistics is well controlled for various combinations of  $d_1$  and  $d_2$  and the power performance is very promising. We also illustrate our methodology by applying our tests to the volatility between two daily nominal U.S. dollar exchange rate-Euro and Japanese Yen and find that there exist strong simultaneous volatilities interaction between them within one month.

This paper is organized as follows. Section 2 provides the theoretical justification of the test statistics, including the proof of approximating a long memory process by an AR ( $k$ ) model. In section 3, we use Monte Carlo studies to demonstrate the usefulness of AIC and Mallow's  $C_p$  criterion for selecting an AR ( $k$ ) model to long memory processes. Section 4 presents the empirical analysis of the correlation in volatility between Euro and Japanese Yen. The last section summarizes this paper. All proofs are in the Appendix.

## 2. THE MODEL AND THE MAIN RESULTS

Haugh (1976) proposed the following test statistics for two weakly dependent processes being uncorrelated :

$$S_M = T \sum_{l=-M}^M \hat{\rho}_{12}(l)^2 \quad \text{and} \quad S_M^* = T^2 \sum_{l=-M}^M (T - |l|)^{-1} \hat{\rho}_{12}(l)^2,$$

where

$$\hat{\rho}_{12}(l) = \frac{\sum_{t=k+1}^{T-l} \hat{e}_{t,k,1} \hat{e}_{t+l,k,2}}{\left( \sum_{t=k+1}^T \hat{e}_{t,k,1}^2 \right)^{1/2} \left( \sum_{t=k+1}^T \hat{e}_{t,k,2}^2 \right)^{1/2}},$$

and  $\hat{e}_{t,k,1}$  and  $\hat{e}_{t,k,2}$  are residuals from fitting  $Y_{1t}$  and  $Y_{2t}$  with an ARMA model, respectively. Haugh (1976) proved that  $S_M$  and  $S_M^*$  are both asymptotically distributed as  $\chi_{2M+1}^2$ . Hong (1996) suggested constructing the Haugh's statistics with residuals from an AR( $k$ ) filter and showed that Haugh's statistics perform very well.

The objective of this paper is to further extend the two test statistics,  $S_M$  and  $S_M^*$ , to the long memory case, i.e, the ARFIMA ( $p, d, q$ ) processes. We suppose the DGP,  $y_t$ , an ARFIMA ( $p, d, q$ ) process, satisfies the following Assumption 1 throughout this paper.

**Assumption 1.**  $y_t$  is generated as:

$$\phi(L)(1-L)^d y_t = \theta(L)e_t,$$

where (i)  $d \in (0, 0.5)$ ; (ii)  $\phi(L)$ , and  $\theta(L)$  are finite degree polynomials, and the zeroes of  $\phi(L)$ , and  $\theta(L)$  all lie outside the unit circle; (iii)  $\phi(L)$  and  $\theta(L)$  have no common zeroes; (iv)  $e_t$  is an independently and identically distributed process, with  $E(e_t) = 0$ ,  $E(e_t^2) = \sigma^2$ , and  $E(e_t^4) < \infty$ .

Hosking (1996) shows that a stationary and invertible ARFIMA process with  $d \neq 0$  has an autocovariance function that satisfies

$$\gamma_j \sim \frac{\sigma^2 f_y(0) \Gamma(1-2d)}{\Gamma(d) \Gamma(1-d)} j^{2d-1}, \quad \text{as } j \rightarrow \infty,$$

where  $\Gamma(\cdot)$  is the Gamma function, and

$$f_y(0) = \frac{(1 + \theta_1 + \dots + \theta_q)^2}{(1 - \phi_1 - \dots - \phi_p)^2}.$$

Assumption 1 guarantees that the conditions in Theorem 3 of Hosking (1996) hold, and allows us to represent an ARFIMA process  $y_t$  as:

$$y_t = \sum_{j=0}^{\infty} \psi_j e_{t-j}, \quad \text{where } \psi_j = O(j^{d-1}) \quad \text{as } j \rightarrow \infty,$$

or

$$y_t = \sum_{j=1}^{\infty} \beta_j y_{t-j} + e_t, \quad \text{where } \beta_j = O(j^{-d-1}) \quad \text{as } j \rightarrow \infty.$$

Accordingly, we resort to analyzing the asymptotic properties of Haugh's statistics when  $Y_{1t}$  and  $Y_{2t}$  are generated as

$$ARFIMA(p_1, d_1, q_1) : \quad \phi_1(L)(1-L)^{d_1} Y_{1t} = \theta_1(L)e_{1,t},$$

and

$$ARFIMA(p_2, d_2, q_2) : \quad \phi_2(L)(1-L)^{d_2} Y_{2t} = \theta_2(L)e_{2,t},$$

where  $d_1 \in (0, 0.5)$  and  $d_2 \in (0, 0.5)$ , respectively, and  $e_{1,t}$  and  $e_{2,t}$  are both zero mean white noise processes with variances  $\sigma_1^2$  and  $\sigma_2^2$ , respectively. We first show that an ARFIMA  $(p, d, q)$  process can be approximated by an AR model, then show that the residual correlations of two prewhitened ARFIMA processes are asymptotically Chi-square distributed under the null hypothesis.

### 2.1. Approximating an ARFIMA $(p, d, q)$ by an AR Model

Based on a binomial expansion, an ARFIMA  $(p, d, q)$ ,  $y_t$ , satisfying Assumption 1 can be represented by an infinite order AR process (Brockwell and Davis (1991)), i.e.,  $y_t = \sum_{j=1}^{\infty} \beta_j y_{t-j} + e_t$ , where  $\beta_j \sim \delta j^{-d-1}$ , as  $j \rightarrow \infty$ ,  $\delta$  is a constant, and  $e_t$  is a zero mean white noise process with variance  $\sigma^2$ . Our analysis centers on figuring out the conditions needed for the growth rate of lag length  $k$  such that  $y_t$  can be approximated well by an AR( $k$ ) model, i.e.,

$$y_t = e_t + \sum_{j=1}^k \beta_j y_{t-j} + \sum_{j=k+1}^{\infty} \beta_j y_{t-j} = e_{t,k} + \sum_{j=1}^k \beta_j y_{t-j},$$

and

$$\sum_{j=k+1}^{\infty} \beta_j y_{t-j} = o_p(1).$$

Theorems 1-4 establish that the OLS residuals  $\hat{e}_{t,k}$  from an AR( $k$ ) filter are able to mimic the statistical properties of  $e_t$  asymptotically. For ease of reading, we put all the proofs in the appendix.

To show that :

$$\hat{e}_{t,k} = e_t + o_p(1), \tag{2}$$

we need to investigate the asymptotic properties of  $\hat{\beta}' = (\hat{\beta}_1, \dots, \hat{\beta}_k)$  and  $\sum_{j=k+1}^{\infty} \beta_j y_{t-j}$ .

We note that the OLS residuals  $\hat{e}_{t,k}$  can be rewritten as:

$$\hat{e}_{t,k} = e_t + \sum_{j=k+1}^{\infty} \beta_j y_{t-j} - \sum_{j=1}^k (\hat{\beta}_j - \beta_j) y_{t-j}.$$

Theorem 1 discusses the asymptotic properties of the OLS estimator  $\hat{\beta}$ , which is obtained from regressing  $y_t$  on  $y_{t-1}, \dots, y_{t-k}$  and the asymptotic properties of

the OLS residuals  $\widehat{e}_{t,k}$  and those of the residual variance  $s_{e,k}^2$  defined as  $s_{e,k}^2 = (T - k)^{-1} \sum_{t=k+1}^T \widehat{e}_{t,k}^2$ .

Following Berk (1974), we let the norm of a matrix  $D$  be defined as

$$\|D\| = \sup \|Dx\|, \quad \text{where } \|x\| \leq 1.$$

The Euclidean norm of the column vector  $x$  is used, and  $\|x\|^2 = x'x$ . Berk (1974) shows that  $\|D\|^2 \leq \sum_{i,j} d_{i,j}^2$ , where  $d_{i,j}$  is the  $(i, j)$ th element of matrix  $D$ , and  $\|D\|$  is dominated by the largest modulus of  $D$ 's eigenvalues.

**THEOREM 1.** *If the data generating process satisfies Assumption 1,  $k = o\left(T^{\frac{0.5-d}{1-d}}\right)$ ,  $d \in (0, 0.5)$  then as  $T \rightarrow \infty$ :*

1.  $\|\widehat{\beta} - \beta\| = O(k^{0.5-d}T^{d-0.5})$ , when  $d \in (0, 0.5)$ .
2.  $\widehat{e}_{t,k} = e_t + o_p(1)$ , when  $d \in (0, 0.5)$ .
3.  $s_{e,k}^2 \xrightarrow{p} \sigma^2$ , when  $d \in (0, 0.5)$ .

Item 1 of Theorem 1 shows that the OLS estimator  $\widehat{\beta}$  is a consistent estimator for the population parameter vector  $\beta$  when  $k \rightarrow \infty$  as  $T \rightarrow \infty$ . However, item 1 of Theorem 1 also says that the convergence rate of  $\widehat{\beta}_j$  to  $\beta_j$  could be slow if  $d$  is close to 0.5. Nevertheless, regardless of the convergence rate of  $\widehat{\beta}_j$ , item 2 of Theorem 1 clearly indicates that a stationary and invertible fractional white noise process can be well approximated by an AR( $k$ ) model when the lag length  $k$  is chosen appropriately. We note that the growth rate  $k$  in Theorem 1 is only a sufficient condition to ensure that a stationary and invertible fractional white noise process can be approximated well by an AR model.

## 2.2. Asymptotic Properties of Haugh's Statistics When the DGP are the ARFIMA $(p, d, q)$ processes

**THEOREM 2.** *If the data generating processes satisfy Assumption (1), and all the conditions in Theorem 1 hold, then as  $T \rightarrow \infty$ , both  $S_M$  and  $S_M^*$  are asymptotically distributed as  $\chi_{2M+1}^2$  when  $e_{1,t}$  is independent of  $e_{2,s}$  for all  $t$  and  $s$ .*

Theorems 1 and 2 provide the sufficient condition to ensure that the Haugh's statistics is asymptotically Chi-Square distributed. As a matter of fact, we only need a suitable finite  $k$  to ensure that the Haugh's statistics are asymptotically Chi-Square distributed.

