Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Analysis of Panel Vector Error Correction Models Using Maximum Likelihood, the Bootstrap, and Canonical-Correlation Estimators

Richard Anderson
Hailong Qian
and
Robert Rasche

Working Paper 2006-050A
http://research.stlouisfed.org/wp/2006/2006-050.pdf

August 2006

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Analysis of Panel Vector Error Correction Models Using Maximum Likelihood, the Bootstrap, and Canonical-Correlation Estimators

Richard Anderson*

Hailong Qian**

Robert Rasche***

This Version: August 23, 2006

Abtract

In this paper, we examine the use of Box-Tiao's (1977) canonical correlation method as an alternative to likelihood-based inferences for vector error-correction models. It is now well-known that testing of cointegration ranks based on Johansen's (1995) ML-based method suffers from severe small sample size distortions. Furthermore, the distributions of empirical economic and financial time series tend to display fat tails, heteroskedasticity and skewness that are inconsistent with the usual distributional assumptions of likelihood-based approach. The testing statistic based on Box-Tiao's canonical correlations shows promise as an alternative to Johansen's ML-based approach for testing of cointegration rank in VECM models.

Keywords: panel cointegration, bootstrap tests, canonical correlation

JEL Codes: C13, C14, C15, C33

* Economist and vice president, Research Department, Federal Reserve Bank of St. Louis. Contact: Anderson@stls.frb.org
**Associate professor, Department of Economics, Saint Louis University, and Research Fellow, Research Department, Federal Reserve Bank of St. Louis. Contact: qianh@slu.edu
*** Senior vice president and director of research, Research Department, Federal Reserve Bank of St. Louis. Contact: rasche@stls.frb.org

Views expressed herein are solely the authors, and are not necessarily those of the Federal Reserve Bank of St. Louis, the Board of Governors of the Federal Reserve System, nor their staffs. We thank many colleagues for comments but hold none responsible for any foolish assertions, methods or conclusions in this analysis.

1. Introduction

Inference regarding the number of long-run equilibrium relationships (that is, the cointegration rank) among a set of economic, financial or social variables is most-often based on maximum likelihood (ML) estimation and related asymptotic distributions, perhaps with a small-sample Bartlett correction. Among, likelihood ratio tests, this has been shown to have the best statistical properties; see for example, Johansen (1995, 2000, 2002), Philips (1995), Stock and Watson (1988). However, it is now well-known via simulation studies that the asymptotic distributions of ML-based testing statistics for cointegration ranks are not good approximations to the true distributions of the testing statistics when the sample size is small to moderate; see for example, Toda (1995), Jacobson (1995), and Haug (1996, 2002).[1]

Because Johansen's maximum likelihood approach is based on the assumption that the true data generating process (DGP) is independent and identically distributed (*i.i.d.*) multivariate normal, it is of interest to explore alternative procedures for testing cointegration ranks and estimation of vector error-correction (VEC) models that are robust to departures from these assumptions. It is also well known, for example, that the distributions of economic and financial data often have fat tails, heteroscedasticity, and skewness. A significant literature has arisen focused on residual-based tests of cointegration, both in univariate and panel models.[2] In these studies, inference regarding cointegration is conducted using residual-based tests via a mapping between the "fit" of a residual-based regression and the presence of cointegration. The intuition in this study is similar, as we explain below. In the presence of nonstationarity, lagged values of a time series should have predictive power for future values; in the presence of stationarity, they will not.

Accurate inference regarding the cointegrating rank in multivariate models is important. If the CI rank is incorrectly inferred due to large size distortions and/or low power of ML-based CI rank test statistics, the long-run coefficient matrix of a vector error-correction model (VECM) is misspecified. This in turn results in an incorrect estimation of the number of common stochastic trends of the system and subsequently causes

[1] Johansen (2000, 2002) furnish additional references.
[2] McCoskey and Kao (2001), Westerlund (2005), and references therein.

erroneous estimates of short-run coefficients. This misspecification of CI rank will have serious consequences for empirical applications, especially for applications to macroeconomic models that prescribe policy recommendations.

To increase power and reduce small-sample size distortions, nonstationary panel data models have recently become very popular; see, for example, Pedroni (1996, 2004), McCoskey and Kao (2001), Kao (1999), Banerjee (1999) Banerjee et al. (2004), and Kao and Chiang (2000). However, the current literature on panel cointegration tests and estimation usually assumes that the number of cross-sectional units is large and does not allow for (i) cross-sectional dependence in the error terms, (ii) the interaction of short-run dynamics between cross-sections, (iii) the difference in cointegration ranks across cross-sections, or (iv) the possibility that long-run equilibrium relationships exist between different cross-sections (hereafter referred to as between-cointegration). If any of these four possibilities holds, the conclusions drawn from the existing panel cointegration literature will be likely misleading and erroneous. [3]

To be more specific about the weakness of the existing panel cointegration literature, let us consider the popular example of testing for purchasing power parity (PPP) among G-7 economies. For this special example, (i) if G-7 economies are affected by the same international economic, financial and political shocks, then we expect that the error terms of different regression equations are contemporaneously correlated. (ii) If temporary changes in trading partners' domestic prices and exchange rates also affect each other in the short-run, then a panel vector error-correction model should also allow for the interaction of short-run dynamics across cross-sections. (iii) Since different countries adopt different monetary and fiscal policies, it is also very natural to allow for different number of cointegration (or long-run equilibrium) relationships among the variables for a given country. Finally, (iv) if we use U.S. as the reference country, then in the regressions of the logarithm of a domestic price on exchange rate (measured as the domestic price per US dollar) and the logarithm of US price, US price appears in every regression and is obviously (trivially) cointegrated with itself across different cross-sectional regressions. In

---

[3] A practice recently introduced in the residual-based testing literature is the removal of "common time effects" as a substitute for allowing cointegration across panel units. Essentially, this entails substracting from each series a mean value calculated across all panels members at each date. See Westerlund (2005) for example.

fact, when the nominal U.S. price is integrated of order one (i.e. I(1)), it acts like a common stochastic trend in the panel regressions. O'Connel (1999) showed through simulations that ignoring the cross-sectional dependence in the error terms can cause large size distortions and significant loss of power for existing panel unit root tests. However, no paper published so far has examined the effects on the size and power of panel cointegration tests when there is cross-sectional interaction of short-run dynamics, cross-sectional differences in cointegration rank, or cross-sectional cointegration.

The current paper has three main objectives. First, we seek to relax those restrictive assumptions that are routinely made in the existing panel cointegration literature. Specifically, we will explicitly allow for cross-sectional dependence among shocks (i.e. model disturbances), cross-sectional interactions in short-run dynamics, differences in cointegration rank across cross-sectional units, as well as the existence of long-run equilibrium relationships between different cross-sections. Second, in order to relax the distributional assumptions of Johansen's (1995) ML-based approach, we propose using Box and Tiao's (1977) canonical correlation (CC) analysis to test for cointegration rank and estimate cointegration vectors. Box and Tiao's CC-based inferences and estimators of long-run parameters do not require any distributional assumptions of the data generating process and are found to have better distributional properties; see Bewley, Orden, Yang and Fisher (1994) and Bewley and Yang (1995). Third, since we do not make any distributional assumptions of the true DGP (except the usual regularity conditions about the existence of relevant moments) nor do we assume that the sample available is large, we use a bootstrap method to find the data-dependent and empirically correct finite-sample distributions for our cointegration rank tests and parameter estimators.[4]

The paper is organized as follows. In the next section, we describe the restricted panel VEC models commonly used by the current literature on panel cointegration and then present the unrestricted panel VEC model considered by the current paper (essentially, that of Larsson and Lyhagen, 1999). In Section 3, we motivate the value of an unrestricted panel VEC model specification via Monte Carlo simulations of the size and power properties of a residual-based panel cointegration test statistic. Section 4 introduces

[4] E.g. Goncalves and White (2004).

the Box and Tiao's (1977) canonical correlation statistic and its extension to testing for cointegration rank. Section 5 provides an empirical application using a panel VEC model for the determination of M1 velocities in U.S. and Canada. We conclude in the last section with some comments.

2. Panel Vector Error-Correction Models

Let $y_{it} = (y_{it1}, y_{it2}, \cdots, y_{itp})'$ be a $p \times 1$ vector of interest for cross-section i in period t. Suppose that y_{it} follows a nonstationary VAR(k) process:

(1A) $$y_{it} = \delta_i d_t + \sum_{j=1}^{k} \Phi_{ij} y_{i,t-j} + \varepsilon_{it}, \ \ t = 1, 2, ..., T; \ \ i = 1, 2, ..., N,$$

where Φ_{ij} is a $p \times p$ coefficient matrix, ε_{it} is a $p \times 1$ vector of disturbances, and d_t is a vector of deterministic components; that is, $d_t = 1$ or (1, t)', δ_i is a $p \times 1$ or $p \times 2$ matrix of parameters. Thus $\delta_i d_t$ is a $p \times 1$ vector with the j-th element equal to δ_{1ij} or $\delta_{1ij} + \delta_{2ij} t$ representing the deterministic component of the model. In this paper, we assume that the number of cross-sections (N) is fixed and the number of time periods (T) is relatively large.

Given (1A), we can also equivalently represent y_{it} as a VECM:

(1B) $$\Delta y_{it} = \delta_i d_t + \Pi_i y_{i,t-1} + \sum_{j=1}^{k-1} \Gamma_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \ \ t = 1, 2, ..., T; \ \ i = 1, 2, ..., N,$$

where $\Gamma_{ij} = -\sum_{s=j+1}^{k} \Phi_{is}$ for j=1, 2, ..., (k-1) and $\Pi_i = -(I_m - \sum_{j=1}^{k} \Phi_{ij})$.

Now, define:

(2A) $$\Gamma_i = (\Gamma_{i1}, \Gamma_{i2}, \cdots, \Gamma_{i,k-1})$$

(2B) $$X_{it} = (\Delta y_{i,t-1}', \Delta y_{i,t-2}', \cdots, \Delta y_{i,t-(k-1)}')'.$$

Then, (1B) can be rewritten as:

(3) $$\Delta y_{it} = \delta_i d_t + \Pi_i y_{i,t-1} + \Gamma_i X_{it} + \varepsilon_{it} .$$

For a given t, model (3) can be stacked over cross-sections to obtain:

(4A) $$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \vdots \\ \Delta y_{Nt} \end{bmatrix} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} d_t + \begin{bmatrix} \Pi_1 & & & \\ & \Pi_2 & & \\ & & \ddots & \\ & & & \Pi_N \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} + \begin{bmatrix} \Gamma_1 & & & \\ & \Gamma_2 & & \\ & & \ddots & \\ & & & \Gamma_N \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Nt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix},$$

or more compactly,

(4B) $$\Delta y_t = \delta d_t + \Pi y_{t-1} + \Gamma X_t + \varepsilon_t, \text{ for } t = 1, 2, ..., T,$$

where

(5A) $$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Nt} \end{bmatrix}, \quad \Delta y_t = \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \vdots \\ \Delta y_{Nt} \end{bmatrix}, \quad X_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Nt} \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}$$

(5B) $$\delta = (\delta_1', \delta_2', \cdots, \delta_N')'$$

(5C) $$\Pi = \begin{bmatrix} \Pi_1 & & & \\ & \Pi_2 & & \\ & & \ddots & \\ & & & \Pi_N \end{bmatrix}$$

(5D) $$\Gamma = \begin{bmatrix} \Gamma_1 & & & \\ & \Gamma_2 & & \\ & & \ddots & \\ & & & \Gamma_N \end{bmatrix}.$$

(4A) or (4B) is the usual form of VEC models, with coefficient matrices restricted by (5C) and (5D).

