東海管理評論
民國八十八年七月，第一卷，第一期，111-128

Detection of Common Long-Range Dependence Component for Bivariate Time Series

Nan-Jung Hsu* Ying-Sen Chen[0]

Abstact

In this paper, we proposed a new method to identifying common long-range dependent component in a bivariate time series. A common long-range dependent component exists if individual series are both long-range dependent but there exists a particular linear combination of the process which does not have the long-memory property. We first find the linear combination by the two-stage least squares procedure and then test the long-memory property for the transformed data using the method proposed by Geweke and Porter-Hudak (1983). The performance of the proposed test is investigated via Monte Carlo simulation and compared with the previous method based on the canonical correlation analysis. The wind speed data are used to illustrate the test procedure.

Keywords: Canonical correlation; common components; long memory; fractionally integrated ARMA process; two-stage least squares.

1.Introduction

Many multivariate time series data are collected on a cross-sectional basis. For

* Department of Statistics, Tunghai University

instance, many economic indices of a country that are subject to the same source of variations. Therefore, it is natural to think that the dynamic structures of those individual processes are similar and might share the common features. In the time series literature, the common factor is usually modeled as a non-stationary component and this corresponds to the case of cointegration (Granger, 1981; Engle and Granger, 1987). The cointegrating relationship can be interpreted as the existence of a long run equilibrium between series which is stationary with finite variance, even though the original individual series are non-stationary with infinite variance. Cheung and Lai (1993) considered the purchasing power parity as fractionally cointegrated in which the original process is integrated of order one but a certain linear combination of the process is integrated of order $d<1$. In this paper, we consider the common component as a stationary process with long-range dependence. Many empirical studies have found that many processes of interest in economics and finance exhibit long-range dependence such as interest rate, exchange rate, returns volatilities. Roughly speaking, the long-range dependence is that the correlation between lagged observations decays hyperbolically as the time lag increases. This phenomenon also exists in many other areas, such as the hydrology and the geophysics. See Beran (1994) and Baillie (1996) for more references.

Ray and Tsay (1998) have developed a test for determining whether long-range dependent (CLRD) behavior observed in individual series is common to a cross-section of series. Their test is constructed based on the canonical correlation analysis between the original process and its lagged values which is similar to the approaches of Box and Tiao (1977) and Tiao and Tsay (1989) for detecting the common trend components. They also found the empirical evidence of common long-range dependence in inflation rates for Switzerland and Austria and in stock returns volatility for companies having similar annual sales. However, the implementation of this test requires two pre-selected parameters which are related to the covariance structure of the individual series. In general, there is no good guideline about how to choose the pre-selected parameters. Therefore, the test statistic is varying subject to the different choices. In this paper, we proposed a new test for detecting CLRD. The key idea of our test procedure is that there exists a linear combination of the multivariate process which does not preserve the long-memory property. Therefore, to determine the existence of CLRD component is equivalent to find a particular linear combination of the process which does not have long-memory property. We first estimate the particular linear combination using the two-stage least squares procedure (2SLS) and then test the long-memory property based on the corresponding transformed series using the method by Geweke and Porter-Hudak (1983). This method is referred to as the GPH method hereafter. Our approach is completely different from the ``two-step" approach by Engle and Granger (1987) for testing the co-integration system since they use the ``ordinary least squares" method to estimate the co-integrating vector in the first step and then plug in the estimated co-integrating vector for testing the error correction structure in the second step. In their approach, the ordinary least squared estimator of the co-integrating vector is consistent since the individual series is non-stationary. However, the same estimator won't be consistent in the CLRD component model since the individual series is stationary.

The rest of the paper is organized as follows: in Section 2, we define the CLRD model and describe its properties. In Section 3, the test procedures for detecting CLRD are introduced, including the method proposed by Ray and Tsay (1998) and our approach. In this study, we particularly concentrate on the bivariate fractionally integrated Gaussian model with additive noise process. The finite-sample performance of the proposed test associated with this class of processes is presented in Section 4 via Monte Carlo simulation. In Section 5, the methodology is used to identify the CLRD component for the wind speed data in Ireland. Section 6 is the conclusion.