**THEOREM 3.** *If the data generating processes satisfy Assumption 1, and the shocks of the two processes  $e_{1,t}$  and  $e_{2,s}$  are independent for all  $t$  and  $s$ , there exists a positive finite constant  $c$  such that for  $k \geq c$ , when  $T \rightarrow \infty$ , both  $S_M$  and  $S_M^*$  are asymptotically distributed as  $\chi_{2M+1}^2$ .*

Theorem 3 yields the theoretical foundation for Haugh's statistics using AR( $k$ ) to prewhiten the ARFIMA ( $p, d, q$ ) processes. The significant difference between Theorem 1 and 2 and that of Theorem 3 is on the condition for lag length  $k$ . Theorem 3 implies the appropriate lag length  $k$  needed to approximate ARFIMA ( $p, d, q$ ) is finite and could be very small. This could be useful for empirical analysis. Moreover, the following Theorem 4 shows that based on the reasoning of Theorem 2, the power of both Haugh's statistics approach to 1.

**THEOREM 4.** *If the data generating processes satisfy Assumption (1), and all the conditions in Theorem 1 hold, then as  $T \rightarrow \infty$ , the power of both  $S_M$  and  $S_M^*$  approach to 1 when  $e_{1,t}$  is correlated with  $e_{2,s}$ .*

### 3. Order Selection for an AR Approximation

The adequacy of an approximate model for the DGP depends on the choice of order of the AR approximation. This section considers two commonly used criteria for choosing a suitable AR ( $k$ ) model for approximating ARFIMA processes, the Akaike's (1973) AIC and the Mallows's  $C_p$  criterion. Through a Monte Carlo study, we show that both suggest a lower order AR( $k$ ) is sufficient to prewhiten an ARFIMA process.

Beran (1995) derived a version of the AIC for determining an appropriate autoregressive order for a class of finite order fractional autoregressive processes as

$$AIC(k) = \log \hat{\sigma}_\epsilon^2 + \frac{2(k+2)}{T}, \quad (3)$$

where  $\hat{\sigma}_\epsilon^2$  is the estimate of the variance  $\sigma_\epsilon^2$ , when the true model is  $y_t = \sum_{j=1}^k \beta_j y_{t-j} + e_t$ . However, our true model,  $y_t = \sum_{j=1}^{\infty} \beta_j y_{t-j} + e_t$ , is of infinite dimension. Ng & Perron (2005) shows that such a model can be approximated well by a finite order AR model. Hence, we modify (3) as

$$AIC(k) = \log \hat{s}_{e,k}^2 + \frac{2(k+2)}{N}.$$

where  $N = T - 2k$ .

Similarly, we modify the Mallows's (1973)  $C_p$  criterion to minimize

$$C_p^* = \frac{(T-2k)\hat{s}_{e,k}^2}{\sigma_\epsilon^2} + 2k.$$

### 3.1. Finite Sample Performance of an AR Approximation for ARFIMA model

Monte Carlo experiments are conducted to examine the finite sample properties of our analytical results. The Monte Carlo experiment for each model is based on 1,000 replications with different sample size  $T$ . We follow McLeod and Hipel (1978) to first generate  $T$  independent values from the standard normal distribution and form a  $T \times 1$  column vector  $e$ . We then calculate the  $T$

autocovariances of the  $I(d)$  process, from which we construct the  $T \times T$  variance-covariance matrix  $\Sigma$  and compute its Cholesky decomposition  $C$  (i.e.,  $\Sigma = CC'$ ). Finally, the vector  $p$  of the  $T$  realized values of the  $I(d)$  process is defined by  $p = Ce$ . We discard the first 200 values.

In light of the empirical evidence provided by Cheung (1993), we simulate several ARFIMA ( $p.d.q$ ) processes with  $d = (0.1, 0.2, 0.3, 0.4, 0.45)$  and  $\sigma^2 = 1$ .

$$\text{DGP (a). } (1 - L)^d y_t = e_t,$$

$$\text{DGP (b). } (1 - 0.7L)(1 - L)^d y_t = (1 + 0.5e_t),$$

$$\text{DGP (c). } (1 - 0.7L - 0.3L^2)(1 - L)^d y_t = e_t,$$

$$\text{DGP (d). } (1 - 0.7L - 0.3L^2)(1 - L)^d y_t = (1 + 0.5e_t),$$

$$\text{DGP (e). } (1 - 0.7L - 0.3L^2 - 0.4L^3)(1 - L)^d y_t = e_t$$

All simulations are performed for  $T = 200, 500$  with  $k$  ranging from 1 to some maximal order  $k_{max}$ . According to Theorem 1,  $k_{max}$  increases with sample size  $T$  and we can set  $k_{max} = o(T^{\frac{0.5-d}{1-d}})$ . Thus, with respect to the different combinations of  $d$  and  $T$ , the values of  $k_{max}$  can be calculated. We set  $k_{max} = 9$  and  $k_{max} = 14$  for all cases of  $T = 200$  and  $T = 500$ , respectively, regardless of  $d$ .

Because nearly identical results between AIC and  $C_p$  Table 1 only displays the frequency of the order selected by AIC for selecting AR ( $k$ ) fitted to the DGP (a)-(e). To save the space, only the detailed model selection results for  $d = 0.1$ ,  $d = 0.4$  and  $d = 0.45$  are displayed. The results for  $d = 0.2$  and  $d = 0.3$  are available from authors. Table 1 shows the lag length needed to approximate a fractional white noise processes is indeed very small. Particularly, for  $d = 0.1$ , no matter what the sample size and model selection criteria used, the value of  $k$  which minimizes these two selection criteria is 1. The largest value of  $k$  selected by  $C_p$  for the combination of  $T = 500$  and  $d = 0.45$  is also only 3. In

particular, for  $T = 200, d = 0.4$ , the order chosen by the AIC is very similar to those reported in Beran (1995) as well. In Table 2, we report the mean of residual variance  $\widehat{s}_{e,k}^2$  when  $k = k^*$  which is the suitable lag length chosen by AIC. The results reveal that when  $k = k^*$ , the mean of  $\widehat{s}_{e,k}^2$  is indeed very close to the true  $\sigma^2$  which is set to be  $\sigma^2 = 1$  in our experiment. It suggests the performance of  $\widehat{e}_{t,k}$  asymptotically approaches to that of the white noise error  $e_t$ .

Likewise, the results show that the lag length needed to approximate the more general ARFIMA  $(1, d, 1)$  process is close to that found in the case of DGP (a). For instance, in Table 1, the lag length  $k$  needed to approximate GDP (b) is 2 when  $T = 200$ , while the value of  $k$  is only 1 in the case of the fractional white noise process. Further, as  $T = 500$ , the largest lag length  $k$  to approximate the DGP (c) is only 2 for  $d \leq 0.3$  and 4 for  $d = 0.4$  and  $d = 0.45$ . In spite of a relatively higher autoregressive order in DGP (c), DGP (d) and DGP (e), the appropriate lag length  $k$  to approximate those three models are still very small. Apparently, the results reveal that for  $T = 200$  and  $T = 500, d = 0.1$ ,  $k = 2$  is large enough to approximate model (c)-(e) well; for the case of  $T = 200, d \in (0, 0.5)$  and the case of  $T = 500, d \leq 0.2$ ,  $k = 3$  and  $k = 4$  are what we need for each case, respectively. Further evidence on how well an appropriate lag length  $k$  for approximating Model (b)-Model (e) are also reported in Table 2. The mean of the residual variance is again close to 1.0, when  $k = k^*$ .

Our simulation evidence appears to indicate that the suitable lag length  $k$  will rise slightly with the increase of the AR order in ARFIMA  $(p, d, q)$  processes. In most cases,  $k = p + 1$  or  $k = p + 2$  would be the appropriate lag length for approximating an ARFIMA  $(p, d, q)$  process with the finite sample size.

### **3.2. Finite sample Properties of Haugh's Statistics When DGP are the ARFIMA Processes**

In addition to reporting the finite sample performance of Haugh's statistics, we present the conventional procedures of examining the independence between two stationary long memory processes. We first estimate the differencing parameter,  $d$  with the maximum likelihood method, then use t-ratio to be the test statistic. Those results are shown in Table 3. It shows that with the increase of the differencing parameter,  $d$ , the bias of maximum likelihood estimates of  $d$  becomes larger. At 5% significant level, there are significant size distortions for  $t$  ratio test. Table 3 also confirms that the significant spurious regression appears to be present, when  $T = 200$ . To save space, only the results for DGP (b) are displayed in this paper and the results for other DGPs are available from the authors. Table 4 and 5 examine the size and power of Haugh's  $S_M$  and  $S_{M^*}$  test, based on  $AR(k)$  approximation at 5% significant level where  $Y_{1t}$  and  $Y_{2t}$  are the ARFIMA  $(1, d, 1)$ . They show that the empirical sizes are very close to 5% for every pair of  $d_1$  and  $d_2$  where the two processes are uncorrelated with each other. In other words, the size control of Haugh's two tests is extremely well for the DGP to be the ARFIMA  $(1, d, 1)$  processes. We also consider the power of Haugh's statistic by considering processes whose pattern of short cross correlation is same as those of many financial time series, as described in Hong (1996). We assume that the error term  $e_{1,t}$  and  $e_{2,t}$  are generated as:  $\rho_{12}(j) = 0.2$  for  $j = 0$  and  $\rho_{12}(j) = 0$  for  $j \neq 0$  where  $\rho_{12}(j)$  denotes the cross correlation function of  $e_{1t}$  and  $e_{2t}$  at lag  $j$ . The results in Table 5 also shows that the rejection percentage of 1000 replications at 5% are similar to those of Table 2 in Hong (1996), where two time series he considered are short memory processes. The power performance of Haugh's statistics does not depend on the long memory characteristics of the data series, as long as the data series themselves are stationary. The predictions made in Theorem 2 are clearly borne out in our simulation studies as they clearly demonstrates that the Haugh's statistics derived from prewhitening ARFIMA  $(p, d, q)$  processes by  $AR(k)$  works well in finite sample.