We now make the following assumption:

(A.1) ε_t is iid, with mean equal to a zero vector and covariance matrix equal to

$$\Omega = \begin{bmatrix} \Omega_{11} & \cdots & \Omega_{1N} \\ \vdots & \ddots & \vdots \\ \Omega_{N1} & \cdots & \Omega_{NN} \end{bmatrix}, \tag{6}$$

which is a $Np \times Np$ positive definite matrix, with $\Omega_{ij} \equiv var(\varepsilon_{it})$.

Here we particularly notice that the covariance matrix in (6) allows for arbitrary cross-sectional dependence across cross-sections, which is a significant relaxation of the cross-sectional independence assumption made by almost all of the current nonstationary panel data literature.

Now, suppose that the long-run coefficient matrix, Π_i, has the following reduced rank decomposition:

$$\Pi_i = \alpha_i \beta_i', \tag{7}$$

where α_i and β_i are of dimension $p \times r_i$, with $r_i = rank(\Pi_i) < p$. Here we note that we allow the cointegration rank to be different among cross-sections, which is also an extension of the existing panel cointegration literature, since the current literature on panel cointegration always assumes that different cross-sections have the same cointegration rank: $r_i = r$ for all i. Given (7), we can factor the long-run coefficient matrix Π of (5C) as:

$$\Pi = \alpha \beta', \tag{8}$$

where

(9) $$\alpha = \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_N \end{bmatrix}, \ \beta = \begin{bmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_N \end{bmatrix}.$$

Then, model (4B) above can be expressed as a familiar panel VEC model:

(10) $$\Delta y_t = \delta d_t + \alpha\beta' y_{t-1} + \Gamma X_t + \varepsilon_t, \ \text{for } t = 1, 2, ..., T.$$

This is the model typically specified by almost all of the papers in the relatively new literature on panel VEC models; see for example, Groen and Kleibergen (2000) and Larsson and Lyhagen (1999). In this specification: (i) short-run dynamics are assumed to be unrelated between cross-sections; that is, the matrix Γ is assumed to be block diagonal, as given in (5D). (ii) There are no long-run equilibrium relationships between cross-sections; in other words, cross-sectional cointegration is not permitted since β is restricted to be block diagonal, as given in (9) above. (iii) The cointegration ranks are assumed to be the same for all cross-sections. And (iv) temporary deviation from long-run equilibrium in one cross-section is not allowed to influence the other members of the panel; that is, the adjustment matrix α is assumed to be block diagonal, as given in (9) above. We believe that these four assumptions are unrealistic and very restrictive. Thus, in this paper, we seek to relax these restrictive assumptions. More precisely, we will allow the short-run dynamic matrix (Γ), the adjustment matrix (α) and the cointegration matrix (β) to be unrestricted, as follows.

(11A) $$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \cdots & \Gamma_{1N} \\ \Gamma_{21} & \Gamma_{22} & \cdots & \Gamma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{N1} & \Gamma_{N2} & \cdots & \Gamma_{NN} \end{bmatrix}$$

(11B) $$\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{N1} & \alpha_{N2} & \cdots & \alpha_{NN} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \cdots & \beta_{NN} \end{bmatrix},$$

where both matrices α and β are of dimension $Np \times r$, with $r \equiv r_1 + r_2 + \cdots + r_N < Np$. Under this specification of unrestricted matrices Γ, α and β, the panel VECM (10) allows: (i) interactions of short-run dynamics between cross-sections, (ii) influence of one cross-section's temporary long-run equilibrium error on other members of the panel, (iii) the difference in cointegration ranks across cross-sections, and (iv) cross-sectional cointegration. Using specification (10)-(11), we can also test whether the conventional block-diagonality restrictions on the short-run coefficient matrix (Γ), the cointegrating matrix (β) and the adjustment matrix (α) are valid once we have estimated the unrestricted matrices Γ, β and α.

3. Monte Carlo Simulations

To motivate our unrestricted panel VECM specification, we now conduct Monte Carlo simulations to examine the effect of cross-sectional correlation and/or cross-sectional cointegration on the size and power of a panel cointegration test. For simplicity, we use the residual-based panel KPSS test for cointegration, which is a direct extension of the residual-based univariate KPSS test for cointegration; see, for example, Shin (1994).[5]

More specifically, the panel KPSS testing statistic for the null of cointegration is calculated by:

$$\overline{LM} = \frac{N^{-1}\sum_{i=1}^{N} T^{-2} \sum_{t=1}^{T} S_{it}^2}{\hat{\sigma}_\varepsilon^2},$$

[5] Eriksson (2004) emphasizes that Monte Carlo experiments that compare ML estimators to a KPSS-style tests are more reasonable than the more-common experiments based on Dickey-Fuller or Phillips-Perron tests. Both ML (likelihood ratio) and KPSS tests have null hypotheses of cointegration, while Dickey-Fuller-type tests have a null hypothesis of no cointegration, making subtle any conclusions drawn from the experiments using DF or PP tests.

where $S_{it} = \sum_{j=1}^{t} \hat{\varepsilon}_{ij}$ and $\hat{\sigma}_{\varepsilon}^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^2$, with $\hat{\varepsilon}_{it}$ being the estimated residual of cross-section i at time t.

The simulation design for the data generating processes (DGPs) is as follows:

$$y_{it} = \alpha_i + x_{it}\beta_i + z_t\gamma_i + \varepsilon_{it}$$

$$\alpha_i \sim U[0, 10],\ \ \beta_i \sim U[0, 2],\ \ \gamma_i \sim U[0, 2]$$

$$x_{it} = x_{i,t-1} + v_{it},\ \ v_{it} \sim IN(0,\ \sigma_i^2),\ \sigma_i^2 \sim U[0.5, 1.5]$$

$$z_t = z_{t-1} + e_t,\ \ e_t \sim IN(0, \lambda^2),\ \ \lambda^2 \sim U[0.5, 1.5]$$

$$\varepsilon_{it} = \theta\sum_{k=1}^{t} u_{ik} + u_{it}$$

$$(u_{1t}, ..., u_{Nt}) = (\eta_{1t}, ..., \eta_{Nt}) \cdot L$$

$$\eta_{it} \sim IN(0, \delta_i^2),\ \ \delta_i^2 \sim U[0.5, 1.5],$$

where α_i is the fixed effect for cross-section i, x_{it} is the I(1) regressor of cross-section i that varies over cross-sections, z_t is the common stochastic trend across cross-sections that captures the cross-sectional cointegration among the regressors of different cross-sections. The parameter θ controls the degree of nonstationarity in the regression error terms, and the lower triangular matrix L ($N \times N$) controls the cross-sectional correlation. The parameter values used in the simulations are:

- Sample Size: T={50, 100}
- No. of cross-sections: N={2, 5, 10}
- $\theta \in \{0.00, 0.05, 0.10, 0.15, 0.20\}$.

McCoskey and Kao (2001) use a similar simulation design, though they do not consider the effects on KPSS tests of cross-sectional correlation or cross-sectional cointegration.

When $\theta = 0$, then $\varepsilon_{it} = u_{it}$, which is stationary. Thus, $\theta = 0$ implies a cointegration relationship between y_{it}, x_{it} and z_t. Also, if $\gamma_i = 0$ for all i, there is no cross-sectional cointegration between the regressors of different cross-sections; so $\beta_i \neq 0$ and $\gamma_i = 0$ (for all i) corresponds to the case of within cointegration only.[6]

[6] Banerjee, Marcellino and Osbat (2004) consider similar parameterizations.

We consider both the dynamic OLS (DOLS) and the dynamic seemingly unrelated regressions (DSUR) estimation methods in our simulations. Note: DSUR is asymptotically equivalent to MLE if all the members of the panel are cointegrated; see Mark, Ogaki and Sul (2005).

For DOLS, the cases considered are:

Case 1: with no cross-sectional correlation or cross-sectional cointegration; that is:

$L = I_N$, $\gamma_i = 0$ for all i.

Case 2: with cross-sectional correlation but with no cross-sectional cointegration; that is:

L is a lower triangular matrix, and $\gamma_i = 0$ for all i.

Case 3: with no cross-sectional correlation but with cross-sectional cointegration; that is:

$L = I_N$, $\gamma_i \neq 0$ for some of the i.

Case 4: with both cross-sectional correlation and cross-sectional cointegration; that is:

L is a lower triangular matrix, and $\gamma_i \neq 0$ for some of the i.

For DSUR, the cases considered are:

Case 1: with cross-sectional correlation but with no cross-sectional cointegration,

Case 2: with no cross-sectional correlation but with cross-sectional cointegration,

Case 3: with both cross-sectional correlation and cross-sectional cointegration.

Our simulations are conducted in GAUSS 3.6 and the number of replications used (i.e. R) is 5,000. The simulation results are reported in Tables 1A-4B. A brief summary of our main findings from the simulations is as follows.

A. Tables 1A-1B indicate that:

(i) When there is cross-sectional correlation, the panel KPSS testing statistic (i.e. LM-bar) is severely over-sized; that is, it over-rejects panel cointegration. On the other hand, the KPSS statistic applied separately to each cross-section has the proper size.

(ii) When there is cross-sectional cointegration, the LM-bar and the individual LM statistics are all severely under-sized; that is, they over-accept panel cointegration.

(iii) When there are both cross-sectional correlation and cointegration, the LM-bar and the individual LM statistics continue to be severely under-sized but the degree of size distortion is less than in case 3.

B. Tables 2A-2B indicate that:

Both cross-sectional correlation (case 2) and cross-sectional cointegration (case 3) cause severe loss of power. However, case 3 is much more severe than case 2, especially when the error term is nearly stationary (that is, when the value of theta is low).

C. Tables 3A-4B indicate that:

The huge size distortion and severe power loss of the LM statistics based on DOLS that we found in Tables 1A-2B are exacerbated for dynamic GLS estimator (or MLE) because of the cross-equation contamination.

In summary, cross-sectional cointegration causes more severe size distortion and power loss for the pooled (i.e. panel) and individual LM-statistics than does cross-sectional correlation. Our finding warns practitioners of the limitations of the existing panel cointegration testing procedures, since almost all of these procedures neglect the possibility of long-run equilibrium relationships between cross-sections.[7] For example, on the ongoing debate about whether PPP holds or not, practitioners usually do not pay any attention to the obvious cross-sectional cointegration: the same price level of the reference country appears in all equations of the panel.

Finally, a word on our simulation design. Our simulation is conducted for single equations not for a VAR or a VECM. This is mainly for simplicity in designing the DGPs and programming the relevant calculations. However, we believe that our findings based on this relatively simple DGP design will carry over to VAR or VEC models, since existing estimation and tests for panel cointegration almost always neglect cross-sectional dependence and/or long-run cross-sectional equilibrium relationships.