2.The Bivariate Common Long-Range Dependent Component Model

Without loss of generality, we consider a zero-mean process since the mean level is not relevant to the test procedure. Let $\{y_t = (y_{1t}, y_{2t})'\}$ be a zero-mean bivariate process with one common long-range dependence component satisfying

$$y_t = \Lambda x_t + \varepsilon_t \tag{1}$$

where $\Lambda = (1, \lambda)'$ is a 2×1 matrix of constants, $\{x_t\}$ is a univariate long-memory process with long-memory parameter d and ε_t is a bivariate short-range dependent disturbance. The first component of Λ is restricted to be one to ensure that the model is identifiable (Harvey, 1989). Clearly, the individual series $\{y_{1t}\}$ and $\{y_{2t}\}$ share the same component $\{x_t\}$, therefore both series are long-range dependent with the same long-memory parameter d. However their linear combination $r_t = y_{2t} - \lambda y_{1t}$ becomes short-range dependent since the common component is canceled out.

The most commonly used long-memory process is the fractionally integrated ARMA process (Granger and Joyeux, 1980; Hosking, 1981), which satisfies the following equation

$$\Phi(B)(1-B)^d x_t = \Theta(B)\eta_t \tag{2}$$

where $d \in (-\frac{1}{2}, \frac{1}{2})$ is the long-memory parameter, $\{\eta_t\}$ are iid from $N(0, \sigma_\eta^2)$ and independent of $\{x_t\}$, $\Phi(z)$ and $\Theta(z)$ are polynomials specifying the short-range dependence. A special characteristics of long-memory process is that the autocovariance function $\gamma_x(h)$ decays very slowly with a hyperbolic rate; that is $\gamma_x(h) \sim h^{2d-1}$ as h approaches infinity. As the equivalent result, the corresponding spectral density $f_x(\omega)$ is unbounded at the zero frequency and behaves like $f_x(\omega) \sim \omega^{-2d}$ as ω approaches zero.

It is reasonable to consider a common component model as opposed to a full vector long-range dependent model for many reasons. First, the CLRD model is more parsimonious than a model that assumes separate long-memory components for each series. The corresponding statistical inference may be simplified because of the dimension reduction. Moreover, the longer term dynamics between series are preserved for forecasting.

3.The Test for Detecting CLRD Component

In this section, two approaches are introduced for detecting the existence of CLRD component—the test proposed by Ray and Tsay in Section 3.1 and the test based on the 2SLS procedure and the GPH method in Section 3.2.

3.1 The Canonical Correlation Approach

First, we define some notations. Let $Y_{h,j,t} = (y_{t-j}^{'}, y_{t-j-1}^{'}, \ldots, y_{t-j-h+1}^{'})^{'}$ where $h > 0$, $j > 0$. It represents a collection of h past variables after lagging j times. Define

$$\Gamma_y(h) = \mathrm{cov}(y_t, y_{t+h}),$$

$$\Gamma_\varepsilon(h) = \mathrm{cov}(\varepsilon_t, \varepsilon_{t+h}),$$

$$G(h,j) = \mathrm{cov}(y_t, Y_{h,j,t}).$$

Under the model defined in (1) – (2), we have $\Gamma_y(h) = \Lambda\gamma_x(h)\Lambda^{'} + \Gamma_\varepsilon(h)$. Since Γ_ε decays at a faster rate than $\gamma_x(h)$, $\Gamma_y(h) \approx \Lambda\gamma_x(h)\Lambda^{'}$ for h large enough. Similarly, we have

$$G(h,j) \approx \Lambda[\gamma_x(j), \gamma_x(j+1), \ldots, \gamma_x(j+h-1)]\Lambda^{'},$$

for a sufficiently large j. Fixed h and j, the canonical correlations between y_t and $Y_{h,j,t}$ are the squared eigenvalues of the matrix

$$A(h,j)=[\mathrm{var}(y_t)]^{-1}G(h,j)[\mathrm{var}(Y_{h,j,t})]^{-1}[G(h,j)]' \equiv \Sigma_{yy}^{-1}\Sigma_{yY}\Sigma_{YY}^{-1}\Sigma_{Yy} \quad (3)$$