## 4. Application to foreign exchange rates volatility

In this section, we illustrate our methodology by considering the volatility of foreign exchange rates. The emergence of the Euro (EUR) and the Japanese Yen (JPY) as the two major currencies of the world has raised questions about how the two currencies will interact and what it will mean for foreign exchange rate markets. The data series under this study are daily spot exchange rates from January 3, 2000 to August 31, 2005, with totally 1426 observations. Define the volatilities of the EUR ( $Y_{Et}$ ) and the JPY ( $Y_{Yt}$ ) as squared returns of those currencies,

$$Y_{Et} = R_{Et}^2 = (\ln(P_{Et,t}/P_{Et,t-1}))^2 \quad Y_{Yt} = R_{Yt}^2 = (\ln(P_{Yt,t}/P_{Yt,t-1}))^2$$

where  $P_{Et}$  and  $P_{Yt}$  are the daily prices of two important nominal U.S dollar exchange rates—the Euro and the Japanese Yen, respectively.

Our first task is to check the statistics properties of  $Y_{Et}$  and  $Y_{Yt}$ . We use the Conditional Sum of Squares (CSS) estimation method discussed in Chung and Baillie (1993). The results of estimating ARFIMA  $(1, d, 1)$  models for  $Y_{Et}$  and  $Y_{Yt}$  are reported in Table 7. The volatility of the Euro and the Japanese Yen are estimated as  $I(0.3224)$  and  $I(0.1209)$  with highly significant AR and MA parameter estimates  $(\phi, \theta)$ .

We then use AIC to select the order of an AR ( $k$ ) approximations of stationary long memory processes. Setting  $k_{Max} = 25$ , the chosen order of AR ( $k$ ) models fitted to  $Y_{Et}$  and  $Y_{Yt}$  by AIC is at  $k = 22$  for both series.

In spite of noting that the lag length  $k = 22$  is more suitable order of an AR( $k$ ) for approximating  $Y_{Et}$  and  $Y_{Yt}$ , we also check the robustness of the inference by considering empirical evidence at  $k, k = 16, \dots, 25$ . Table 7 reports the results. The performances of  $S_M$  and  $S_M^*$  are all significant at 5% level for  $M = 5, 10, 15, 20, 22, 25$ . The test results appear to suggest there exists "strong contemporaneous correlation" between the volatility of Euro and Yen within 25 days, even though both of these two series are long memory processes.

## 5. Conclusion

In this paper, we have proved that a stationary long memory ARFIMA  $(p, d, q)$  process can be approximated well by an AR( $k$ ) model when  $k$  is chosen appropriately and show the applicability of Haugh's (1976) statistics to test for two long memory processes being uncorrelated based on the sample cross-correlation function of AR prewhitened ARFIMA processes. The new test is easy to implement and avoids issues arising from inaccurate estimation of  $d$  nor the issues of spurious regression induced by the long memory processes considered in Tsay and Chung (2000). We also demonstrated the desirability of using Akaike information criterion to select the order of an autoregression for approximating long memory processes.

Monte Carlo experiments conducted in this paper confirm our theoretical prediction. We find that Haugh's statistics based on an AR approximation is an accurate and powerful method to detect the independence between two ARFIMA  $(p, d, q)$  processes. We also applied our methodology to investigate the independence between volatilities of two daily nominal dollar exchange rates- Euro and Japanese Yen. We found that there existed " strong contemporaneous correlation " between the volatilities of Euro and Yen within 30 days.

**Table 1. The Frequencies of Selected Lag Lengths for AIC respectively  
When the DGP are models (a)-(e)**

DGP	$k$			1	2	3	4	5	6	7	8	9
	$d$	$T$										
(a)	0.1	200	<i>AIC</i>	758*	134	57	25	12	8	1	1	1
		500	<i>AIC</i>	639*	192	81	45	24	6	5	6	1
	0.4	200	<i>AIC</i>	238	361*	201	101	53	23	10	7	2
		500	<i>AIC</i>	7	154	262*	214	143	95	47	27	21
	0.45	200	<i>AIC</i>	198	365*	222	110	53	23	16	7	2
		500	<i>AIC</i>	2	129	259*	223	146	99	53	32	23
(b)	0.1	200	<i>AIC</i>	271	505*	121	57	24	9	4	4	1
		500	<i>AIC</i>	22	650*	113	137	36	21	9	7	1
	0.4	200	<i>AIC</i>	7	634*	94	159	48	31	13	6	1
		500	<i>AIC</i>	1	270	32	345*	106	108	53	39	14
	0.45	200	<i>AIC</i>	4	630*	87	169	49	32	13	8	2
		500	<i>AIC</i>	1	230	30	366*	104	124	56	39	16
(c)	0.1	200	<i>AIC</i>	54	678*	169	56	27	5	4	3	1
		500	<i>AIC</i>	1	518*	268	112	56	22	5	9	2
	0.4	200	<i>AIC</i>	91	9	412*	273	112	46	29	16	5
		500	<i>AIC</i>	1	1	82	245*	236	175	119	69	56
	0.45	200	<i>AIC</i>	47	8	404*	295	125	57	36	21	9
		500	<i>AIC</i>	1	1	57	233*	245	185	131	79	61
(d)	0.1	200	<i>AIC</i>	143	241	428*	96	49	25	9	4	1
		500	<i>AIC</i>	4	42	624*	124	141	24	19	12	2
	0.4	200	<i>AIC</i>	74	7	639*	72	117	41	22	14	5
		500	<i>AIC</i>	1	1	375*	45	317	80	81	37	27
	0.45	200	<i>AIC</i>	47	2	658*	78	125	40	22	14	7
		500	<i>AIC</i>	1	1	344*	39	331	90	89	58	45
(e)	0.1	200	<i>AIC</i>	7	1	725*	165	64	20	6	4	2
		500	<i>AIC</i>	1	1	549*	251	114	46	14	12	12
	0.4	200	<i>AIC</i>	6	8	46	465*	277	106	53	29	14
		500	<i>AIC</i>	1	11	95	255*	249	177	115	47	26
	0.45	200	<i>AIC</i>	3	8	14	466*	289	119	53	27	8
		500	<i>AIC</i>	1	1	2	69	247*	254	179	116	49

*Notes:* The results are all based on 1,000 replications. \* denotes the lag length  $k^*$  at which simulated numbers chosen by AIC and  $C_p$  are the maximum

**Table 2. Results on Mean of Residual Variance  $S_{\epsilon,k}^2$  when  $k = k^*$**

DGP	$d$	0.1	0.4	0.45
	$T$			
(a)	200	0.997	1.007	0.997
	500	1.002	1.009	1.010
(b)	200	0.995	1.007	1.003
	500	1.000	1.007	1.008
(c)	200	0.999	1.001	1.009
	500	1.014	1.007	1.008
(d)	200	1.000	1.001	1.002
	500	1.011	1.008	1.009
(e)	200	1.000	1.001	1.002
	500	1.013	1.011	1.012

**Table 3. The Rejection Percentages of the  $t$  When  $Y_{1t} = \text{ARFIMA}(0, d_1, 0)$ ,  $Y_{2t} = \text{ARFIMA}(0, d_2, 0)$ , and  $\rho_{12}(j) = 0$  for all  $j$**

$Y_{1t}$	$d_1$	0.1	0.2	0.3	0.4	0.45
$Y_{2t}$	$d_2$					
	0.1	10.4	14.5	15.3	18.6	24.1
	0.2	14.3	15.0	15.5	19.7	27.3
	0.3	15.7	16.0	16.8	24.7	30.2
	0.4	18.0	20.2	24.2	28.9	32.7
	0.45	24.2	27.8	30.8	33.1	33.4

*Notes:*  $T = 200$ . The critical value of the above two-tailed  $t$  test is at 5% level of significance.

**Table 4. The Size and Power Performance of the  $S_M$  Test at 5% Level of Significance When  $Y_{1t} = \text{ARFIMA}(-0.7, d_1, 0.5)$ ,  $Y_{2t} = \text{ARFIMA}(-0.7, d_2, 0.5)$**

	$Y_{1t}$	$d_1$	$k$	$M$	0.1	0.2	0.3	0.4	0.45
Size	$Y_{2t}$	$d_2$							
		0.1	$k_{AIC}$	5	5.3	5.4	5.1	5.2	5.5
				9	5.1	5.1	5.3	5.2	5.1
				15	4.6	4.7	4.8	5.1	5.9
		0.2	$k_{AIC}$	5	5.2	5.5	5.7	5.6	5.4
				9	5.4	5.3	5.2	5.4	5.3
				15	4.7	5.0	5.6	5.4	4.4
		0.3	$k_{AIC}$	5	4.4	4.7	4.8	4.9	4.8
				9	4.3	4.1	4.2	4.7	4.6
				15	3.9	4.0	4.1	4.1	4.6
		0.4	$k_{AIC}$	5	5.6	5.8	4.9	4.6	5.5
				9	4.9	4.6	4.3	4.8	4.9
				15	4.0	3.9	4.2	4.2	4.4
		0.45	$k_{AIC}$	5	4.4	4.3	4.6	5.0	4.9
				9	4.9	4.6	4.7	4.5	4.5
				15	4.0	4.0	4.1	4.2	4.6
Power		0.1	$k_{AIC}$	5	40.8	41.0	40.5	40.3	40.0
				9	30.2	30.4	30.4	30.1	30.1
				15	23.0	23.3	23.6	23.2	23.8
		0.2	$k_{AIC}$	5	40.7	41.4	41.0	40.9	40.6
				9	29.4	28.9	29.1	30.3	30.0
				15	21.6	21.8	24.4	23.7	23.4
		0.3	$k_{AIC}$	5	40.0	40.2	40.3	40.5	40.7
				9	30.2	30.2	29.7	30.5	30.3
				15	21.3	21.2	21.6	21.8	21.9
		0.4	$k_{AIC}$	5	39.1	39.6	40.2	39.9	40.2
				9	29.3	28.6	30.3	30.3	30.1
				15	21.2	21.3	21.7	21.9	21.9
		0.45	$k_{AIC}$	5	38.9	39.5	39.6	40.0	40.1
				9	29.9	29.8	28.8	29.3	30.0
				15	20.0	20.4	21.9	22.1	22.0

Notes:  $T = 200$ ,  $k = k_{AIC}$ , and the results are all based on 1,000 replications. The critical value used is obtained from  $\chi_{2M+1}^2$  at 5% level of significance.