Our simulation results indicate that when there is cross-sectional correlation or cross-sectional cointegration, existing panel cointegration tests have large size distortion and low power, while the estimates for long-run parameters may be inconsistent if cross-sectional cointegration is neglected. Now a challenging question is how to find a valid estimation and testing procedure for panel cointegration models that have cross-sectional

[7] Similar results are obtained and warnings made by Banerjee, Marcellino and Osbat (2004).

correlation and/or cross-sectional cointegration. This is the main motivation for our unrestricted panel VECM estimation and testing approach to be defined in the next section.

4. Canonical Correlation Analysis

Box and Tiao (1977) consider the predictability of linear combinations of a multivariate time series from the history of the linear combinations concerned. More specifically, suppose that we have a random sample of T observations on a p dimensional time-series $y_t = (y_{1t}, y_{2t}, ..., y_{pt})'$, which has mean zero and variance equal to Σ_{00}. Box and Tiao consider the predictability of $z_t \equiv c' y_t$ based on the past values of z_t, where c is a $p \times 1$ vector of constants. Then, if the linear combination z_t is stationary and independently distributed (for example, a white noise process), the past values of z_t are not informative for forecasting the current value of z_t. On the other hand, if z_t is very persistent over time (for example, z_t is nonstationary or has long memory) then the past values of z_t can forecast the current value of z_t very well. Precisely the same motivation/ intuition underlies the recent error-correction, residual-based cointegration tests of Westerlund (2005, 2006) and Westerlund and Edgerton (2005).

More precisely, define the linear projection of y_t on its own history as

(12) $$\hat{y}_t = E(y_t \mid y_{t-1}, y_{t-2}, \cdots) = \sum_{i=1}^{k} \Gamma_i y_{t-i} ,$$

where the projection coefficient matrices Γ_i's are $p \times p$. Then, we have:

(13) $$y_t = \hat{y}_t + e_t ,$$

where e_t is the projection error that is uncorrelated with $\hat{y}_t$ (or the lagged values of y_t). Define:

(14) $$\Omega(y_t) = \text{var}(y_t) ,\ \Omega(\hat{y}_t) = \text{var}(\hat{y}_t) ,\ \Sigma = \text{var}(e_t) .$$

Assume that $\Omega(y_t)$ is positive definite and $\Omega(\hat{y}_t)$ and Σ are positive semi-definite.

Now consider forecasting the linear combination $c'y_t$. Box and Tiao (1977) use

$$\mu \equiv \frac{\mathrm{var}(c'\hat{y}_t)}{\mathrm{var}(c'y_t)} \tag{15}$$

to measure the forecastability of the linear combination $c'y_t$. Notice that

$$\mu = \frac{c'\Omega(\hat{y}_t)c}{c'\Omega(y_t)c}, \tag{16}$$

then, under the normalization $c'\Omega(y_t)c = 1$, it is easy to show that the c that minimizes (maximizes) μ (measuring the degree of predictability of $c'y_t$ based on $c'\hat{y}_{t-1}$) satisfies the following first order condition:

$$\Omega(\hat{y}_t)c = \lambda\Omega(y_t)c, \tag{17}$$

where λ is an eigenvalue of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$.[8] Thus, the c that minimizes (maximizes) μ is just the eigenvector corresponding to the smallest (largest) eigenvalue of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$. Notice that (17) also implies:

$$c'\Omega(\hat{y}_t)c = \lambda c'\Omega(y_t)c \tag{18}$$

or equivalently:

$$\lambda = \frac{c'\Omega(\hat{y}_t)c}{c'\Omega(y_t)c}. \tag{19}$$

Thus, the minimum (maximum) value of μ achieved also equals the smallest (largest) eigenvalue. We now want to show that the maximum value of μ defined in (15) above also equals the squared maximum canonical correlation (CC) coefficient between y_t and $\hat{y}_t$. To this end, let us suppose that λ_{max} is the maximum eigenvalue of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$ and v_{max} is the corresponding eigenvector. Let $a_t = v_{max}'y_t$, $\hat{a}_t = v_{max}'\hat{y}_t$ and $u_t = a_t - \hat{a}_t$. Then, using (16) and (19), we have:

$$\text{(20)} \qquad \lambda_{max} = \mu = \frac{var(\hat{a}_t)}{var(a_t)}.$$

On the other hand, the maximum canonical correlation coefficient between y_t and $\hat{y}_t$ is given by:

$$\text{(21)} \qquad \rho_{max} \equiv \frac{cov(a_t, \hat{a}_t)}{[var(a_t)\,var(\hat{a}_t)]^{1/2}} = \frac{cov(\hat{a}_t + u_t, \hat{a}_t)}{[var(a_t)\,var(\hat{a}_t)]^{1/2}}$$

$$= \frac{var(\hat{a}_t)}{[var(a_t)\,var(\hat{a}_t)]^{1/2}} = \frac{[var(\hat{a}_t)]^{1/2}}{[var(a_t)]^{1/2}}$$

where we use $a_t = \hat{a}_t + u_t$ and $cov(\hat{a}_t, u_t) = 0$. Comparing (20) and (21), we obtain $\lambda_{max} = \rho^2_{max}$; that is, the maximum (minimum) eigenvalue of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$ equals the squared maximum (minimum) canonical correlation between y_t and $\hat{y}_t$.

We now wish to extend the above idea to the hypothesis testing for the CI rank of the VECM for y_t. To this end, we assume that the eigenvalues of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$ are ordered as:

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_p$$

and the corresponding eigenvectors are:

---

[8] See, for example, Dhrymes (1978), p.72. Hamilton (1994) explains that this language reflects no more than the most-common normalization used when calculating eigenvalues. Here, for precision, we retain the classic language.

$$v_1, v_2, \cdots, v_p,$$

with normalization

(22) $$v_i'\Omega(y_t)v_i = 1, \text{ for } i = 1, 2, ..., p.$$

Now, we establish Lemma 1, as follows.

Lemma 1: (i) The eigenvalues of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$ satisfy $0 \le \lambda_i \le 1$ for all i. (ii) The p eigenvectors $v_1, v_2, \cdots, v_p$ are linearly independent.

Proof: See Appendix. ■

We now define:

(23a) $$\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p), \text{ with } 0 \le \lambda_i \le 1 \text{ for i=1, 2, ..., p},$$

(23b) $$M = (v_1, v_2, ..., v_p)',$$

(23c) $$z_t = My_t.$$

Then,

(24) $$z_t = \begin{bmatrix} z_{1t} \\ \vdots \\ z_{pt} \end{bmatrix} = \begin{bmatrix} v_1'y_t \\ \vdots \\ v_p'y_t \end{bmatrix}.$$

Thus, using $y_t = \hat{y}_t + e_t$ we have:

(25) $$z_t = \hat{z}_t + q_t,$$

where $\hat{z}_t = M\hat{y}_t = E(z_t | z_{t-1}, z_{t-2}, ...)$ and $z_t = Me_t$. (25) implies:

(26) $$\Omega(z_t) = \Omega(\hat{z}_t) + \text{var}(q_t)\ ,$$

since $\hat{z}_t$ and q_t are uncorrelated.

Now, we are ready to establish Lemma 2, as follows.

Lemma 2: We can transform the original $p \times 1$ vector y_t into a $p \times 1$ canonical vector $z_t = My_t$ such that $z_t = \hat{z}_t + q_t$ (that is, (25)), where

(i) $\text{var}(z_t) = I_p$;

(ii) $\text{cov}(z_t, \hat{z}_t) = \text{var}(\hat{z}_t) = \Lambda \equiv \text{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$, with $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_p \le 1$; and

(iii) $\text{var}(r_t) = I_p - \Lambda$.

Proof: (i) Under the normalization (22), we can easily verify that $\Omega(z_t) = M\Omega(y_t)M'$ $= I_p$.

(ii) Since the columns M' are the eigenvectors of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$, we have:

(27) $$\Omega(\hat{y}_t)M' = \Omega(y_t)M'\Lambda\ .$$

Then, premultiplying it by M and using $M\Omega(y_t)M' = I_p$, we obtain:

(28) $$M\Omega(\hat{y}_t)M' = \Lambda\ .$$

That is, $\text{var}(\hat{z}_t) = M\Omega(\hat{y}_t)M' = \Lambda$.

(iii) Using $\Omega(z_t) = \Omega(\hat{z}_t) + \text{var}(q_t)$, we obtain:

(29) $$\text{var}(q_t) = \Omega(z_t) - \Omega(\hat{z}_t) = I - \Lambda\ ,$$

by results (i) and (ii) above. This completes the required proof. ■

- Remark 1: The eigenvalues ($\lambda_1, \lambda_2, \cdots, \lambda_p$) are just the squared canonical correlation coefficients between y_t and $\hat{y}_t$ and are ordered from the smallest to the largest; for example, λ_1 is the square of the smallest canonical correlation coefficient between the covariates $z_{1t} = v_1' y_t$ and $\hat{z}_{2t} = v_1' \hat{y}_t$, and λ_p is the square of the largest canonical correlation coefficient between the covariates $z_{pt} = v_p' y_t$ and $\hat{z}_{pt} = v_p' \hat{y}_t$. A large canonical correlation between $z_{pt} = v_p' y_t$ and $\hat{z}_{pt} = v_p' \hat{y}_t$ implies that the canonical covariate $z_{pt} = v_p' y_t$ is highly predictable according to Box and Tiao (1977); in other words, if λ_p is very close to 1, then the canonical covariate $z_{pt} = v_p' y_t$ is highly predictable. Then, using the fact that an I(1) process is highly predictable, we can now use the hypothesis of $\lambda_p = 1$ to test whether the canonical covariate $z_{pt} = v_p' y_t$ is I(1) or not. On the other hand, if the canonical covariate $z_{1t} = v_1' y_t$ is highly unpredictable, then λ_1 must be very close to zero or at least significantly less than 1. Thus, we can also use the hypothesis of $\lambda_1 < 1$ to test whether the first canonical covariate $z_{1t} = v_1' y_t$ is I(0) or not.
- Remark 2: The number of eigenvalues (that is, the λ's) that are close to 1 is the same as the number of linear combinations that can be almost perfectly forecasted; see, Box and Tiao (1977). Thus, the number of eigenvalues that are close to 1 is the same as the number of common stochastic trends in the VAR system for y_t; see for example, Bewley and Yang (1995).
- Remark 3: The difference between Johansen's MLE-based method and Box-Tiao's canonical correlation method is that Johansen's method calculates the canonical correlation between Δy_t and y_{t-1}, while the Box-Tiao canonical correlation analysis uses the canonical correlation between y_t and y_{t-1}. The main reason for this difference is that Johansen's MLE approach uses the prior information that the elements of y_t are I(1), while the Box-Tiao canonical correlation analysis does not.
- Remark 4: Based on Remark 2 above, if the CI rank equals r, then we have:

$$0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_r < 1, \text{ and } \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_p = 1.$$

Thus, we can directly test the hypotheses about the number of common trends of a VAR system based on hypotheses about the number of eigenvalues of $\Omega(\hat{y}_t)$ (in the metric of $\Omega(y_t)$) that are equal to one. Given the fact that the eigenvalues are already ranked from the smallest to the largest, we can test the hypothesis that there are (k-r) common trends based on the following null and alternative:

$$H_0:\ \lambda_{r+1} = 1 \ \text{ v.s. } \ H_1:\ \lambda_{r+1} < 1.$$

Bewley and Yang (1995) proposed several CI tests based on this idea.