Assume $0 \le \rho_1^2 \le \rho_2^2 \le 1$ are the ordered eigenvalues of $A(h,j)$. Ray and Tsay proposed the following test hypothesis

$$H_0 : \rho_1^2 = 0 \text{ against } H_1 : \rho_1^2 > 0,$$

with the test statistic

$$T_n = -(n-h)\ln(1-\hat{\rho}_1^2) \to \chi_{2h-1}^2 \quad (4)$$

under H_0 and some regularity conditions, where $\hat{\rho}_1^2$ is the smaller eigenvalue of the matrix $\hat{A}(h,j)=\hat{\Sigma}_{yy}^{-1}\hat{\Sigma}_{yY}\hat{\Sigma}_{YY}^{-1}\hat{\Sigma}_{Yy}$ which is the product of the sample variances and covariances based on the observed data. The existence of a CLRD component is concluded if the test is not rejected. The similar test procedure is also hold for k-dimensional process with r CLRD component for $k>2$ and $k>r$. Besides, this test procedure is also applicable for the nonstationary case with $d \ge 0.5$.

There are two pre-selected parameters j and h in the test procedure. The value of j should be chosen so that Γ_ε does die out after lag j. The value of h should be chosen so that the block of series $Y_{h,j,t}$ can capture the most of the dynamic structure of $\{y_t\}$. However, it is a hard problem for choosing the values of j and h in real applications.

3.2 The GPH Approach

Assuming the model defined in Equations (1)－(2) holds which exists a CLRD component, then $\{r_t = y_{2t} - \lambda y_{1t}\}$ is short-range dependent. The same model can be considered as an errors-in-variables model. It is well known that the ordinary least squared estimator of λ is biased. Instead, a consistent estimator can be obtained by using the instrumental variables. Let $Y_i = (y_{i1}, y_{i2}, \ldots, y_{in})'$ and $E_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in})'$ be the vectors of the individual series and the noise process for the i-th series, Z be the vector of instrumental variable for Y_1. By using the 2SLS procedure, we have a consistent estimator of λ

$$\hat{\lambda} = (\hat{Y}_1'\hat{Y}_1)^{-1}\hat{Y}_1'Y_2 \tag{5}$$

where $\hat{Y}_1 = Z(Z'Z)^{-1}Z'Y_1$.

The instrumental variable Z can be chosen arbitrarily but should satisfy the following two conditions: (i) $\lim(Z'Y_1/n)$ is a non-zero constant; (ii) $\lim(Z'E_2/n) = 0$. The natural choice of instrument would be the lagged variables $\{y_{1,t-j}\}$; that is to set $Z = (0,...,0, y_{11}, y_{12},..., y_{1,n-j})'$ for some j. Fixed j, the value of λ is estimated by Equation (5) and the corresponding transformed data $\hat{r}_t = y_{2t} - \hat{\lambda}y_{1t}$ can be obtained. Then, the GPH method is applied to $\{\hat{r}_t\}$ for determining its memory property. Simply speaking, the GPH method is an estimation procedure in which the property of the spectral density around frequency zero are used. The estimate of the long-memory parameter d is obtained through a regression of the log periodogram ordinates on the log frequencies

$$\ln I(\omega_k) = \beta_0 + \beta_1 \ln[4\sin^2(\omega_k/2)] + u_k, \qquad k = 1,2,...,m,$$

where $I(\omega_k)$ is the periodogram ordinate at the k-th Fourier frequency for $\{\hat{r}_t\}$, $m = [n^{\alpha}]$ with $0 < \alpha < 1$. Geweke and Porter-Hudak (1983) have showed that the least squares estimator of β_1 provides a consistent estimator of $-d$. Therefore, to test $H_0: d \le 0$ is equivalent to test $H_0: \beta_1 \ge 0$ based on the t statistic of the regression coefficient. Although we use the estimated $\{\hat{r}_t\}$ instead of the real interested but unobserved variables $\{r_t\}$ in the testing procedure, the GPH test is still valid for making inference about $\{r_t\}$ because of the consistency of $\hat{\lambda}$. Initially, the GPH test was built for the fractionally integrated process but the same test is also valid for various long-memory processes satisfying $f(\omega) \sim |1-e^{iw}|^{-2d} f^*(\omega)$ with a bounded function f^* as ω approaches to zero.