**Table 5. The Size and Power Performance of the  $S_M^*$  Test at 5% Level of Significance When  $Y_{1t} = \text{ARFIMA}(-0.7, d_1, 0.5)$ ,  $Y_{2t} = \text{ARFIMA}(-0.7, d_2, 0.5)$**

	$Y_{1t}$	$d_1$	$k$	$M$	0.1	0.2	0.3	0.4	0.45
Size	$Y_{2t}$	$d_2$							
		0.1	$k_{AIC}$	5	5.0	4.9	5.0	5.0	5.0
				9	5.2	5.3	5.3	5.0	5.0
				15	5.7	5.7	5.7	5.4	5.3
		0.2	$k_{AIC}$	5	5.3	4.8	4.9	5.0	5.1
				9	5.1	5.2	5.1	5.2	5.1
				15	5.6	5.4	5.6	5.8	5.7
		0.3	$k_{AIC}$	5	5.3	5.3	5.1	5.1	5.1
				9	5.2	5.3	5.5	5.8	5.3
				15	5.6	6.0	5.7	5.7	5.8
		0.4	$k_{AIC}$	5	5.1	5.2	5.4	5.2	5.0
				9	5.5	5.4	5.5	5.7	5.6
				15	5.6	5.7	5.5	5.8	5.8
		0.45	$k_{AIC}$	5	5.0	5.0	5.2	5.1	5.1
				9	5.6	5.4	5.3	5.6	5.6
				15	5.3	5.7	5.6	5.9	5.9
Power		0.1	$k_{AIC}$	5	41.7	41.7	41.4	41.2	41.3
				9	32.2	33.1	30.9	30.9	32.2
				15	23.5	25.6	26.2	25.2	25.8
		0.2	$k_{AIC}$	5	41.6	41.9	41.7	41.7	41.4
				9	33.8	32.3	33.3	32.3	31.4
				15	25.9	26.1	25.6	25.7	25.6
		0.3	$k_{AIC}$	5	40.8	41.1	41.7	40.9	41.5
				9	32.7	33.0	32.7	32.3	32.3
				15	25.8	26.6	26.2	26.0	25.7
		0.4	$k_{AIC}$	5	39.7	40.5	41.0	40.7	40.9
				9	31.5	32.3	32.7	32.3	32.4
				15	25.9	26.4	26.2	25.9	25.9
		0.45	$k_{AIC}$	5	39.6	40.1	40.9	40.7	40.5
				9	32.3	32.4	32.7	32.1	32.2
				15	25.6	26.1	26.6	25.8	26.3

Notes:  $T = 200$ ,  $k = k_{AIC}$ , and the results are all based on 1,000 replications. The critical value used is obtained from  $\chi_{2M+1}^2$  at 5% level of significance.

**Table 6. Estimation of ARFIMA (1,  $d$ , 1) models for the volatility of Euro and Japanese Yen**

Estimates	Japanese Yen	Euro
$d$	0.1209	0.3224
$\phi$	0.5948	0.3703
$\theta$	-0.6824	-0.6982
$t_d$	1.6699	4.1900
$t_\phi$	4.0956	6.2149
$t_\theta$	-4.8412	-10.2539

**Table 7. Test statistics for correlation in volatility between Euro and Japanese Yen**

k		17	18	19	20	21	22	23
	$M$							
$S_M$	5	38.17*	37.10*	37.15*	39.77*	40.07*	39.75*	39.27*
	10	51.46*	50.33*	50.25*	53.28*	53.82*	53.82*	53.62*
	15	51.61*	55.35*	55.31*	57.76*	58.33*	58.32*	58.20*
	20	65.74*	64.61*	64.60*	66.99*	67.50*	67.83*	67.77*
	22	65.90*	65.78*	66.10*	67.10*	68.80*	69.10*	68.96*
	25	69.36*	68.26*	68.59*	71.23*	71.85*	72.17*	72.17*
$S_M^*$	5	38.21*	37.14*	37.18*	39.81*	40.10*	39.78*	39.31*
	10	51.57*	50.43*	50.35*	53.38*	53.92*	53.93*	53.73*
	15	56.76*	55.50*	55.46*	57.90*	58.47*	58.46*	58.35*
	20	66.00*	64.87*	64.87*	67.25*	67.76*	68.10*	68.04*
	22	66.21*	66.52*	67.10*	68.07*	69.01*	68.73*	69.01*
	25	69.28*	68.58*	68.92*	71.56*	72.18*	72.50*	72.51*

## REFERENCES

- Abramowitz, M. and I. A. Stegun (1970): *Handbook of Mathematical functions*, Dover, New York.
- Akaike, H. (1969): "Power Spectrum Estimation through Autoregressive Model Fitting," *Annals of Institute of Statistics Mathematics*, 21, 407-419.
- Baillie, R. T. (1996): "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5-59.
- Beran, J., R. J. Bhansali and D. Ocker (1995): "On unified model selection for stationary and nonstationary short -and long-memory autoregressive processes," *Biometrika*, 85, 4, 921-934.
- Berk, K. N. (1974): "Consistent Autoregressive Spectral Estimates," *The Annals of Statistics*, 2, 489-502.
- Brockwell, P. J. and R. A. Davis (1991): *Time Series: Theory and Methods*, 2nd Edn. New York: Springer-Verlag.
- Bollerslev, T. and J. H. Wright (2000): "Semiparametrics Estimation of Long-memory Volatility Dependencies: The Role of High-Frequency Data," *Journal of Econometrics*, 98, 81-106.
- Cheung, Y.-W. AND K. S. Lai (1993): "A Fractional Cointegration Analysis of Purchasing Power Parity," *Journal of Business and Economic Statistics*, 11, 103-112.
- Gallant, R. A., Hsieh, D., and G. Tauchen (1997): "Estimation of Stochastic Volatility Models with Diagnostics," *Journal of Econometrics* 81, 159-192.
- Haugh, L. D. (1976): "Checking the Independence of Two Covariance Stationary Time Series: A univariate Residual Cross-Correlation Approach," *Journal of American Statistical Association*, 71, 378-385.

- Hong, Y. (1996): "Testing for Independence Between Two Covariance Stationary Time Series," *Biometrika*, 83, 615-625.
- Hosking, J. R. M. (1996): "Asymptotic Distributions of the Sample Mean, Autocovariances, and Autocorrelations of Long-Memory Time Series," *Journal of Econometrics*, 73, 261-284.
- Mallows, C. L. (1973): "Come Comments on  $C_p$ ," *Technometrics*, 15, 661-675.
- Martin, V. L. and N. P. Wilkins (1999): "Indirect estimation of ARFIMA and VARFIMA models," *Journal of Econometrics* 93, 149-175.
- McLeod, A. I. AND K. W. Hipel (1978): "Preservation of the Rescaled Adjusted Range 1: A Reassessment of the Hurst Phenomenon," *Water Resources Research*, 14, 491-508.
- Ng, S. and P. Perron (2005): "Practitioners' Corner A Note on the Selection of Time Series Models," *Oxford Bulletin of Economics and Statistics*, 67, 1, 115-134.
- Ray, B. K. (1993): "Modeling Long-Memory Processes for Optimal Long-Range Prediction," *Journal of Time Series Analysis*, 14, 511-525.
- Robinson, P. M. (1993): "Highly Insignificant F-Ratios," *Econometrica*, 61, 687-696.
- Robinson, P. M. (2001): "Higher-order Kernel Semiparametric M-estimation of Long memory," *Journal of Econometrics*, 114, 1-27.
- Tsay, W. J. and C.-F. Chung (2000): "The Spurious Regression of Fractionally Integrated Processes," *Journal of Econometrics* 96, 155-182.
- Yajima, Y. (1991): "Asymptotic Properties of the LSE in a Regression Model with Long-Memory Stationary Errors," *The Annals of Statistics*, 19, 158-177.

## APPENDIX

For clarity of exposition, let us define the OLS estimator  $\widehat{\beta}_{k,T}$  as:

$$\widehat{\beta}_{k,T} = \left( \sum_{t=k+1}^T Y_t Y_t' \right)^{-1} \sum_{t=k+1}^T Y_t y_t,$$

where  $\widehat{\beta}'_{k,T} = (\widehat{\beta}_{1,k}, \dots, \widehat{\beta}_{k,k})$ , and  $Y_t' = (y_{t-1}, \dots, y_{t-k})$ . Let

$$\widehat{R}(k) = \sum_{t=k+1}^T Y_t Y_t', \quad \text{and} \quad \widehat{\gamma}(k) = \sum_{t=k+1}^T Y_t y_t,$$

then

$$\widehat{\beta}_{k,T} - \beta = \widehat{R}(k)^{-1} \sum_{t=k+1}^T Y_t e_{t,k},$$

where

$$e_{t,k} = e_t + \sum_{j=k+1}^{\infty} \beta_j y_{t-j}.$$

Suppose

$$R = \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{k-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{k-1} & \gamma_{k-2} & \dots & \gamma_0 \end{bmatrix},$$

where  $\gamma_s$  is the autocovariance function of  $y_t$  at lag  $s$ .

$$\sum_{t=k+1}^T Y_t e_{t,k} = M \quad \text{and} \quad \frac{1}{T-k} \widehat{R}(k) = R^*,$$

then  $\widehat{\beta}_{k,T} - \beta$  can be rewritten as

$$\begin{aligned} \widehat{\beta}_{k,T} - \beta &= \frac{1}{T-k} \left( R^{*-1} - R^{-1} \right) M + \frac{1}{T-k} R^{-1} \left( \sum_{t=k+1}^T Y_t e_t \right) \\ &\quad + \frac{1}{T-k} R^{-1} \left\{ \sum_{t=k+1}^T Y_t (e_{t,k} - e_t) \right\}. \end{aligned} \quad (\text{A.1})$$

Let  $C$  denote an arbitrary finite positive constant throughout this Appendix..