- Remark 5: Notice that

$\lambda_i = v_i'\Omega(\hat{y}_t)v_i / v_i'\Omega(y_t)v_i = \text{var}(v_i'\hat{y}_t)/\text{var}(v_i'y_t) = \text{var}(\hat{z}_{it})/\text{var}(z_{it})$. Then if $z_{it} = v_i'y_t$ is almost unpredictable, λ_i must be less than 1. On the other hand, if $z_{it} = v_i'y_t$ is almost perfectly predictable, λ_i must be equal to (or at least very close to) 1. Thus, we can test the hypothesis that $\lambda_i < 1$ based on testing whether the corresponding canonical covariate $z_{it} \equiv v_i'y_t$ is stationary or not. Therefore, we can apply KPSS testing statistic to the canonical covariate $z_{it} = v_i'y_t$. This approach is analogous to the residual-based test for CI; see, for example, Shin (1994). Similarly, we can test whether $\lambda_i = 1$ or not based on testing whether the corresponding canonical covariate $z_{it} \equiv v_i'y_t$ has a unit root or not, which can be executed by applying Dickey-Fuller (DF) or augmented DF (ADF) unit-root test to z_{it}.

As an illustration, let us now examine the VAR(1) model in detail. Suppose that the $p \times 1$ vector y_t follows VAR(1):

(30) $$y_t = \phi y_{t-1} + u_t,$$

where ϕ is the $p \times p$ coefficient matrix. Then, using the notation of this section, we have:

(31) $$y_t = \hat{y}_t + u_t,$$

where $\hat{y}_t = E(y_t | y_{t-1}, y_{t-2},...) = \phi y_{t-1}$. Thus, $\Omega(y_t) = \Omega(\hat{y}_t) + \Sigma$, where $\Omega(y_t) = E(y_t y_t') \equiv \Gamma_{0t}$ and $\Omega(\hat{y}_t) = \text{var}(\phi y_{t-1}) = \phi \Gamma_{0t} \phi'$.

Let $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_p \le 1$ be the ranked eigenvalues of $\Omega(\hat{y}_t) = \phi \Gamma_{0t} \phi'$ in the metric of $\Omega(y_t) = \Gamma_{0t}$ and $v_1, v_2, \cdots, v_p$ be the corresponding eigenvectors that are linearly independent. As in the previous section, let $M = (v_1, \cdots, v_p)'$.

Then, premultiplying (30) by M, we have:

(32) $$z_t = \tilde{\phi} z_{t-1} + q_t \equiv \hat{z}_t + q_t,$$

where $z_t = My_t$, $\tilde{\phi} = M\phi M^{-1}$ and $q_t = Mu_t$, and where, by Lemma 2 above,

$$\text{var}(z_t) = I_p,$$

$$\text{cov}(z_t, \hat{z}_t) = \text{var}(\hat{z}_t) = \Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p),$$

$$\text{var}(q_t) = I_p - \Lambda.$$

The VAR(1) model in (32) is usually referred to as the canonical model.

Now, we turn to examining the properties of $\tilde{\phi}$ when some of the eigenvalues approach the unit circle. Specifically, suppose that (p-r) eigenvalues of ϕ approach points on the unit circle. Partition z_t, q_t and $\tilde{\phi}$ as:

$$z_t = \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix}, \quad q_t = \begin{bmatrix} q_{1t} \\ q_{2t} \end{bmatrix}, \quad \tilde{\phi} = \begin{bmatrix} \tilde{\phi}_{11} & \tilde{\phi}_{12} \\ \tilde{\phi}_{21} & \tilde{\phi}_{22} \end{bmatrix},$$

where z_{1t} and q_{1t} are $r \times 1$ vectors and $\tilde{\phi}_{11}$ is $r \times r$.

Box and Tiao (1977) showed the following ***important results***:

(1) If, and only if, (p-r) eigenvalues of ϕ (or equivalently, $\tilde{\phi}$) approach values on the unit circle, then (p-r) eigenvalues of $\Omega(\hat{y}_t) = \phi \Gamma_{0t} \phi'$ in the metric of $\Omega(y_t) = \Gamma_{0t}$ approach 1.

(2) The canonical model for z_t in (32) becomes, in the limit,

(33) $$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} \tilde{\phi}_{11} & 0 \\ \tilde{\phi}_{21} & \tilde{\phi}_{22} \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} + \begin{bmatrix} q_{1t} \\ q_{2t} \end{bmatrix}$$

where z_{1t} follows a stationary VAR(1) process and z_{2t} is non-stationary. Thus, the canonical transformation ($z_t = My_t$) decomposes the original p-dimensional vector y_t into two subvectors: one stationary subvector (z_{1t}: $r \times 1$) and the other non-stationary subvector (z_{2t}: $(p-r) \times 1$) that also depends on $z_{1,t-1}$.

- Remark 6: Expression (33) can be thought of as the triangular representation for the vector of canonical covariates $z_t = M' y_t$. In contrast, Phillips's (1991) triangular representation is for the vector y_t itself.
- Remark 7: Given (33) above, we can also test hypotheses about the CI rank (for example, H_0: CI rank = r v.s. H_1: CI rank < r) based on testing:

  $$H_0: \ \tilde{\phi}_{12} = 0 \quad \text{v.s.} \quad H_1: \ \tilde{\phi}_{12} \neq 0,$$

  because $\tilde{\phi}_{12} = 0$ holds if, and only if, (p-r) of the p eigenvalues of $\Omega(\hat{y}_t) = \phi\Gamma_{0t}\phi'$ in the metric of $\Omega(y_t) = \Gamma_{0t}$ are equal to 1; that is, $\tilde{\phi}_{12} = 0$ holds if, and only if (p-r) of the p canonical correlation coefficients between y_t and $\hat{y}_t$ are equal to 1. Moreover, the testing statistics for $\tilde{\phi}_{12} = 0$ are just the Wald-type statistics and are usually distributed as a Chi-square.

5. Application: A Panel Cointegration Model for M1 Velocities

5.1. Testing for Cointegration Rank Based on Johansen's Statisitcs

In this subsection, we apply Johansen's MLE method to the unrestricted panel VECM model (12) using a panel data set of M1 demand in United States and Canada. More specifically, we specify the following model for M1 velocity:

(34) $$\log(GDP_{it} / M1_{it}) = \beta_0 + \beta_1 (1/R_{it}) + \varepsilon_{it},$$

where i indexes country (Canada and U.S.), t indexes year (1919-1980), R is the long rate, and ε is the error term of the model.

Using ADF tests, we found that log(GDP/M1) and 1/R are unit root processes for both US and Canada at 5% significance level. Thus, model (34) posits a cointegration relationship between M1 velocity and inverse long rate. However, when we use Johansen's (1995) ML-based trace and maximum-eigenvalue statistics and their corresponding asymptotic distributions to test for the cointegration rank of the panel VECM in (34), we accept the null hypothesis that the cointegration rank is equal to 2, where we use Akaike and Schwarz information criteria to determine the specification of the deterministic components and the number of lagged differences included; more specifically, based on the information criteria, we include the constant term in both the cointegration equations and the VECM, and we include two lagged differences in the VECM. The detailed testing results are summarized in Table 5.

If we apply the same tests to (34) for US and Canada separately, we can confirm that the cointegration rank is equal to one for each country. Thus, it is now clear that applying Johansen's asymptotic tests to panel VECM (34) fails to find any cross-country cointegration. One possible reason for this failure may be that Johansen's asymptotic distributional approximation is not very accurate for a sample size of 62.

One way to overcome the inaccuracy of the asymptotic distributional approximation to Johansen's trace and maximum-eigenvalue statisitcs is to bootstrap Johansen's testing statistics. Table 6 reports the bootstrapped distributions of Johansen's trace and maximum-eigenvalue testing statistics for testing the cointegration rank of panel VECM (34), where the number of bootstrap used is 10,000. Based on the bootstrapped distributions, we accept the null hypothesis that the cointegration rank is equal to 2 at 10% significance level. Again, bootstrapping Johansen's testing statistics still fails to find any cross-country cointegration.

One alternative method to test for the cointegration rank of panel VECM (34) is to bootstrap the trace statistic based on canonical correlations of Bewley and Yang (1995).

5.2. Bootstrapping the Trace Statistic Based on Box-Tiao's Canonical Correlations

Suppose that the squared canonical correlations are ordered as $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_p \le 1$. We now wish to test whether the cointegration rank is r or not; that is, we wish to test the following hypothesis, H_0: CI rank = r and H_1: CI rank > r. This is equivalent to testing, H_0: $\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_p = 1$ and H_1: $\lambda_{r+1} \le \lambda_{r+2} \le \cdots \le \lambda_{r+j} < 1$ for some j<(n-r). We consider the following trace testing statistic:

The trace statistic: $\lambda_{tr1} = \sum_{i=r+1}^{p} T(1-\lambda_i)$

The trace statistic is considered by Bewley and Yang (1995), and it is analogous to Johansen's trace statistic.

Since we do not want to assume that the true DGP is multivariate normal with the same covariance matrix over time, nor do we want to assume that the sample size used is large, we choose to bootstrap the trace statistic to find empirically correct critical values and *p*-values. We follow the bootstrap procedure of van Giersbergen (1996). More specifically, we follow the following six steps to bootstrap its finite sample distribution.

Step 1: For a given sample, calculate the empirical value of $\lambda_{tr1} = \sum_{i=r+1}^{p} T(1-\lambda_i)$.

Step 2: Estimate parameter values under the joint null hypothesis H_0' : CI rank = r and $\mu = 0$ by running the restricted VECM regression:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{k-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t ,$$

where α and β are of dimension $p \times r$. Let $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\Gamma}_1$, ..., $\tilde{\Gamma}_{k-1}$ be the restricted coefficient estimators and $\tilde{\varepsilon}_t$ be the corresponding residuals. Let $\tilde{\mu} = T^{-1}\sum_{t=1}^{T}\tilde{\varepsilon}_t$ be the sample mean of the residuals. Let $\hat{\varepsilon}_t$ be the scaled and centered residuals, $\hat{\varepsilon}_t = [T/(T-p(k-1))]^{1/2}(\tilde{\varepsilon}_t - \tilde{\mu})$, which is stationary under the null hypothesis H_0. Thus, we can now use stationary bootstrap method to resample these adjusted residuals.

Step 3: Let $\{\varepsilon_t^* : t = 1, ..., T\}$ be T resampled residuals from the adjusted residuals, $\{\hat{\varepsilon}_t : t = 1, ..., T\}$. Given a resample of $\{\varepsilon_t^* : t = 1, ..., T\}$, generate a bootstrap sample $\{y_t^* : t = 1, ..., T\}$ from the restricted model:

$$\Delta y_t^* = \tilde{\alpha}\tilde{\beta}' y_{t-1}^* + \sum_{j=1}^{k-1} \tilde{\Gamma}_j \Delta y_{t-j}^* + \varepsilon_t^* \text{, for t = 1, 2, ..., T,}$$

where the initial values are given by $y_s^* = y_s$ for s = (1-k), (2-k), ..., 1, 0.