In this testing procedure, the instrument Z should satisfy

$$\frac{1}{n}Z'Y_1 = \frac{1}{n}\sum_{t=j+1}^{n} y_{1t}y_{1,t-j} \rightarrow \Gamma_y(j),$$

$$\frac{1}{n}Z'E_2 = \frac{1}{n}\sum_{t=j+1}^{n} \varepsilon_{2t}y_{1,t-j} \rightarrow \text{cov}(\varepsilon_{2t}, \varepsilon_{1,t-j}),$$

in which the matrix $\Gamma_y(j)$ is always a non-zero scalar for arbitrary fixed j. But j should be chosen carefully so that $\text{cov}(\varepsilon_{1,t-j}, \varepsilon_{2t}) = 0$; that is the covariance between two noise components should die out after lagging j times. As we can see that j here plays the same role as that in Ray and Tsay's method. There is a guideline on finding a suitable value of j by looking at the sample correlation of z_t and $\hat{r}_t$ under a fixed j. A small value of the sample correlation could indicate that the choice of j would be appropriate.

4.Simulation Results

To investigate the performance of the proposed test for detecting CLRD component in the bivariate Gaussian case, a finite-sample simulation is conducted. We consider a bivariate CLRD model defined in Equations (1) and (2). The design parameters in the simulation study include the long-memory parameter d, the signal-to-noise ratio (the ratio of the variances of the long-memory component to the short-memory component) and those specified for the short-range dependent structure. We consider the following setups. The long-memory parameter is examined for the case with $d = 0.3$ and $d = 0.4$. Three models for $\{\varepsilon_t\}$ are studied, which are (i) $\{\varepsilon_t\}$ is a white noise process; (ii) $\{\varepsilon_t\}$ is a MA(1); (iii) $\{\varepsilon_t\}$ is an AR(1). The signal-to-noise ratio (SNR) is set at either one or five. The value of λ is set at 0.5 in all cases. For each case, 500 realizations are simulated independently, each having sample size $n = 1000$.

Both tests described in Section 3 are applied to each realization for testing the existence of CLRD component. In Ray and Tsay's procedure, the values of j and h are both set to be five which are suggested by the authors. The value of m is set to be

$[n^{0.6}] = 63$ in the GPH procedure. First, the value of λ is estimated based on both the canonical correlation approach (through the eigenvector corresponding to the smallest eigenvalue) and the 2SLS procedure. The bias and the root mean square error (RMSE) of the estimator of λ for each process are shown in Tables 1—3. The actual sizes of both tests are also recorded and compared with the nominal size $\alpha = 0.05$. The results are also presented in Table 1-3.

First, we consider $\{\varepsilon_t\}$ being iid from $N(0,\Sigma)$ with a diagonal matrix Σ. The results are summarized in Table 1. The estimator of λ based on the 2SLS procedure performs better than the estimator in the other approach in terms of the smaller RMSE. The actual sizes of both tests are close to their nominal size 0.05. The performance of $\hat{\lambda}$ is slightly improved as d increases which is the case exhibiting stronger long-range dependence. Besides, both estimators have better performance as the SNR increases. However, there is no clear difference in the actual size when the SNR or d changes. Secondly, we consider $\{\varepsilon_t\}$ being a MA(1) process satisfying $\varepsilon_t = (1+\Theta B)u_t$ where $\{u_t\}$ are iid $N(0,\sigma_u^2 I)$ and Θ is a diagonal matrix with elements 0.5 and 0.8. The results are summarized in Table 2. Most of the results are similar to the previous case. However, the actual size based on Ray and Tsay's method seems to be smaller than the nominal size. That is, the corresponding test tends to reject the hypothesis of existing CLRD component not as often as it should be. Then, we consider the AR(1) process satisfying $(1-\Phi B)\varepsilon_t = u_t$ where $\{u_t\}$ are iid $N(0,\sigma_u^2 I)$ and Φ is a diagonal matrix with elements 0.5 and 0.8. The results are summarized in Table 3. For AR cases, most of results are similar to the previous two cases except the RMSE of 2SLS estimator is much smaller than that of the other estimator.