**LEMMA A.1.** *If the data generating process is the ARFIMA  $(p, d, q)$  process, then as  $k \rightarrow \infty$ ,  $T \rightarrow \infty$  and  $k/T \rightarrow 0$ , we have the following results:*

1.  $\|R^{-1}\|$  is bounded for  $d \in (0, 0.5)$ .
2.  $\|R^{*-1} - R^{-1}\| = O_p(kT^{2d-1})$ , when  $d \in (0.25, 0.5)$  and  $k = o(T^{1-2d})$ ;  
 $\|R^{*-1} - R^{-1}\| = O_p(kT^{-0.5}(\text{LOG } T)^{0.5})$  when  $d \in (0, 0.25]$  and  $k = o(T^{0.5}(\text{LOG } T)^{-0.5})$ .

### A.1. Proof of Lemma A.1

The spectral density of the the ARFIMA( $p, d, q$ ) process  $y_t$  is

$$f(\lambda) = \frac{\sigma^2 \left| \theta(e^{-i\lambda}) \right|^2}{2\pi \left| \phi(e^{-i\lambda}) \right|^2} \left| 1 - e^{-i\lambda} \right|^{-2d}.$$

Because the roots of  $\phi(L)$  and  $\theta(L)$  are assumed to be outside the unit circle,  $f(\lambda) > 0$  for all  $\lambda$  in our model setting. Moreover,

$$f(\lambda) = \frac{\sigma^2 [\theta(1)]^2}{2\pi [\phi(1)]^2} \lambda^{-2d}, \quad \text{as } \lambda \rightarrow 0.$$

We thus have

$$\min_{\lambda} f(\lambda) > 0$$

By Proposition 4.5.3 of Brockwell and Davis (1991), the eigenvalues of the covariance matrix  $R$  are all greater than 0, and the eigenvalues of  $R^{-1}$  is uniformly bounded for all  $k$ . This proves item 1 of Lemma A.1.

To prove item 2 of Lemma A.1, let  $Q = R^* - R$ ,  $q = \|R^{*-1} - R^{-1}\|$  and  $p = \|R^{-1}\|$ . Observe that

$$\begin{aligned} q &= \|R^{*-1} - R^{-1}\| = \|R^{*-1}(R - R^*)R^{-1}\| \\ &\leq \|R^{*-1}\| \|R - R^*\| \|R^{-1}\| \\ &\leq (p + q) \|Q\| p, \end{aligned}$$

and

$$q \leq p^2 \|Q\| (1 - p \|Q\|)^{-1}. \quad (\text{A.2})$$

Equation (A.2) indicates that  $\|Q\| \xrightarrow{p} 0$  is the necessary condition for  $q \xrightarrow{p} 0$  to hold.

Let  $q_{i,j}$  be the  $(i, j)$ th element of  $Q$ , where  $i, j = 1, \dots, k$ , i.e.,  $Q$  is a  $k \times k$  matrix with

$$q_{i,j} = (T - k)^{-1} \sum_{t=k+1}^T y_{t-i} y_{t-j} - \gamma_{i-j} = \widehat{\gamma}_{i-j} - \gamma_{i-j},$$

where  $\widehat{\gamma}_s$  denotes the sample autocovariance function of  $y_t$  at lag  $s$ . We first note that  $E(q_{i,j}) = E(\widehat{\gamma}_{i-j}) - \gamma_{i-j} = 0$ , it follows that

$$\begin{aligned} E(q_{i,j}^2) &= E\{(\widehat{\gamma}_{i-j} - \gamma_{i-j})^2\} = \mathbf{E}[\widehat{\gamma}_{i-j} - \mathbf{E}(\widehat{\gamma}_{i-j}) + \mathbf{E}(\widehat{\gamma}_{i-j}) - \gamma_{i-j}]^2 \\ &= \mathbf{E}[\widehat{\gamma}_{i-j} - \mathbf{E}(\widehat{\gamma}_{i-j})]^2 = \text{VAR}(\widehat{\gamma}_{i-j}). \end{aligned}$$

To derive the asymptotic properties of  $\mathbf{E}(q_{i,j}^2)$  or those of  $\text{VAR}(\widehat{\gamma}_{i-j})$ , we first rewrite  $y_{t-i}$  and  $y_{t-j}$  into moving average representations:

$$y_{t-i} = \sum_{s=0}^{\infty} \psi_s e_{t-i-s}. \quad (\text{A.3})$$

Both  $y_{t-i}$  and  $y_{t-j}$  satisfy the conditions in equations (1), (2), and (3) of Hosking (1996, p.262) automatically. Given that item (iii) of Assumption 1 is satisfied, i.e.,  $\mathbf{E}(e_t^4) < \infty$ , then the conditions imposed in Theorem 3 of Hosking (1996) are all fulfilled. Following the arguments in the proof of Theorem 3 of Hosking (1996), we show that  $\mathbf{E}(q_{i,j}^2)$  is uniformly bounded by

$$\text{MAX}_{1 \leq i, j \leq k} \mathbf{E}(q_{i,j}^2) = \begin{cases} O(T^{4d-2}), & \text{if } 0.25 < d < 0.5, \\ O(T^{-1}(\text{LOG } T)), & \text{if } 0 < d \leq 0.25. \end{cases} \quad (\text{A.4})$$

Note that both  $i$  and  $j$  might increase with the sample size  $T$ .

Given the results in (A.4) and  $\|D\|^2 \leq \sum_{i,j} d_{i,j}^2$ , where  $d_{i,j}$  is the  $(i, j)$ th element of matrix  $D$ , it follows that

$$\begin{aligned} \mathbf{E}\|Q\|^2 &\leq \sum_{1 \leq i, j \leq k} \text{MAX}_{1 \leq i, j \leq k} \mathbf{E} \left\| (T - k)^{-1} \sum_{t=k+1}^T y_{t-i} y'_{t-j} - \gamma(i - j) \right\|^2 \\ &= \begin{cases} O(k^2 T^{4d-2}), & \text{if } 0.25 < d < 0.5, \\ O(k^2 T^{-1}(\text{LOG } T)), & \text{if } 0 < d \leq 0.25. \end{cases} \end{aligned}$$

As  $d \in (0.25, 0.5)$ , following the arguments of Lütkepohl and Poskitt (1996, p.69),  $\left\| k^{-1} T^{1-2d} \left( \widehat{\Gamma}_k - \Gamma_k \right) \right\|$  is bounded in probability, i.e.,  $\left\| \widehat{\Gamma}_k - \Gamma_k \right\| = O_p(kT^{2d-1})$ . The preceding arguments can be similarly applied to the cases  $d \in (0, 0.25]$  to show that

$$\|Q\| = \begin{cases} O_p(kT^{2d-1}), & \text{if } 0.25 < d < 0.5, \\ O_p(kT^{-0.5}(\text{LOG } T)^{0.5}), & \text{if } 0 < d \leq 0.25. \end{cases} \quad (\text{A.5})$$

Given the results in (A.5),  $k = o(T^{1-2d})$  must be imposed to make  $\|Q\|^2 \xrightarrow{p} 0$  as  $d > 0.25$ ; when  $0 < d \leq 0.25$ ,  $k = o(T^{0.5}(\text{LOG } T)^{-0.5})$  is needed to ensure that  $\|Q\|^2 \xrightarrow{p} 0$ . Conditional on these restrictions being imposed, by (A.2)

$$\|Q\| \xrightarrow{p} 0 \quad \text{when} \quad \begin{cases} k = o(T^{1-2d}), & \text{if } 0.25 < d < 0.5, \\ k = o(T^{0.5}(\text{LOG } T)^{-0.5}), & \text{if } 0 < d \leq 0.25. \end{cases} \quad (\text{A.6})$$

Thus, when these conditions are all fulfilled,  $\|Q\| \xrightarrow{p} 0$ , and item 2 of Lemma 1 is established.

**LEMMA A.2.** *Given that the conditions in Lemma A.1 hold, and for some positive constant  $C$  and  $K_p$  as  $T \rightarrow \infty$ , we have the following results:*

1.  $\left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_t e_t \right\| = C\sigma^2 k^{1/2} T^{1/2} = O_p(k^{1/2} T^{-1/2})$ , if  $d \in (0, 0.5)$ .
2.  $\left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_t (e_{t,k} - e_t) \right\| = CK_p \sigma^2 k^{0.5-d} T^{d-0.5} = O_p(k^{0.5-d} T^{d-0.5})$ , if  $d \in (0, 0.5)$ .
3.  $\left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_t e_{t,k} \right\| = O_p(k^{-d} T^{0.5+d})$ , if  $d \in (0, 0.5)$ .

## A.2. Proof of Lemma A.2

We note that

$$\begin{aligned} E \left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_t e_t \right\|^2 &= (T-k)^{-2} \sum_{i=1}^k E \left( \sum_{t=k+1}^T y_{t-i} e_t \right)^2 \\ &= (T-k)^{-1} \sum_{i=1}^k \gamma_0 \sigma^2 = C \gamma_0 \sigma^2 k T^{-1} = O(k T^{-1}), \end{aligned}$$

where  $C$  is a positive finite constant. Additionally, because  $y_{t-i}$  is independent of  $e_t$  when  $i > 0$ , and item 1 of Lemma A.2 is obtained.