Step 4: Use the bootstrapped sample $\{y_t^* : t = 1, ..., T\}$ from step 3 to calculate a bootstrap realization of the trace statistic, denoted by λ_{trb}^*.

Step 5: Repeat steps 3 and 4 a large number of times, say B times, for b=1, 2, ..., B.

Step 6: For a given significance level α, we reject the null hypothesis of cointegration rank equal to r, if the empirical value (calculated in step 1) is larger than the $[(1-\alpha)B]$-th largest bootstrap realization; that is, we reject the null if $\lambda_{tr} > \lambda_{tr[(1-\alpha)B]}^*$, where $\lambda_{tr[1]}^* < \lambda_{tr[2]}^* < \cdots < \lambda_{tr[B]}^*$.

Our bootstrap is implemented in GAUSS 6.0 and the number of bootstrap (that is, B in step 5 above) used is 10,000. Since our data is annual and also as we explained in the previous subsection, we choose to include two lagged difference; that is, we estimate a VECM(2) model. Table 7 provides the bootstrapped distribution of the trace statistic for testing the cointegration rank. Based on the bootstrap distribution in Table 7, we conclude that the null hypothesis that the CI rank is equal to 3 is accepted at significance level 10%. The three normalized cointegration vector estimators are:

$$\hat{\beta}_1 = \begin{bmatrix} 1 \\ \hat{\beta}_{12} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.031 \\ 0 \\ 0 \end{bmatrix}, \quad \hat{\beta}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \hat{\beta}_{24} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0.041 \end{bmatrix}, \quad \hat{\beta}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \hat{\beta}_{34} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1.538 \end{bmatrix},$$

where we normalized the cointegration vectors in such a way that $\hat{\beta}_1$ is the within-U.S. cointegration vector, $\hat{\beta}_2$ is the within-Canada cointegration vector, and $\hat{\beta}_3$ is the between-cointegration vector (that is, the cointegration vector between U.S. and canada).

Using the inferences based on bootstrapped distribution of the CC-based trace statistic, we successfully uncover the cross-country cointegration, which Johansen's asymptotic and bootstrapped tests fail to find. Thus, based on this simple empirical application, we believe that the bootstrapped canonical correlation analysis approach is superior to the ML-based approach for testing hypotheses of cointegration rank, since the bootstrapped canonical correlation method does not depend on distributional assumption of the true DGP, does not require the covariance matrix of the VECM errors to be homoscedastic (in fact, our bootstrap procedure accommodates for possible heteroscedasticity in the VECM errors), nor does it require that the sample size used is large.

6. Conclusions

Given the poor small sample performance of Johansen's ML-based asymptotic approach to testing for cointegration ranks of vector error correction models and its critical dependence on the distributional assumptions, we believe that there is a genuine need to find alternative methods for testing cointegration ranks that do not depend on the restrictive distributional assumptions or inaccurate asymptotics. In this paper, we consider an alternative statistic for testing cointegration ranks based on Box and Tiao's (1977) canonical correlation approach. To ensure that our canonical correlation based test has correct empirical size, we use bootstrap method to find the finite-sample distribution of the testing statistic.

The current literature on panel cointegration tests almost always assumes that (i) there is no contemporaneous cross-sectional correlation in the error terms; (ii) there is no interaction of short-run dynamics between cross-sections; (iii) different cross-sections have the same cointegration rank; and (iv) there are no long-run equilibrium relationships between different cross-sections. In this paper, we argue that cross-sectional dependence in short-run is pervasive since different cross-sectional units are influenced by the same types of domestic and international macro shocks, and that long-run equilibrium relationships between cross-sections are also very common since different economies (or cross-sections) tend to move together in the long-run, especially in this age of globalization.

Our Monte Carlo simulations demonstrate that the presence of short-run and/or long-run cross-sectional dependence causes very severe size distortions and power loss for the panel KPSS cointegration test. To overcome the weakness of the current panel cointegration tests, we propose in this paper an unrestricted panel VECM that allows for arbitrary contemporaneous correlation, cross-sectional interaction of short-run dynamics, heterogeneous cointegration ranks across cross-sections, as well as cointegration between different cross-sections.

In our empirical application of an unrestricted panel VECM for the long-run determination of M1 velocities in U.S. and Canada, using bootstrap method, we unequivocally find three cointegration relationships; two for within-country cointegrations (one for each country) and the other for between-country cointegration. This is in sharp contrast to the conclusion of cointegration rank equal to 2 that is reached by Johansen's ML-based testing statistics using asymptotic distributions or bootstrapped distributions.

The unrestricted panel VECM approach advocated in this paper can be easily applied to many other interesting economic and financial problems; for examples, the testing of economic convergence of OECD countries and the estimation of consumption and investment functions across regions and states.

Appendix

Proof of Lemma 1: (i) See Propositions 62 and 64 of Dhrymes (1978, pp. 72-74).

(ii) Because $\Omega(y_t)^{-1/2}\Omega(\hat{y}_t)\Omega(y_t)^{-1/2}$ is a real symmetric matrix, there exist p orthogonal eigenvectors, say b_1, b_2, ..., b_p. Let $B = (b_1, \cdots, b_p)$. Then,

$$\Omega(y_t)^{-1/2}\Omega(\hat{y}_t)\Omega(y_t)^{-1/2}B = B\Lambda . \tag{A.1}$$

Premultiplying it by $\Omega(y_t)^{1/2}$, we have:

$$\Omega(\hat{y}_t)[\Omega(y_t)^{-1/2}B] = \Omega(y_t)[\Omega(y_t)^{-1/2}B]\Lambda , \tag{A.2}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p)$. This means that the columns of $\Omega(y_t)^{-1/2}B$ are the eigenvectors of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$ and the diagonal elements of Λ are the corresponding eigenvalues. Now, let $M' = \Omega(y_t)^{-1/2}B$. Then,

$$M\Omega(y_t)M' = B'\Omega(y_t)^{-1/2}\Omega(y_t)\Omega(y_t)^{-1/2}B = B'B = I_p , \tag{A.3}$$

since the columns of B are orthogonal. Thus, rank(M)=p, which implies that the eigenvectors of $\Omega(\hat{y}_t)$ in the metric of $\Omega(y_t)$ must be linearly independent. This completes the required proof. ■

References

Anderson R.G. and R.H. Rasche (2001), "The Remarkable Stability of Monetary Base Velocity in the United States, 1919-1999," manuscript, St. Louis Federal Reserve Bank.

Aznar, A. and M. Salvador (2002), " Selecting the Rank of the Cointegration Space and the Form of the Intercept Using an Information Criterion," *Econometric Theory* 18, 926-947.

Baltagi, B. H., and C. Kao (2000), "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey," *Advances in Econometrics* 15, 7-51.

Banerjee, A. (1999), "Panel Data Unit Roots and Cointegration: An Overview," *Oxford Bulletin of Economics and Statistics*, Special Issue, 607-629.

Banerjee, A., M. Marcellino, and C. Osbat (2004), "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-economic Data," *Econometrics Journal*, 7, pp. 322-340.

Bewley, R.A., Orden, D., Yang, M. and L.A. Fisher (1994), "Comparison of Box-Tiao and Johansen Canonical Estimators of Cointegrating Vectors in VEC(1) Models," *Journal of Econometrics* 64, 3-27.

Bewley, R. and M. Yang (1995), "Tests for Cointegration Based on Canonical Correlation Analysis," *Journal of the American Statistical Association* 90, pp. 990-996.

Binder, M., C. Hsiao and M. Pesaran (2000), "Estimation and Inference in Short Panel Vector Autoregressions with Unit Roots and Cointegration," working paper, University of Cambridge.

Box, G.E.P. and G.C. Tiao (1977), "A Canonical Analysis of Multiple Time Series," *Biometrika* 64, 355-65.

Chan, N.H. and C.Z. Wei (1987), "Asymptotic Inferences for Nearly Nonstationary AR(1) Processes," *Annals of Statistics* 15, 1050-1063.

Dhrymes, P.J. (1978), *Mathematics for Econometrics*, Springer-Verlag.

Eliot, G. (2000), "Estimating Restricted Cointegrating Vectors," *Journal of Business and Economic Statistics* 18(1), 91-99.

Eriksson, Asa (2004). "Testing Structural Hypotheses on Cointegration Vectors: A Monte Carlo Study," mimeo, Department of Economics, Lund University, December.

Flores, R., P. Jorion, P. Pruemont and A. Szafarz (1999), "Multivariate Unit Root Tests of the PPP Hypothesis," *Journal of Empirical Finance*, 335-353.

Goncalves, Silvia and Halbert White (2004). “Maximum likelihood and the bootstrap for nonlinear dynamic models,” *Journal of Econometrics*, 119, pp. 199-219.

Gonzalo, J. and J-Y Pitarakis (1988), “Specification via Model Selection in Vector Error Correction Models,” *Economics Letters* 60, 321-328.

Groen, J. and F. Kleibergen (2000), “Likelihood-Based Cointegration Analysis in Panels of VECMs,” working paper, Tinbergen Institute.

Hamilton, James D. (1994) *Time Series Analysis*. Princeton University Press.

Harris, D. (1997), “Principal Components Analysis of Cointegrated Time Series,” *Econometric Theory* 13, 529-557.

Haug, A. (1996), “Tests for Cointegration: A Monte Carlo Comparison,” *Journal of Econometrics* 71, 89-115.

___ . (2002). “Testing Linear Restrictions on Cointegrating Vectors: Sizes and Powers of Wald and Likelihood Ratio Tests in Finite Samples,” *Econometric Theory,* 18(2), pp. 505-524.

Horvath, M. and M. Watson (1995), “Testing for Cointegration When Some of the Cointegrating Vectors are Prespecified,” *Econometric Theory* 11, 984-1014.

Im, K.M., H. Pesaran, and Y. Shin (1997), "Testing For Unit Roots in Heterogeneous Panels," working paper, Department of Applied Economics, University of Cambridge, Trinity College.

Jacobson, T. (1995), “Simulating Small-sample Performance of the Maximum Likelihood Cointegration Method: Estimation and Testing,” *Finnish Economic Papers* 8, 96-107.

Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press: Oxford.

___ . (2000) “A Bartlett Correction Factor for Tests on the Cointegrating Relations,” *Econometric Theory*, 16, pp. 740-778.

___ . (2002) “A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model,” *Econometrica,* 70(5), September, pp. 1929-1961.

Kao, C. (1999), "Spurious Regression and Residual-Based Tests For Cointegration in Panel data,” *Journal of Econometrics* 90, 1-44.

Kao, C and M.-H. Chiang (2000), "On the Estimation and Inference of a Cointegrated

Regression in Panel Data," *Advances in Econometrics* 15, 179-222.

Karlsson, S., and M. Lothgren (2000), "On The Power and Interpretation of Panel Unit Root Tests," *Economics Letters* 66, 249-255.