The second part of the simulation study concentrates on examining the power of both tests. We consider the following model

$$y_t = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} x_t + \varepsilon_t \tag{6}$$

where $x_t = (x_{1t}, x_{2t})'$ in which each component x_{it} satisfies $(1-B)^{d_i} x_{it} = \eta_{it}$ with

$\text{var}(x_{it}) = 1$; $\{x_{1t}\}$ is independent of $\{x_{2t}\}$, the variance of each disturbance component is also one. We examined the power of the tests under the model in (6) for several situations: (i) $d_1 = d_2 = 0.4$ or $d_1 = 0.4$,; $d_2 = 0.2$ (ii), $\lambda = 0.5$ 1 or 2; (iii) $\{\varepsilon_t\}$ is white noise, MA(1), MA(3), or AR(1). Under (6), the asymptotic distribution of the test statistic in Ray and Tsay's procedure is not the same as that in the model (1)--(2), therefore the theoretical power is unknown. But the t-test in GPH procedure is still valid, so that the power is exactly 0.95 under the given size $\alpha = 0.05$. The empirical results of the power for both procedures are summarized in Table 4. The empirical powers of both tests are very high. There is no clear evidence that one test is superior to the other in terms of the power.

In this simulation study, the long-memory parameters are set to be 0.3 and 0.4. For the case with larger d which is encountered in most of empirical studies, the performance of both tests is slightly improved. However, the SNR is the important factor to affect the performance but not the value of d .

5.Application

We apply our test procedure to determine the existence of CLRD component for the wind speed data in Ireland which were first studied by Haslett and Raftery (1989). The data are hourly wind speeds at 12 meteorological stations during 1961-1978; each series contain 6570 observations. The data are adjusted for the seasonal effects by subtracting the smoothed daily average over all years and stations. We examined the memory property for each of the 12 seasonally adjusted series using GPH method and found that all of them are long-range dependent. The estimated long-memory parameters are reported in Table 5. The long-memory characteristics at Birr, Dublin, Claremorris and Clones are most significant and some of them are similar. These four stations are all located in central Ireland; see the map in Haslett and Raftery (1989). We apply both tests to each pair $(k = 2)$ of these four stations and found that none of them can be reduced to only one CLRD component. In the southeast area, the wind speeds at Roche's Pt., Rosslare and Kilkenny have smaller but similar long-memory parameters. Since these three stations are located closely, we are interested to examine if they share the same source of long-memory process. Both tests are applied and the results are summarized in Table 6. Both tests cannot reject the existence of only one CLRD component for wind speeds at Roche's Pt. and Rosslare. In fact, these two stations are very close and both nearby the Ocean therefore it seems reasonable to have the common structure. For the other bivariate processes, two test results are different. Our test result indicates that there is another source of variation with long-memory property but the corresponding evidence does not show up in Ray and Tsay's method.

6.Conclusions

We have proposed a new test for detecting the CLRD component for a bivariate long-memory process. The method is very easy to implement and its finite-sample performance is competitive with the previous method by Ray and Tsay. In addition, the estimation for λ can be made accurately at the same time in the proposed procedure which can not be achieved by the previous method. The proposed test can be extended for higher-dimensional process under the similar procedures but with sequential steps. For further research, we will derive our test procedures for the more general case which is to detect r CLRD components for the k-dimensional model with $r < k$.