To prove item 2 of Lemma A.2, we observe that

Observe that the 2nd moment norm of  $Z = \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j y_{t-j} y_{t-i}$  satisfies

$$\{\mathbf{E}|Z|^2\}^{1/2} \leq |EZ| + \{\mathbf{E}|Z - EZ|^2\}^{1/2} \equiv A^* + B^* \quad (\text{A.9})$$

To calculate the asymptotic properties of  $A^*$  and  $B^*$  in (A.5), let us first define  $\gamma_{s-t} = \mathbf{E}(y_{t-i} y_{s-i})$  and  $\gamma_{s-t}^* = \mathbf{E} \left\{ \left( \sum_{j=k+1}^{\infty} \beta_j y_{t-j} \right) \left( \sum_{j=k+1}^{\infty} \beta_j y_{s-j} \right) \right\}$ . Using First Moment Bound Theorem in Findley and Wei (1993), we find that

$$B^* \leq K_p \left( \sum_{s=k+1}^T \sum_{t=k+1}^T \gamma_{s-t} \gamma_{s-t}^* \right)^{1/2} \quad (\text{A.8})$$

where  $K_p$  denotes some positive finite constant.

Furthermore, we note that

$$\sum_{s=k+1}^T \sum_{t=k+1}^T |\gamma_{s-t} \gamma_{s-t}^*| \leq \gamma_0^* \sum_{s=k+1}^T \sum_{t=k+1}^T |\gamma_{s-t}| \leq \gamma_0^* (T-k) \sum_{i=-(T-k-1)}^{T-k-1} |\gamma_i|. \quad (\text{A.9})$$

The definition of the spectral density of  $\{y_t\}$  is

$$f_y(\lambda) = \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2} |1 - e^{-i\lambda}|^{-2d} \sigma^2 / (2\pi) = \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2} |2\sin(\lambda/2)|^{-2d} \sigma^2 / (2\pi)$$

and the autocovariances are

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda = \frac{\sigma^2}{\pi} \int_0^{\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2} \cos(h\lambda) (2\sin(\lambda/2))^{-2d} d\lambda. \quad h = 0, 1, 2, \dots$$

Thus,

$$\begin{aligned}
\gamma_0^* &= \int_{-\pi}^{\pi} \left| \sum_{j=k+1}^{\infty} \beta_j e^{-ij\lambda} \right|^2 f(\lambda) d\lambda \\
&\leq \left| \sum_{j=k+1}^{\infty} \sum_{h=k+1}^{\infty} \beta_j \beta_h \right| \int_{-\pi}^{\pi} e^{i(h-j)\lambda} f(\lambda) d\lambda \\
&= \left| \sum_{j=k+1}^{\infty} \sum_{h=k+1}^{\infty} \beta_j \beta_h \right| \gamma_{h-j} \\
&\leq \gamma_0 \left( \sum_{j=k+1}^{\infty} \beta_j \right)^2 + \left| \sum_{j=k+1}^{\infty} \sum_{h=k+1, h \neq j}^{\infty} \beta_j \beta_h \gamma_{h-j} \right| \\
&\leq \gamma_0 \left( \sum_{j=k+1}^{\infty} \beta_j \right)^2 + C \sum_{j=k+1}^{\infty} \sum_{h=k+1, h \neq j}^{\infty} j^{-d-1} h^{-d-1} (h-j)^{2d-1} \\
&\leq \gamma_0 \left( \sum_{j=k+1}^{\infty} \beta_j \right)^2 + C \sigma^2 k^{-2d} \\
&= C k^{-2d} \sigma^2
\end{aligned} \tag{A.10}$$

By (A.10), we can get the order of the second term on the right side of (A.9) being  $O(T^{1+2d} k^{-2d})$ . Thus,

$$B^* \leq K_p \left( \sum_{s=k+1}^T \sum_{t=k+1}^T \gamma_{s-t} \gamma_{s-t}^* \right)^{1/2} = C \sigma^2 K_p T^{0.5+d} k^{-d} \tag{A.11}$$

Let

$$y_{t-j} = \sum_{l_1=0}^{\infty} \phi_{l_1} e_{t-j-l_1}, \quad y_{t-i} = \sum_{l_2=0}^{\infty} \phi_{l_2} e_{t-i-l_2}.$$

Then

$$\begin{aligned}
A^* &= |EZ| = \left| \mathbf{E} \left( \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j y_{t-j} y_{t-i} \right) \right| \\
&= \left| \left\{ \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j \left( \sum_{l_1=0}^{\infty} \phi_{l_1} \phi_{(j-i)+l_1} \right) \sigma^2 \right\} \right| \\
&\sim \left| \left\{ \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j \left( \int_{l_1=0}^{\infty} l_1^{d-1} (l_1 + (j-i))^{d-1} dl_1 \right) \sigma^2 \right\} \right|
\end{aligned} \tag{A.12}$$

Using the integration by parts, we note

$$\int_{l_1=0}^{\infty} l_1^{d-1} (l_1 + (j-i))^{d-1} dl_1 + \frac{d-1}{d} \int_{l_1=0}^{\infty} l_1^d (l_1 + (j-i))^{d-2} dl_1 = \frac{1}{d} l_1^d (l_1 + (j-i))^{d-1} \Big|_0^{\infty}.$$

Thus,

$$O\left(\int_{l_1=0}^{\infty} l_1^{d-1} (l_1 + (j-i))^{d-1} dl_1\right) \leq O\left(\frac{1}{d} l_1^d (l_1 + (j-i))^{d-1} \Big|_0^{\infty}\right)$$

or

$$\left(\int_{l_1=0}^{\infty} l_1^{d-1} (l_1 + (j-i))^{d-1} dl_1\right) = C^* \left(\frac{1}{d} l_1^d (l_1 + (j-i))^{d-2} \Big|_0^{\infty}\right),$$

where  $C^*$  denotes some finite positive constant. Hence, following the argument of Robinson (1993, p.693),

$$\sum_{j=k+1}^{\infty} j^{-d-1} \sim \int_{k+1}^{\infty} t^{-d-1} dt = O(k^{-d}).$$

Then, we can observe (A.12) and have

$$\begin{aligned} A^* &= \left\{ \sigma^2 \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j \left( \int_{l_1=0}^{\infty} l_1^{d-1} (l_1 + (j-i))^{d-1} dl_1 \right) \right\} \\ &\leq \sigma^2 C \left\{ \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j \left( \frac{1}{d} T^d (T + (j-i))^{d-1} \right) \right\} \\ &\sim \sigma^2 C \left\{ \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} j^{-d-1} \left( \frac{1}{d} T^d (T + (j-i))^{d-1} \right) \right\} \\ &\leq C \sigma^2 \left\{ T k^{-d} \left( \frac{1}{d} T^d (T + ((k+1) - i))^{d-1} \right) \right\} \end{aligned} \tag{A.13}$$

Combining (A.7),(A.11) and (A.13), we have

$$\begin{aligned} &\left\{ \sum_{t=k+1}^T \sum_{j=k+1}^{\infty} \beta_j y_{t-j} y_{t-i} \right\}^2 \leq C^2 \sigma^4 \left\{ T^2 k^{-2d} \left( \frac{1}{d^2} T^{2d} (T + ((k+1) - i))^{2d-2} \right) \right\} \\ &+ C \sigma^4 \left\{ T k^{-d} \left( \frac{1}{d} T^d (T + ((k+1) - i))^{d-1} \right) \right\} C K_p T^{0.5+d} k^{-d} \\ &+ C^2 K_p^2 \sigma^4 T^{1+2d} k^{-2d} \end{aligned} \tag{A.14}$$

Therefore,

$$\begin{aligned} \left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_t(e_{t,k} - e_t) \right\|^2 &= (T-k)^{-2} \sum_{i=1}^k \left\{ \sum_{t=k+1}^T y_{t-i} \sum_{j=k+1}^{\infty} \beta_j y_{t-j} \right\}^2 \\ &= O(k^{1-2d} T^{2d-1}). \end{aligned}$$

Thus, item 2 of Lemma 3 is established.

Because the order of magnitude of item 3 of Lemma A.2 cannot be greater than that of the sum of the items 1 and 2 of Lemma A.2, item 3 of Lemma 2 follows.

#### A.4. Proof of Theorem 1

We note that

$$\begin{aligned} \|\widehat{\beta} - \beta\| &\leq \left\| (R^{*-1} - R^{-1}) \frac{1}{T-k} M \right\| + \left\| R^{-1} \frac{1}{T-k} \left( \sum_{t=k+1}^T Y_t e_t \right) \right\| \\ &+ \left\| R^{-1} \frac{1}{T-k} \left\{ \sum_{t=k+1}^T Y_t (e_{t,k} - e_t) \right\} \right\| = Z_{1T} + Z_{2T} + Z_{3T} = O(k^{0.5-d} T^{d-0.5}), \end{aligned}$$

because

$$Z_{1T} \leq \|R^{*-1} - R^{-1}\| \left\| \frac{1}{T-k} M \right\| = \begin{cases} O(k^{1.5-d} T^{3d-1.5}) & \text{if } d \in (0.25, 0.5), \\ O(k^{1.5-d} T^{d-1} (\text{LOG } T)^{0.5}) & \text{if } d \in (0, 0.25], \end{cases}$$

$$Z_{2T} \leq \|R^{-1}\| \left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_t e_t \right\| = O(k^{1/2} T^{-1/2}),$$

$$Z_{3T} \leq \|R^{-1}\| \left\| \frac{1}{T-k} \left\{ \sum_{t=k+1}^T Y_t (e_{t,k} - e_t) \right\} \right\| = O(k^{0.5-d} T^{d-0.5})$$

Item 1 of Theorem 1 follows immediately.