Larsson R. and J. Lyhagen (1999), "Likelihood-Based Inference in Multivariate Panel Cointegration Models," working paper No. 331, Stockholm School of Economics.

Larsson R., J. Lyhagen and M. Lothgren (2001), "Likelihood-based Cointegration in Heterogeneous Panels," *Econometrics Journal* 4, 109-142.

Levin, A. and C.F. Lin (1993), "Unit root test in Panel Data: Asymptotic and Finite-Sample Properties," Discussion Paper 93-56, Department of Economics, UCSD.

Lutkepohl, H. and D. Poskitt (1998), "Consistent estimation of the number of cointegration relations in a vector autoregressive model." In *Econometrics in Theory and Practice,* Galant, R. and H. Kuchenhoff Eds.; Physica-Verlag: Heidelberg, 87-100.

Lyhagen, J. (2000), "Why not use standard panel unit root test for testing PPP," working paper no. 413, Stockholm School of Economics.

Maddala, G.S., and S. Wu (1999), "A Comparative Study of Unit Root Tests With PanelData and a New Simple Test," *Oxford Bulletin of Economics and Statistics*, pp. 631-652.

Mark, N., M. Ogaki and D. Sul (2005), "Dynamic Seemingly Unrelated Cointegrating Regression," *Review of Economic Studies*, 72(3), pp. 797-820.

McCoskey, S., and C. Kao (2001), "A Monte Carlo Comparison of Tests for Cointegration in Panel Data," *Journal of Propagations in Probability and Statistics*, 1, pp. 165-198.

Moon, R. and B. Perron (2004), "Testing for A Unit Root in Panels with Dynamic Factors," *Journal of Econometrics,* September, 122(1), pp. 81-126.

Ng, S. and P. Perron (1997), "Estimation and Inference in Nearly Unbalanced Nearly Cointegrated Systems," *Journal of Econometrics*, 53-81.

O'Connell, P. (1998), "The overvaluation of purchasing power parity," *Journal of International Economics,* 1-19.

Pedroni, P. (1999), "Critical Values for Cointegration in Heterogeneous Panels with Multiple Regressors," *Oxford Bulletin of Economics and Statistics*, 61, 653-678.

Pedroni, P. (2004), "Panel Cointegration: Asymptotic and Finite Sample Properties of

Pooled Time Tests with an Application to the PPP Hypothesis," *Econometric Theory*, 20, pp. 597-625.

Pesaran, M.H., Y. Shin and R. Smith (2000), "Structural Analysis of VECMs with exogenous I(1) Variables," *Journal of Econometrics,* August, 97(2), pp. 293-343.

Pesaran, M.H., Y. Shin and R. Smith (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels," *Journal of American Statistical Association* 94, 621-634.

Phillips, P.C.B. (1991), "Optimal Inference in Cointegrated Systems," *Econometrica* 59, 283-306.

Phillips, P.C.B. (1995), "Fully Modified Least Squares and Vector Autoregression," *Econometrica* 63, 1023-1078.

Phillips, P. and H. Moon (2000), "Nonstationary Panel Data Analysis: An Overview of Some Recent Developments," *Econometric Reviews*, 19(3), pp. 263-286.

Shin, Y. (1994), "A residual-based test of the null of cointegration against the alternative of no cointegration," *Econometric Theory* 10, 91-115.

Stock, J. and M. Watson (1988), "Testing for Common Trends," *Journal of the American Statistical Association* 83, 1097-1107.

Taylor, M. and L. Sarno (1998), "The Behavior of Real Exchange Rates During the Post-Bretton Woods Period," *Journal of Internal Economics* 46, 281-312.

Toda, H.Y. (1995), "Finite Sample Performance of Likelihood Ratio Tests for Cointegrating Ranks in Vector Autoregressions," *Econometric Theory* 11, 1015-1032.

van Giersbergen, N.P.A. (1996), "Bootstrapping the Trace Statistic in VAR Models: Monte Carlo Results and Applications," *Oxford Bulletin of Economics and Statistics* 58(2), 391-415.

Westerlund, Joakim (2005). "New Simple Tests for Panel Cointegration," *Econometric Reviews*, 24(3), pp. 297-316.

___ . (2006). "Testing for Panel Cointegration with Multiple Structural Breaks," *Oxford Bulletin of Economics and Statistics,* forthcoming.

Westerlund, Joakim and David Edgerton (2005). "Panel Cointegration Tests with Deterministic Trends and Structural Breaks," Working Paper 2005:42, Department of Economics, Lund University, October.

Table 1A: Actual Size of LM Test

Estimation Method: Dynamic OLS (DOLS)
Sample Size (T)=50
No. of Replications (R)=5000

Nominal Size:	10%				5%			
	Case 1	Case 2	case 3	case 4	Case 1	Case 2	case 3	case 4
N=2								
pooled	0.1000	0.1114	0.0188	0.0364	0.0500	0.0618	0.0076	0.0158
unit 1	0.1000	0.1000	0.0288	0.0288	0.0500	0.0502	0.0090	0.0090
unit 2	0.1000	0.1074	0.0322	0.0324	0.0500	0.0498	0.0104	0.0126
N=5								
pooled	0.1000	0.1382	0.0124	0.0458	0.0500	0.0900	0.0058	0.0254
unit 1	0.1000	0.1002	0.0328	0.0328	0.0500	0.0502	0.0082	0.0082
unit 2	0.1000	0.1110	0.0336	0.0356	0.0500	0.0546	0.0116	0.0124
unit 3	0.1000	0.0922	0.0268	0.0266	0.0500	0.0474	0.0118	0.0114
unit 4	0.1000	0.0970	0.0286	0.0306	0.0500	0.0520	0.0118	0.0114
unit 5	0.1000	0.1026	0.0284	0.0320	0.0500	0.0494	0.0090	0.0114
N=10								
pooled	0.1000	0.1698	0.0090	0.0552	0.0500	0.1222	0.0050	0.0346
unit 1	0.1000	0.1000	0.0276	0.0276	0.0500	0.0502	0.0098	0.0098
unit 2	0.1000	0.1008	0.0348	0.0338	0.0500	0.0546	0.0114	0.0122
unit 3	0.1000	0.0886	0.0280	0.0188	0.0500	0.0478	0.0106	0.0068
unit 4	0.1000	0.1036	0.0296	0.0320	0.0500	0.0516	0.0118	0.0112
unit 5	0.1000	0.1000	0.0320	0.0308	0.0500	0.0456	0.0100	0.0092
unit 6	0.1000	0.1058	0.0324	0.0294	0.0500	0.0524	0.0084	0.0110
unit 7	0.1000	0.1050	0.0318	0.0326	0.0500	0.0550	0.0114	0.0128
unit 8	0.1000	0.1062	0.0288	0.0292	0.0500	0.0546	0.0110	0.0114
unit 9	0.1000	0.0968	0.0300	0.0272	0.0500	0.0508	0.0126	0.0092
unit 10	0.1000	0.1072	0.0322	0.0384	0.0500	0.0530	0.0108	0.0134

Table 1B: Actual Size of LM Test

Estimation Method: DOLS
T=100
R=5000

Nominal Size:	10%				5%			
	Case 1	Case 2	case 3	case 4	Case 1	Case 2	case 3	case 4
N=2								
pooled	0.1000	0.1128	0.0060	0.0194	0.0500	0.0662	0.0016	0.0074
unit 1	0.1000	0.1000	0.0164	0.0164	0.0500	0.0502	0.0046	0.0046
unit 2	0.1000	0.1052	0.0150	0.0144	0.0500	0.0532	0.0032	0.0040
N=5								
pooled	0.1000	0.1354	0.0008	0.0270	0.0500	0.0890	0.0006	0.0140
unit 1	0.1000	0.1002	0.0146	0.0146	0.0500	0.0502	0.0030	0.0030
unit 2	0.1000	0.0950	0.0124	0.0156	0.0500	0.0502	0.0032	0.0048
unit 3	0.1000	0.1066	0.0180	0.0190	0.0500	0.0574	0.0042	0.0054
unit 4	0.1000	0.0946	0.0168	0.0170	0.0500	0.0520	0.0036	0.0052
unit 5	0.1000	0.0988	0.0158	0.0142	0.0500	0.0458	0.0038	0.0028
N=10								
pooled	0.1000	0.1680	0.0002	0.0304	0.0500	0.1254	0.0000	0.0196
unit 1	0.1000	0.1002	0.0152	0.0152	0.0500	0.0500	0.0026	0.0026
unit 2	0.1000	0.0964	0.0124	0.0130	0.0500	0.0472	0.0026	0.0022
unit 3	0.1000	0.1018	0.0176	0.0144	0.0500	0.0488	0.0038	0.0050
unit 4	0.1000	0.0970	0.0156	0.0142	0.0500	0.0476	0.0024	0.0032
unit 5	0.1000	0.0980	0.0174	0.0146	0.0500	0.0480	0.0040	0.0020
unit 6	0.1000	0.0958	0.0150	0.0132	0.0500	0.0516	0.0044	0.0036
unit 7	0.1000	0.1076	0.0194	0.0154	0.0500	0.0538	0.0038	0.0038
unit 8	0.1000	0.1072	0.0142	0.0176	0.0500	0.0558	0.0036	0.0036
unit 9	0.1000	0.1002	0.0156	0.0136	0.0500	0.0500	0.0046	0.0028
unit 10	0.1000	0.0984	0.0112	0.0166	0.0500	0.0516	0.0022	0.0042

Case 1: with no cross-sectional dependence or cross-sectional cointegration
Case 2: with cross-sectional dependence but with no cross-sectional cointegration
Case 3: with no cross-sectional dependence but with cross-sectional cointegration
Case 4: with both cross-sectional dependence and cross-sectional cointegration