Table 1. The actual sizes of the Ray and Tsay's test and the GPH test for detecting the existence of CLRD component for bivariate fractionally integrated Gaussian process with uncorrelated noise ($n = 1000$, 500 replications).

SNR=1				
Long-memory Parameter	Test Procedure	Estimation for λ		Actual Size of Test
		Bias	RMSE	
$d = 0.3$	Ray & Tsay	0.044	0.378	0.040
	2SLS & GPH	0.015	0.094	0.048
$d = 0.4$	Ray & Tsay	0.012	0.118	0.042
	2SLS & GPH	0.005	0.057	0.046
SNR=5				
Long-memory Parameter	Test Procedure	Estimation for λ		Actual Size of Test
		Bias	RMSE	
$d = 0.3$	Ray & Tsay	0.018	0.141	0.044
	2SLS & GPH	0.003	0.028	0.048
$d = 0.4$	Ray & Tsay	0.001	0.042	0.058
	2SLS & GPH	0.001	0.020	0.042

Table 2. The actual sizes of the Ray and Tsay's test and the GPH test for detecting the existence of CLRD component for bivariate fractionally integrated Gaussian process with MA(1) noise ($n = 1000$, 500 replications).

SNR=1				
Long-memory Parameter	Test Procedure	Estimation for λ		Actual Size of Test
		Bias	RMSE	
$d = 0.3$	Ray & Tsay	-0.029	0.534	0.028
	2SLS & GPH	0.046	0.407	0.054
$d = 0.4$	Ray & Tsay	0.014	0.121	0.032
	2SLS & GPH	0.005	0.092	0.058
SNR=5				
Long-memory Parameter	Test Procedure	Estimation for λ		Actual Size of Test
		Bias	RMSE	
$d = 0.3$	Ray & Tsay	0.008	0.091	0.034
	2SLS & GPH	0.001	0.020	0.042
$d = 0.4$	Ray & Tsay	0.002	0.044	0.052
	2SLS & GPH	0.002	0.030	0.048

Table 3. The actual sizes of the Ray and Tsay's test and the GPH test for detecting the existence of CLRD component for bivariate fractionally integrated Gaussian process with AR(1) noise ($n = 1000$, 500 replications).

SNR=1				
Long-memory Parameter	Test Procedure	Estimation for λ		Actual Size Of Test
		Bias	RMSE	
$d = 0.3$	Ray & Tsay	0.107	1.313	0.036
	2SLS & GPH	0.006	0.079	0.064
$d = 0.4$	Ray & Tsay	0.009	0.111	0.052
	2SLS & GPH	0.003	0.064	0.058
SNR=5				
Long-memory Parameter	Test Procedure	Estimation for λ		Actual Size Of Test
		Bias	RMSE	
$d = 0.3$	Ray & Tsay	0.026	0.369	0.040
	2SLS & GPH	0.001	0.026	0.058
$d = 0.4$	Ray & Tsay	0.001	0.039	0.058
	2SLS & GPH	0.001	0.019	0.052

Table 4. The actual power of the Ray and Tsay's test and the GPH test for detecting the existence of CLRD component for bivariate fractionally integrated Gaussian process with additive noise ($n = 1000$, 500 replications).

Long-memory Parameter	Test Procedure	WN			MA(1)	MA(3)	AR(1)
		$\lambda = 0.5$	$\lambda = 1$	$\lambda = 2$	$\lambda = 0.5$	$\lambda = 0.5$	$\lambda = 0.5$
$d_1 = 0.4$	Ray & Tsay	0.946	0.952	0.944	0.952	0.954	0.966
$d_2 = 0.4$	2SLS & GPH	0.938	0.964	0.944	0.956	0.948	0.946
$d_1 = 0.4$	Ray & Tsay	0.946	0.940	0.956	0.972	0.958	0.972
$d_2 = 0.2$	2SLS & GPH	0.932	0.946	0.962	0.958	0.958	0.946

Table 5. The GPH estimates of the long-memory parameter for the seasonally adjusted wind speed data in Ireland.