To prove item 2 of Theorem 1, we note that the OLS residuals are defined as:

$$\widehat{e}_{t,k} = e_t + \sum_{j=k+1}^{\infty} \beta_j y_{t-j} - \sum_{j=1}^k (\widehat{\beta}_j - \beta_j) y_{t-j}. \quad (\text{A.15})$$

We prove that

$$\mathbf{E} \left( \sum_{j=k+1}^{\infty} \beta_j y_{t-j} \right)^2 \leq \sup_t E(y_{t-j}^2) \left( \sum_{j=k+1}^{\infty} \beta_j \right)^2 = C\gamma_0 k^{-2d}$$

and  $\sum_{j=k+1}^{\infty} \beta_j y_{t-j} = O_p(k^{-d})$ . Furthermore, by Cauchy-Schwarz's Inequality, we obtain

$$\sum_{j=1}^k (\hat{\beta}_j - \beta_j) y_{t-j} \leq \left[ \sum_{j=1}^k (\hat{\beta}_j - \beta_j)^2 \right]^{1/2} \left[ \sum_{j=1}^k y_{t-j}^2 \right]^{1/2} = O(k^{1-d} T^{d-0.5}) \quad (\text{A.16})$$

and it shows that  $k = o(T^{\frac{0.5-d}{1-d}})$  is needed to make  $\sum_{j=1}^k (\hat{\beta}_j - \beta_j) y_{t-j} \xrightarrow{p} 0$ .

Because  $\sum_{j=k+1}^{\infty} \beta_j y_{t-j} = O_p(k^{-d})$  and  $\sum_{j=1}^k (\hat{\beta}_j - \beta_j) y_{t-j} = o_p(1)$ , item 2 of Theorem 1 follows immediately.

Item 3 of Theorem 1 can be easily proved with the help of item 2 of Theorem 1, and the details are omitted.

#### A.4. Proof of Theorem 2

To save the space, some detailed proofs of Theorem 2 are omitted and available to authors.

We define

$$\begin{aligned} \hat{e}_{t,k,1} &= e_{t,1} + \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} - \sum_{j=1}^k (\hat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1}, \\ \hat{e}_{t+l,k,2} &= e_{t+l,2} + \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} - \sum_{j=1}^k (\hat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2}. \end{aligned}$$

Together with equation (2), we know that

$$\left( \sum_{t=k+1}^T \hat{e}_{t,k,1}^2 \right)^{1/2} \left( \sum_{t=k+1}^T \hat{e}_{t+l,k,2}^2 \right)^{1/2} \sim C(T-k)$$

Hence, we obtain

$$\begin{aligned}
\widehat{\rho}_{12}(l) &= \frac{\sum_{t=k+1}^{T-l} \widehat{e}_{t,k,1} \widehat{e}_{t+l,k,2}}{\left(\sum_{t=k+1}^T \widehat{e}_{t,k,1}^2\right)^{1/2} \left(\sum_{t=k+1}^T \widehat{e}_{t+l,k,2}^2\right)^{1/2}} \\
&= \frac{\sum_{t=k+1}^{T-l} e_{t,1} e_{t+l,2}}{\left(\sum_{t=k+1}^T \widehat{e}_{t,1}^2\right)^{1/2} \left(\sum_{t=k+1}^T \widehat{e}_{t,2}^2\right)^{1/2}} \\
&\quad - \frac{C}{T-k} \left( \sum_{t=k+1}^{T-l} e_{t,1} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right) \\
&\quad + \frac{C}{T-k} \left( \sum_{t=k+1}^{T-l} e_{t,1} \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) \right) \\
&\quad + \frac{C}{T-k} \left( \left\{ \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) \right\} \right) \\
&\quad - \frac{C}{T-k} \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \\
&\quad + \frac{C}{T-k} \left( \sum_{t=k+1}^{T-l} e_{2,t+l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \right) \\
&\quad - \frac{C}{T-k} \left( \sum_{t=k+1}^{T-l} e_{2,t+l} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1} \right\} \right) \\
&\quad - \frac{C}{T-k} \sum_{t=k+1}^{T-l} \left( \sum_{j=1}^k (\widehat{\beta}_j - \beta_j) y_{t-j} \left( \sum_{j=k+1}^{\infty} \beta_j y_{t+l-j} \right) \right) \\
&\quad + \frac{C}{T-k} \left( \sum_{t=k+1}^{T-l} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1} \right\} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right)
\end{aligned} \tag{A.17}$$

Before we calculate the asymptotic properties of the terms at the right hand side of (A.19), we note that: for any conformable matrix  $A$  and  $B$ , we have

$$\|AB\|^2 \leq \|A\|^2 \|B\|^2, \tag{A.18}$$

We denote

$$\beta(k+1, \infty) = (\beta_{k+1}, \beta_{k+2}, \dots, \beta_\infty) \quad Y_{t-k} = (y_{t-k-1}, y_{t-k-2}, \dots, y_{t-\infty})'$$

Then, under the independence assumption between  $\hat{e}_{t,k,1}$  and  $\hat{e}_{t+l,k,2}$ ,

$$\begin{aligned} & \left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} e_{t,1} \left\{ \sum_{j=1}^k (\hat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right\|^2 \\ & \leq \left\| \left[ \frac{C}{T-k} \sum_{t=k+1}^{T-l} e_{t,1} Y'_{t+l} \right] \right\|^2 \left\| \hat{\beta}(k) - \beta(k) \right\|^2 = O_p(T^{-1}k) O_p(k^{1-2d} T^{2d-1}) \end{aligned}$$

and

$$\left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} e_{t,1} \left\{ \sum_{j=1}^k (\hat{\beta}(k, 2) - \beta(k, 2)) y_{t+l-j,2} \right\} \right\| = O_p(T^{-0.5} k^{1/2}) O_p(k^{0.5-d} T^{d-0.5}), \quad (\text{A.19})$$

by item 1 of Lemma A.2 and item 1 of Theorem 1. Similarly, we observe that

$$\left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} e_{t+l,2} \left\{ \sum_{j=1}^k (\hat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1} \right\} \right\| = O_p(T^{-0.5} k^{1/2}) O_p(k^{0.5-d} T^{d-0.5}) \quad (\text{A.20})$$

For the third term at the right hand of (A.17), notice that

$$\begin{aligned} \mathbf{E} \left\| \frac{C}{T-k} \left\{ \sum_{t=k+1}^{T-l} e_{t,1} \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) \right\} \right\|^2 & \leq \frac{C^2}{T-k} \gamma_0 \sigma^2 \left( \sum_{j=k+1}^{\infty} \beta_j \right)^2 \\ & = O_p(k^{-2d} (T-k)^{-1}) \end{aligned}$$

and

$$\left\| \frac{C_2}{T-k} \left\{ \sum_{t=k+1}^{T-l} e_{t,1} \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) \right\} \right\| = O_p(k^{-d} (T-k)^{-0.5}) \quad (\text{A.21})$$

Following the preceding reasoning for (A.21), we also have

$$\left\| \frac{C}{T-k} \left\{ \sum_{t=k+1}^{T-l} e_{t+l,2} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \right\} \right\| = O_p(k^{-d} (T-k)^{-0.5}) \quad (\text{A.22})$$

To compute the fourth term at (A.17), we find that

$$\begin{aligned} & \left\{ \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) \right\}^2 \\ & \leq \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right)^2 \left\{ \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \right\}^2 \end{aligned} \quad (\text{A.23})$$

Thus, using the same reasoning for  $A^*$ , we have

$$\begin{aligned} \mathbf{E} & \left\{ \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \right\}^2 = \sum_{t=k+1}^{T-l} \sum_{j=k+1}^{\infty} \beta_{j,1}^2 \gamma_0 \\ & + 2 \sum_{t=k+1}^{T-l} \sum_{j=k+1}^{\infty} \sum_{v=j+1}^{\infty} \beta_{j,1} \beta_{v,1} \gamma_{j-v} + 2 \sum_{t=k+1}^{T-l-1} \sum_{s=t+1}^{T-l} \sum_{j=k+1}^{\infty} \sum_{v=j+1}^{\infty} \beta_{j,1} \beta_{v,1} \mathbf{E}(y_{t-j,1} y_{s-v,1}) \\ & = O((T-k)k^{-2d}) \end{aligned}$$

then

$$\left\{ \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \right\} = O((T-k)^{1/2} k^{-d}) \quad (\text{A.24}).$$

Thus, from (A.23) and (A.24), we have

$$\frac{C}{T-k} \left\{ \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) \right\} = O(T^{-1/2} k^{-2d}) \quad (\text{A.25})$$

For the fifth term at the right hand side of (A.17), by the item 2 of Lemma A.2, we observe

$$\begin{aligned} & \left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right\|^2 \\ & = C^2 \left\| \frac{1}{T-k} \sum_{t=k+1}^{T-l} [\widehat{\beta}(k, 2) - \beta(k, 2)] Y_{t+l,2} Y'_{t-k,1} \beta(k+1, \infty, 1)' \right\|^2 \\ & \leq \left\| \widehat{\beta}(k, 2) - \beta(k, 2) \right\|^2 \left\| \frac{1}{T-k} \sum_{t=k+1}^{T-l} Y_{t+l,2} Y'_{t-k,1} \beta(k+1, \infty, 1)' \right\|^2 \\ & = O(k^{1-2d} T^{2d-1}) O(k^{1-2d} T^{2d-1}) \end{aligned}$$

by item 2 of Lemma A.2 and item 1 of Theorem 1.