Table 2A: The actual power of LM Test when the nominal size is at 5%

Estimation Method: DOLS
T=50
R=5000

	Case 1	Case 2	case 3	case 4	Case 1	Case 2	case 3	case 4
Theta	0.05				0.1			
pooled	0.0970	0.1046	0.0158	0.0254	0.2178	0.2136	0.0458	0.0578
unit 1	0.0784	0.0784	0.0190	0.0190	0.1614	0.1614	0.0426	0.0426
unit 2	0.0882	0.0842	0.0176	0.0180	0.1634	0.1644	0.0438	0.0420
Theta	0.15				0.2			
pooled	0.3764	0.3534	0.0868	0.1072	0.4764	0.4546	0.1592	0.1792
unit 1	0.2544	0.2544	0.0728	0.0728	0.3356	0.3356	0.1258	0.1258
unit 2	0.2688	0.2656	0.0772	0.0774	0.3312	0.3420	0.1256	0.1242
theta	0.05				0.10			
pooled	0.1328	0.1588	0.0142	0.0420	0.3550	0.3188	0.0550	0.0946
unit 1	0.0888	0.0888	0.0174	0.0174	0.1676	0.1676	0.0416	0.0416
unit 2	0.0840	0.0860	0.0174	0.0182	0.1670	0.1524	0.0408	0.0350
unit 3	0.0820	0.0818	0.0212	0.0202	0.1604	0.1598	0.0404	0.0396
unit 4	0.0920	0.0902	0.0236	0.0200	0.1670	0.1772	0.0464	0.0552
unit 5	0.0842	0.0890	0.0198	0.0176	0.1642	0.1624	0.0440	0.0402
theta	0.15				0.2			
pooled	0.5940	0.4910	0.1364	0.1698	0.7630	0.6442	0.2368	0.2704
unit 1	0.2514	0.2514	0.0766	0.0766	0.3456	0.3456	0.1270	0.1270
unit 2	0.2534	0.2506	0.0840	0.0766	0.3280	0.3292	0.1178	0.1174
unit 3	0.2530	0.2494	0.0814	0.0796	0.3284	0.3428	0.1262	0.1212
unit 4	0.2594	0.2630	0.0900	0.0888	0.3426	0.3370	0.1340	0.1344
unit 5	0.2542	0.2486	0.0750	0.0774	0.3302	0.3444	0.1236	0.1326

theta	0.05				0.10			
pooled	0.1728	0.2048	0.0162	0.0512	0.5324	0.4330	0.0806	0.1310
unit 1	0.0920	0.0920	0.0158	0.0158	0.1676	0.1676	0.0436	0.0436
unit 2	0.0932	0.0888	0.0180	0.0188	0.1638	0.1622	0.0396	0.0406
unit 3	0.0880	0.0820	0.0186	0.0142	0.1628	0.1658	0.0406	0.0392
unit 4	0.0876	0.0814	0.0242	0.0200	0.1712	0.1724	0.0398	0.0418
unit 5	0.0822	0.0800	0.0134	0.0152	0.1508	0.1566	0.0442	0.0414
unit 6	0.0826	0.0856	0.0192	0.0186	0.1616	0.1612	0.0408	0.0422
unit 7	0.0898	0.0846	0.0182	0.0178	0.1722	0.1732	0.0506	0.0464
unit 8	0.0824	0.0790	0.0186	0.0174	0.1598	0.1640	0.0430	0.0404
unit 9	0.0856	0.0828	0.0188	0.0164	0.1628	0.1626	0.0428	0.0436
unit 10	0.0884	0.0840	0.0180	0.0166	0.1568	0.1622	0.0392	0.0426

theta	0.15				0.20			
pooled	0.8094	0.6340	0.1914	0.2362	0.9362	0.7758	0.3590	0.3470
unit 1	0.2722	0.2722	0.0878	0.0878	0.3428	0.3428	0.1304	0.1304
unit 2	0.2550	0.2510	0.0764	0.0760	0.3428	0.3350	0.1182	0.1180
unit 3	0.2468	0.2606	0.0794	0.0820	0.3274	0.3242	0.1234	0.1212
unit 4	0.2634	0.2570	0.0860	0.0834	0.3384	0.3356	0.1260	0.1266
unit 5	0.2370	0.2434	0.0732	0.0780	0.3140	0.3154	0.1184	0.1124
unit 6	0.2510	0.2482	0.0834	0.0812	0.3334	0.3350	0.1250	0.1270
unit 7	0.2480	0.2608	0.0778	0.0818	0.3382	0.3464	0.1326	0.1276
unit 8	0.2476	0.2562	0.0802	0.0816	0.3356	0.3290	0.1122	0.1116
unit 9	0.2374	0.2480	0.0780	0.0836	0.3336	0.3256	0.1222	0.1188
unit 10	0.2516	0.2484	0.0752	0.0826	0.3248	0.3160	0.1248	0.1126

Table 2B: The actual power of LM Test when the nominal size is at 5%

Estimation Method: DOLS
T=100
R=5000

	Case 1	Case 2	case 3	case 4	Case 1	Case 2	case 3	case 4
Theta	0.05				0.1			
pooled	0.2584	0.2442	0.0142	0.0298	0.5732	0.5354	0.1218	0.1498
unit 1	0.1938	0.1938	0.0200	0.0200	0.4200	0.4200	0.1066	0.1066
unit 2	0.1918	0.1902	0.0196	0.0176	0.3818	0.3992	0.1034	0.1040
Theta	0.15				0.2			
pooled	0.7472	0.6956	0.3280	0.3088	0.8598	0.8026	0.5328	0.4672
unit 1	0.5518	0.5518	0.2482	0.2482	0.6606	0.6606	0.3894	0.3894
unit 2	0.5428	0.5544	0.2302	0.2340	0.6598	0.6536	0.3678	0.3688
theta	0.05				0.10			
pooled	0.4298	0.3784	0.0150	0.0614	0.8438	0.7184	0.1966	0.2360
unit 1	0.1838	0.1838	0.0218	0.0218	0.3854	0.3854	0.1096	0.1096
unit 2	0.2060	0.2028	0.0252	0.0210	0.4094	0.4052	0.1188	0.1108
unit 3	0.1830	0.1898	0.0224	0.0190	0.3928	0.3776	0.1006	0.1034
unit 4	0.1862	0.2012	0.0226	0.0256	0.4020	0.3940	0.1154	0.1090
unit 5	0.1816	0.1826	0.0234	0.0214	0.3850	0.3874	0.1080	0.1110
theta	0.15				0.2			
pooled	0.9642	0.8880	0.5676	0.4642	0.9938	0.9526	0.8232	0.6472
unit 1	0.5480	0.5480	0.2386	0.2386	0.6610	0.6610	0.3660	0.3660
unit 2	0.5598	0.5666	0.2566	0.2634	0.6678	0.6864	0.3762	0.3820
unit 3	0.5490	0.5418	0.2322	0.2450	0.6622	0.6428	0.3634	0.3584
unit 4	0.5502	0.5536	0.2418	0.2498	0.6646	0.6640	0.3866	0.3800
unit 5	0.5498	0.5434	0.2402	0.2524	0.6552	0.6450	0.3686	0.3644

theta	0.05				0.10			
pooled	0.6308	0.4782	0.0178	0.0914	0.9736	0.8514	0.301	0.3192
unit 1	0.1974	0.1974	0.0248	0.0248	0.4116	0.4116	0.1194	0.1194
unit 2	0.197	0.1908	0.0258	0.0254	0.3924	0.403	0.1132	0.11
unit 3	0.1782	0.183	0.0226	0.0202	0.3882	0.395	0.1082	0.104
unit 4	0.1858	0.184	0.0222	0.022	0.3856	0.39	0.1032	0.1086
unit 5	0.1994	0.1886	0.0232	0.0228	0.3918	0.3854	0.1122	0.1138
unit 6	0.187	0.198	0.026	0.027	0.4014	0.3954	0.1174	0.1224
unit 7	0.1884	0.1974	0.0222	0.0218	0.3962	0.3926	0.1022	0.1152
unit 8	0.1768	0.1768	0.0212	0.0194	0.3742	0.3798	0.1026	0.1046
unit 9	0.1858	0.1834	0.0224	0.0268	0.397	0.3998	0.1064	0.1154
unit 10	0.185	0.1862	0.023	0.0244	0.4022	0.3928	0.1	0.1108

theta	0.15				0.20			
pooled	0.9994	0.9692	0.8088	0.5916	1	0.994	0.968	0.7762
unit 1	0.566	0.566	0.2634	0.2634	0.6796	0.6796	0.397	0.397
unit 2	0.56	0.564	0.2546	0.256	0.6466	0.6588	0.3786	0.3844
unit 3	0.5436	0.5494	0.2322	0.242	0.6604	0.669	0.3622	0.3692
unit 4	0.5504	0.5538	0.2406	0.2456	0.6588	0.6584	0.362	0.3678
unit 5	0.5684	0.5652	0.2454	0.2392	0.6654	0.6672	0.3786	0.3732
unit 6	0.5534	0.5628	0.257	0.26	0.6642	0.6816	0.3724	0.3936
unit 7	0.552	0.551	0.2438	0.2454	0.6612	0.6724	0.3744	0.382
unit 8	0.5394	0.5398	0.2316	0.2296	0.6524	0.6526	0.3524	0.3616
unit 9	0.5664	0.5564	0.2498	0.2518	0.6684	0.6634	0.3686	0.3824
unit 10	0.5628	0.5462	0.2502	0.2384	0.6578	0.6606	0.359	0.3802

Table 3A: Actual Size of LM Test

Estimation Method: Dynamic SUR (DSUR)
T=50
R=5000

Nominal size:		10%			5%		
		Case 1	Case 2	case 3	Case 1	Case 2	case 3
N=2	pooled	0.1000	0.0048	0.0246	0.0500	0.0006	0.0110
	unit 1	0.1000	0.0132	0.0250	0.0500	0.0048	0.0108
	unit 2	0.1000	0.0102	0.0252	0.0500	0.0036	0.0110
N=5	pooled	0.1000	0.0006	0.0456	0.0500	0.0002	0.0164
	unit 1	0.1000	0.0098	0.0444	0.0500	0.0036	0.0186
	unit 2	0.1000	0.0110	0.0456	0.0500	0.0034	0.0146
	unit 3	0.1000	0.0092	0.0442	0.0500	0.0046	0.0174
	unit 4	0.1000	0.0142	0.0320	0.0500	0.0058	0.0122
	unit 5	0.1000	0.0106	0.0462	0.0500	0.0054	0.0166
N=10	pooled	0.1000	0.0000	0.0536	0.0500	0.0000	0.0216
	unit 1	0.1000	0.0140	0.0544	0.0500	0.0040	0.0218
	unit 2	0.1000	0.0120	0.0508	0.0500	0.0048	0.0188
	unit 3	0.1000	0.0130	0.0522	0.0500	0.0044	0.0200
	unit 4	0.1000	0.0164	0.0448	0.0500	0.0056	0.0178
	unit 5	0.1000	0.0124	0.0498	0.0500	0.0042	0.0192
	unit 6	0.1000	0.0116	0.0544	0.0500	0.0030	0.0198
	unit 7	0.1000	0.0140	0.0534	0.0500	0.0050	0.0204
	unit 8	0.1000	0.0180	0.0454	0.0500	0.0048	0.0188
	unit 9	0.1000	0.0130	0.0448	0.0500	0.0060	0.0178
	unit 10	0.1000	0.0130	0.0540	0.0500	0.0058	0.0210

Case 1: with cross-sectional correlation but without cross-sectional cointegration
Case 2: without cross-sectional correlation but with cross-sectional cointegration
Case 3: with both cross-sectional correlation and cross-sectional cointegration