Station	$\hat{d}_{GPH}$
Roche's Pt.	0.124
Valentia	0.148
Rosslare	0.125
Kilkenny	0.165
Shannon	0.163
Birr	0.255
Dublin	0.229
Claremorris	0.308
Mullingar	0.137
Clones	0.230
Belmullet	0.106
Main Head	0.172

Table 6. The test statistics and p-values (in parentheses) for examining CLRD component for the seasonally adjusted wind speeds at Roche's Pt., Rosslare and Kilkenny.

Test Procedure	Roche's Pt. vs. Rosslare	Kilkenny vs. Roche's Pt.	Kilkenny vs. Rosslare
Ray & Tsay	T=4.33 (0.888)	T=8.34 (0.500)	T=7.32 (0.604)
2SLS & GPH	$\hat{d}_2 = 0.063$ (0.134)	$\hat{d}_2 = 0.124$ (0.036)	$\hat{d}_2 = 0.124$ (0.040)

References

Baillie, R. (1996), Long memory processes and fractional integration in econometrics, Journal of Econometrics, 73, 5-59.

Beran, J. (1994), *Statistics for long memory processes*, Chapman and Hall, New York.

Box, G.E.P. and Tiao, G.C. (1977), A canonical analysis of multiple time series, *Biometrika*, 64, 355-365.

Cheung, Y.-W. and Lai, K.S. (1993), A fractional cointegration analysis of purchasing power parity, *Journal of the Business and Economic Statistics*, 11, 103-112.

Engle, R. and Granger, C.W.J. (1987), Co-integration and error correction: representation, estimation and testing, *Econometrica*, 55, 251-276.

Geweke, J.F. and Porter-Hudak, S. (1983), The estimation and application of long memory time series models, *Journal of Time Series Analysis*, 4, 221-238.

Granger, C.W.J. and Joyeux, R. (1980), An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, 15-29.

Granger, C.W.J. (1981), Some properties of time series data and their use in econometric model specification, *Journal of Econometrics*, 16, 121-130.

Harvey, A.C. (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, New York.

Haslett, J. and Raftery, A.E. (1989), Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource, *Applied Statistics*, 38, 1-21.

Hosking, J.R.M. (1981), Fractional differencing, *Biometrika*, 68, 165-176.

Ray, B.K. and Tsay, R.S. (1998), Testing for common long-range dependent components in a vector time series, preprint, U. Chicago (available at http://www.math.njit.edu/~borayx/clrdr.ps).

Robinson, P. (1994), Semiparametric analysis of long memory time series, *Annals of Statistics*, 22, 515-539.

Tiao, G.C. and Tsay, R.S. (1989), Model specification in multivariate time series, *Journal of Royal Statistical Society Series B*, 51, 157-213.

Acknowledgments

This work was supported by National Science Council under Grant NSC# 88-2118-M-029-003.

多變量時間數列之共同長期相關成份的檢測

徐南蓉* 陳英森

摘 要

在本篇文章中，我們提倡一種新的方法去辨識多變量時間數列的共同長期相關成份。如果原始數列間個別均具有長期相關，但是卻存在著特定的線性組合不具有長期相關的性質，則我們說在此數列間存在著共同長期相關的成份。首先，我們利用兩階段最小平方法去尋找這種數列間的線性組合，再利用 Geweke 與 Porter-Hudak 的方法對轉換過的資料做長期相關性質的檢定。藉由模擬試驗評估此新方法的表現並與先前 Ray 和 Tsay 所提倡的典型相關檢定法做比較。結果顯示在檢測多變量時間數列之共同長期相關成份的個數上，此新方法比典型相關檢定法的正確率高且對於因子荷載(factor loading)矩陣具有較佳的估計值。在實證分析上，我們以愛爾蘭的風速資料為例，說明整個檢定的流程。

關鍵詞：典型相關；共同成份；長記憶；分數型整合自我迴歸移動平均過程；兩階段最小平方法。

* 東海大學統計學系