Also,

$$\left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} \left( \sum_{j=k+1}^{\infty} \beta_{j,1} y_{t-j,1} \right) \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right\| = O(k^{1-2d} T^{2d-1}) \quad (\text{A.26})$$

Similarly, we prove that

$$\frac{C}{T-k} \sum_{t=k+1}^{T-l} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1} \right\} \left( \sum_{j=k+1}^{\infty} \beta_{j,2} y_{t+l-j,2} \right) = O(k^{1-2d} T^{2d-1}) \quad (\text{A.27})$$

Moreover, for the ninth term at the right hand side of (A.17),

$$\begin{aligned} & \left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1} \right\} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right\|^2 \\ & \leq C^2 \left\| \widehat{\beta}(k) - \beta(k) \right\|^2 \left\| \frac{1}{T-k} \sum_{t=k+1}^{T-l} Y_t Y'_{t+l} \right\|^2 \left\| \widehat{\beta}(k) - \beta(k) \right\|^2 \end{aligned}$$

Moreover, under the independent assumption between  $x_t$  and  $x_{t+l}$  observe that

$$\begin{aligned} \mathbf{E} \left\| \frac{1}{T-k} \sum_{t=k+1}^T Y_{t,1} Y_{t+l,2} \right\|^2 & \leq \left( \frac{1}{T-k} \right)^2 k^2 \max_{1 \leq i, j \leq k} \mathbf{E} \left( \sum_{t=k+1}^T y_{t,i,1} y_{t+l,j,2} \right)^2 \\ & = O(k^2 T^{-1}) \end{aligned}$$

Combining the preceding findings and item 1 of Theorem 1, we prove that

$$\left\| \frac{C}{T-k} \sum_{t=k+1}^{T-l} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,1} - \beta_{j,1}) y_{t-j,1} \right\} \left\{ \sum_{j=1}^k (\widehat{\beta}_{j,2} - \beta_{j,2}) y_{t+l-j,2} \right\} \right\| = O(k^{2-2d} T^{2d-1.5}). \quad (\text{A.28})$$

Consequently, by (A.17), (A.19), (A.20), (A.21), (A.22), (A.25), (A.26), (A.27), (A.28), it can be established that  $\widehat{\rho}_{12}(l) - \rho_l$  is  $o_p(T^{-1/2})$  as follows.

$$\widehat{\rho}_{12}(l) = \frac{\sum_{t=k+1}^{T-l} e_{t,1} e_{t+l,2}}{\left( \sum_{t=k+1}^T e_{t,1}^2 \right)^{1/2} \left( \sum_{t=k+1}^T e_{t,2}^2 \right)^{1/2}} + o_p(T^{-1/2}) = \rho_{12}(l) + o_p(T^{-1/2})$$

Thus, we can interpret the cross-correlation estimates constructed with the OLS residuals behave as those of two white noise series,  $e_{1,t}$  and  $e_{2,t}$ , asymptotically. Furthermore, following equations (1.3) and (1.4) of Haugh (1976), we thus prove Theorem 2.

### A.5. Proof of Theorem 3

To prove Theorem 3, let  $k$  be a positive finite constant, such as  $k = 1, 2, \dots, c$  and  $c$  is a positive finite constant. According to the proof procedure of Theorem 2, Theorem 3 can be established directly. The detailed proof is available to authors.

### A.6. Proof of Theorem 4

Following the theoretical reasoning of Theorem 2, we note that

$$\hat{\rho}_{12}(l) = \frac{\sum_{t=k+1}^{T-l} e_{t,1}e_{t+l,2}}{\left(\sum_{t=k+1}^T e_{t,1}^2\right)^{1/2} \left(\sum_{t=k+1}^T e_{t,2}^2\right)^{1/2}} + o_p(T^{-1/2}) = \rho_{12}(l) + o_p(T^{-1/2}),$$

under the null hypothesis of independence. Hence, the asymptotic power of the two Haugh tests based on the AR approximation of long memory can be obtained directly

$$\lim_{T \rightarrow \infty} Pr(S_M > \chi_{2M+1}^2) = 1 - \Phi(\chi_{2M+1}^2) = 1$$

where  $\Phi$  is the cumulative distribution of  $\chi_{2M+1}^2$ .

## Recent titles

### CORE Discussion Papers

- 2008/11. Pierre PESTIEAU and Uri POSSEN. Prodigality and myopia. Two rationales for social security.
- 2008/12. Tim COELLI, Mathieu LEFEBVRE and Pierre PESTIEAU. Social protection performance in the European Union: comparison and convergence.
- 2008/13. Loran CHOLLETE, Andréas HEINEN and Alfonso VALDESOGO. Modeling international financial returns with a multivariate regime switching copula.
- 2008/14. Filomena GARCIA and Cecilia VERGARI. Compatibility choice in vertically differentiated technologies.
- 2008/15. Juan D. MORENO-TERNERO. Interdependent preferences in the design of equal-opportunity policies.
- 2008/16. Ana MAULEON, Vincent VANNETELBOSCH and Wouter VERGOTE. Von Neumann-Morgenstern farsightedly stable sets in two-sided matching.
- 2008/17. Tanguy ISAAC. Information revelation in markets with pairwise meetings: complete information revelation in dynamic analysis.
- 2008/18. Juan D. MORENO-TERNERO and John E. ROEMER. Axiomatic resource allocation for heterogeneous agents.
- 2008/19. Carlo CAPUANO and Giuseppe DE FEO. Mixed duopoly, privatization and the shadow cost of public funds.
- 2008/20. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Forced saving, redistribution and nonlinear social security schemes.
- 2008/21. Philippe CHEVALIER and Jean-Christophe VAN DEN SCHRIECK. Approximating multiple class queueing models with loss models.
- 2008/22. Pierre PESTIEAU and Uri M. POSSEN. Interaction of defined benefit pension plans and social security.
- 2008/23. Marco MARINUCCI. Optimal ownership in joint ventures with contributions of asymmetric partners.
- 2008/24. Raouf BOUCEKKINE, Natali HRITONENKO and Yuri YATSENKO. Optimal firm behavior under environmental constraints.
- 2008/25. Ana MAULEON, Vincent VANNETELBOSCH and Cecilia VERGARI. Market integration in network industries.
- 2008/26. Leonidas C. KOUTSOUGERAS and Nicholas ZIROS. Decentralization of the core through Nash equilibrium.
- 2008/27. Jean J. GABSZEWICZ, Didier LAUSSEL and Ornella TAROLA. To acquire, or to compete? An entry dilemma.
- 2008/28. Jean-Sébastien TRANCREZ, Philippe CHEVALIER and Pierre SEMAL. Probability masses fitting in the analysis of manufacturing flow lines.
- 2008/29. Marie-Louise LEROUX. Endogenous differential mortality, non monitored effort and optimal non linear taxation.
- 2008/30. Santanu S. DEY and Laurence A. WOLSEY. Two row mixed integer cuts via lifting.
- 2008/31. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Taxing sin goods and subsidizing health care.
- 2008/32. Jean J. GABSZEWICZ, Didier LAUSSEL and Nathalie SONNAC. The TV news scheduling game when the newscaster's face matters.
- 2008/33. Didier LAUSSEL and Joana RESENDE. Does the absence of competition *in* the market foster competition *for* the market? A dynamic approach to aftermarkets.
- 2008/34. Vincent D. BLONDEL and Yurii NESTEROV. Polynomial-time computation of the joint spectral radius for some sets of nonnegative matrices.
- 2008/35. David DE LA CROIX and Clara DELAVALLADE. Democracy, rule of law, corruption incentives and growth.
- 2008/36. Jean J. GABSZEWICZ and Joana RESENDE. Uncertain quality, product variety and price competition.

## Recent titles

### CORE Discussion Papers - continued

- 2008/37. Gregor ZOETTL. On investment decisions in liberalized electricity markets: the impact of price caps at the spot market.
- 2008/38. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Habit formation and labor supply.
- 2008/39. Marie-Louise LEROUX and Grégory PONTIERE. Optimal tax policy and expected longevity: a mean and variance approach.
- 2008/40. Kristian BEHRENS and Pierre M. PICARD. Transportation, freight rates, and economic geography.
- 2008/41. Gregor ZOETTL. Investment decisions in liberalized electricity markets: A framework of peak load pricing with strategic firms.
- 2008/42. Raouf BOUCEKKINE, Rodolphe DESBORDES and H el ene LATZER. How do epidemics induce behavioral changes?
- 2008/43. David DE LA CROIX and Marie VANDER DONCKT. Would empowering women initiate the demographic transition in least-developed countries?
- 2008/44. Geoffrey CARUSO, Dominique PEETERS, Jean CAVAILHES and Mark ROUNSEVELL. Space-time patterns of urban sprawl, a 1D cellular automata and microeconomic approach.
- 2008/45. Taoufik BOUEZMARNI, Jeroen V.K. ROMBOUTS and Abderrahim TAAMOUTI. Asymptotic properties of the Bernstein density copula for dependent data.
- 2008/46. Joe THARAKAN and Jean-Philippe TROPEANO. On the impact of labor market matching on regional disparities.
- 2008/47. Shin-Huei WANG and Cheng HSIAO. An easy test for two stationary long processes being uncorrelated via AR approximations.

### Books

- Y. POCHET and L. WOLSEY (eds.) (2006), *Production planning by mixed integer programming*. New York, Springer-Verlag.
- P. PESTIEAU (ed.) (2006), *The welfare state in the European Union: economic and social perspectives*. Oxford, Oxford University Press.
- H. TULKENS (ed.) (2006), *Public goods, environmental externalities and fiscal competition*. New York, Springer-Verlag.
- V. GINSBURGH and D. THROSBY (eds.) (2006), *Handbook of the economics of art and culture*. Amsterdam, Elsevier.
- J. GABSZEWICZ (ed.) (2006), *La diff erenciation des produits*. Paris, La d ecouverte.
- L. BAUWENS, W. POHLMEIER and D. VEREDAS (eds.) (2008), *High frequency financial econometrics: recent developments*. Heidelberg, Physica-Verlag.
- P. VAN HENTENRYCKE and L. WOLSEY (eds.) (2007), *Integration of AI and OR techniques in constraint programming for combinatorial optimization problems*. Berlin, Springer.

### CORE Lecture Series

- C. GOURI EROUX and A. MONFORT (1995), *Simulation Based Econometric Methods*.
- A. RUBINSTEIN (1996), *Lectures on Modeling Bounded Rationality*.
- J. RENEGAR (1999), *A Mathematical View of Interior-Point Methods in Convex Optimization*.
- B.D. BERNHEIM and M.D. WHINSTON (1999), *Anticompetitive Exclusion and Foreclosure Through Vertical Agreements*.
- D. BIENSTOCK (2001), *Potential function methods for approximately solving linear programming problems: theory and practice*.
- R. AMIR (2002), *Supermodularity and complementarity in economics*.
- R. WEISMANTEL (2006), *Lectures on mixed nonlinear programming*.