Table 3B: Actual Size of LM Test

Estimation Method: Dynamic SUR (DSUR)
T=100
R=5000

Nominal Size:			10%			5%	
		Case 1	Case 2	case 3	Case 1	Case 2	case 3
N=2	pooled	0.1000	0.0004	0.0112	0.0500	0.0000	0.0038
	unit 1	0.1000	0.0038	0.0136	0.0500	0.0008	0.0032
	unit 2	0.1000	0.0038	0.0096	0.0500	0.0012	0.0040
N=5	pooled	0.1000	0.0000	0.0302	0.0500	0.0000	0.0104
	unit 1	0.1000	0.0024	0.0308	0.0500	0.0004	0.0116
	unit 2	0.1000	0.0032	0.0312	0.0500	0.0004	0.0112
	unit 3	0.1000	0.0020	0.0318	0.0500	0.0006	0.0118
	unit 4	0.1000	0.0044	0.0176	0.0500	0.0006	0.0058
	unit 5	0.1000	0.0040	0.0306	0.0500	0.0010	0.0094
N=10	pooled	0.1000	0.0000	0.0432	0.0500	0.0000	0.0198
	unit 1	0.1000	0.0022	0.0468	0.0500	0.0006	0.0194
	unit 2	0.1000	0.0032	0.0322	0.0500	0.0004	0.0130
	unit 3	0.1000	0.0028	0.0426	0.0500	0.0006	0.0182
	unit 4	0.1000	0.0038	0.0330	0.0500	0.0004	0.0114
	unit 5	0.1000	0.0054	0.0404	0.0500	0.0012	0.0146
	unit 6	0.1000	0.0032	0.0414	0.0500	0.0006	0.0178
	unit 7	0.1000	0.0042	0.0442	0.0500	0.0004	0.0178
	unit 8	0.1000	0.0038	0.0330	0.0500	0.0010	0.0108
	unit 9	0.1000	0.0052	0.0358	0.0500	0.0012	0.0130
	unit 10	0.1000	0.0030	0.0360	0.0500	0.0006	0.0144

Case 1: with cross-sectional correlation but without cross-sectional cointegration
Case 2: without cross-sectional correlation but with cross-sectional cointegration
Case 3: with both cross-sectional correlation and cross-sectional cointegration

Table 4A: The actual power of LM Test when the nominal size is at 5%

Estimation Method: DSUR
T=50
R=5000

		Case 1	Case 2	case 3	Case 1	Case 2	case 3
N=2	Theta	0.0500			0.1000		
	pooled	0.1102	0.0056	0.0244	0.2500	0.0186	0.0666
	unit 1	0.1076	0.0118	0.0248	0.2426	0.0286	0.0676
	unit 2	0.1064	0.0092	0.0234	0.2394	0.0276	0.0632
	Theta	0.1500			0.2000		
	pooled	0.3928	0.0456	0.1304	0.5056	0.0808	0.2040
	unit 1	0.3794	0.0566	0.1284	0.4900	0.0946	0.1980
	unit 2	0.3724	0.0600	0.1230	0.4870	0.0980	0.1998
N=5	theta	0.05			0.10		
	pooled	0.1120	0.0014	0.0434	0.2812	0.0116	0.1400
	unit 1	0.1190	0.0106	0.0466	0.2682	0.0310	0.1390
	unit 2	0.1120	0.0098	0.0468	0.2602	0.0250	0.1292
	unit 3	0.1114	0.0086	0.0456	0.2664	0.0286	0.1422
	unit 4	0.0994	0.0122	0.0306	0.2450	0.0344	0.0966
	unit 5	0.1160	0.0106	0.0418	0.2662	0.0270	0.1356
	theta	0.1500			0.2000		
	pooled	0.4408	0.0390	0.2570	0.5698	0.0776	0.3652
	unit 1	0.4072	0.0594	0.2536	0.5114	0.0984	0.3520
	unit 2	0.4006	0.0528	0.2398	0.5114	0.0900	0.3418
	unit 3	0.4054	0.0544	0.2542	0.5040	0.0942	0.3490
	unit 4	0.3810	0.0636	0.1894	0.4822	0.1054	0.2866
	unit 5	0.3988	0.0528	0.2454	0.5074	0.0910	0.3444

		Case 1	Case 2	case 3	Case 1	Case 2	case 3
N=10	theta	0.05			0.10		
	pooled	0.1134	0.0002	0.0604	0.2928	0.0046	0.1774
	unit 1	0.1068	0.0074	0.0568	0.2568	0.0238	0.1604
	unit 2	0.1102	0.0114	0.0442	0.2544	0.0316	0.1430
	unit 3	0.1062	0.0102	0.0564	0.2590	0.0300	0.1616
	unit 4	0.1036	0.0130	0.0398	0.2502	0.0340	0.1284
	unit 5	0.1062	0.0080	0.0490	0.2566	0.0272	0.1520
	unit 6	0.1066	0.0098	0.0542	0.2574	0.0304	0.1616
	unit 7	0.1068	0.0110	0.0532	0.2600	0.0278	0.1546
	unit 8	0.1022	0.0110	0.0448	0.2470	0.0334	0.1374
	unit 9	0.1102	0.0130	0.0450	0.2506	0.0336	0.1334
	unit 10	0.1098	0.0102	0.0486	0.2590	0.0260	0.1492
	theta	0.1500			0.20		
	pooled	0.4690	0.0336	0.3096	0.6246	0.0798	0.4314
	unit 1	0.3914	0.0514	0.2826	0.5042	0.0902	0.3870
	unit 2	0.3950	0.0606	0.2610	0.4984	0.0950	0.3656
	unit 3	0.3992	0.0594	0.2840	0.4998	0.1022	0.3902
	unit 4	0.3890	0.0658	0.2426	0.4946	0.1096	0.3444
	unit 5	0.3958	0.0548	0.2714	0.5070	0.0946	0.3730
	unit 6	0.3914	0.0622	0.2800	0.5026	0.0992	0.3852
	unit 7	0.4002	0.0570	0.2774	0.5072	0.0962	0.3838
	unit 8	0.3904	0.0642	0.2574	0.4974	0.1014	0.3612
	unit 9	0.3956	0.0630	0.2572	0.4946	0.1062	0.3540
	unit 10	0.3882	0.0552	0.2646	0.4984	0.0936	0.3662

Table 4B: The actual power of LM Test when the nominal size is at 5%

Estimation Method: DSUR
T=100
R=5000

		Case 1	Case 2	case 3	Case 1	Case 2	case 3
	Theta	0.0500			0.1000		
N=2	pooled	0.2660	0.0022	0.0300	0.5472	0.0390	0.1452
	unit 1	0.2560	0.0094	0.0294	0.5326	0.0612	0.1424
	unit 2	0.2640	0.0110	0.0314	0.5412	0.0648	0.1458
	Theta	0.1500			0.2000		
	pooled	0.7202	0.1380	0.3126	0.8144	0.2936	0.4628
	unit 1	0.6948	0.1686	0.2964	0.7946	0.2810	0.4436
	unit 2	0.7028	0.1716	0.3082	0.7984	0.2910	0.4596
N=5	theta	0.05			0.10		
	pooled	0.3046	0.0012	0.1126	0.6218	0.0160	0.3546
	unit 1	0.2858	0.0074	0.1114	0.5562	0.0546	0.3454
	unit 2	0.2810	0.0090	0.1090	0.5550	0.0618	0.3396
	unit 3	0.2844	0.0074	0.1112	0.5476	0.0492	0.3432
	unit 4	0.2626	0.0134	0.0726	0.5364	0.0666	0.2710
	unit 5	0.2872	0.0068	0.1078	0.5620	0.0532	0.3410
	theta	0.15			0.20		
	pooled	0.8112	0.0982	0.5604	0.8978	0.2766	0.6916
	unit 1	0.7188	0.1630	0.5370	0.7980	0.2832	0.6564
	unit 2	0.7062	0.1564	0.5292	0.7918	0.2712	0.6560
	unit 3	0.7066	0.1408	0.5338	0.7936	0.2510	0.6548
	unit 4	0.6918	0.1798	0.4710	0.7810	0.2918	0.6026
	unit 5	0.7096	0.1498	0.5302	0.7944	0.2634	0.6544

		Case 1	Case 2	case 3	Case 1	Case 2	case 3
N=10	theta	0.05			0.10		
	pooled	0.3242	0.0000	0.1884	0.6808	0.0132	0.4630
	unit 1	0.2936	0.0094	0.1818	0.5522	0.0616	0.4384
	unit 2	0.2848	0.0084	0.1492	0.5434	0.0704	0.3980
	unit 3	0.2866	0.0090	0.1694	0.5460	0.0606	0.4258
	unit 4	0.2812	0.0100	0.1298	0.5450	0.0678	0.3708
	unit 5	0.2882	0.0080	0.1620	0.5578	0.0616	0.4246
	unit 6	0.2866	0.0094	0.1652	0.5414	0.0544	0.4278
	unit 7	0.2912	0.0084	0.1732	0.5556	0.0624	0.4244
	unit 8	0.2754	0.0088	0.1358	0.5348	0.0518	0.3948
	unit 9	0.2854	0.0096	0.1418	0.5526	0.0642	0.3994
	unit 10	0.2806	0.0092	0.1552	0.5478	0.0660	0.4102
	theta	0.15			0.20		
	pooled	0.8888	0.0890	0.6554	0.9640	0.3372	0.7734
	unit 1	0.7044	0.6820	0.6094	0.7918	0.2896	0.7158
	unit 2	0.6918	0.1776	0.5768	0.7824	0.3014	0.6824
	unit 3	0.6932	0.1590	0.6012	0.7796	0.2758	0.7058
	unit 4	0.6952	0.1806	0.5474	0.7748	0.2958	0.6718
	unit 5	0.7010	0.1662	0.5936	0.7870	0.2894	0.6974
	unit 6	0.6968	0.1596	0.5912	0.7826	0.2694	0.7024
	unit 7	0.7064	0.1692	0.6028	0.7932	0.2904	0.7102
	unit 8	0.6862	0.1450	0.5676	0.7798	0.2600	0.6796
	unit 9	0.7008	0.1770	0.5842	0.7880	0.2916	0.6950
	unit 10	0.7042	0.1668	0.5862	0.7840	0.2924	0.6954

Table 5

Johansen's Tests for Cointegration Rank Based on Asymptotic Distributions

Sample (adjusted): 1922 1980
Included observations: 59 after adjustments
Trend assumption: Linear deterministic trend
Series: LV_US ILONG_US LV_CAN ILONG_CAN
Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.420569	66.51895	47.85613	0.0004
At most 1 *	0.335045	34.32215	29.79707	0.0141
At most 2	0.157967	10.24806	15.49471	0.2621
At most 3	0.001758	0.103817	3.841466	0.7473

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.420569	32.19680	27.58434	0.0118
At most 1 *	0.335045	24.07409	21.13162	0.0187
At most 2	0.157967	10.14425	14.26460	0.2026
At most 3	0.001758	0.103817	3.841466	0.7473

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Table 6

A. Bootstrap Distribution of Johansen's Maximum-Eigenvalue Statistic

Hypothesized CI Rank	Empirical Values	Bootstrapped Percentiles								
		1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
r=0	32.20	11.99	13.13	14.23	15.71	22.10	31.37	34.32	37.57	41.26
r=1	24.07	7.27	8.20	9.14	10.30	15.57	22.78	25.08	27.06	29.75
r=2	10.14	2.81	3.39	4.00	4.80	8.39	13.46	15.10	16.52	18.22

Notes:
When the eigenvalues are ordered in ascending order and the hypothesized CI rank is r, the maximum eigenvalue statistic is: $\lambda_{\max} = -T \ln(1-\hat{\lambda}_{4-r})$. It is an upper tail test.

B. Bootstrap Distribution of Johansen's Trace Statistic

Hypothesized CI Rank	Empirical Values	Bootstrapped Percentiles								
		1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
r=0	66.52	25.68	27.91	29.99	32.41	42.94	56.44	60.62	64.69	69.96
r=1	34.32	12.40	13.93	15.31	17.04	24.71	34.55	37.86	40.20	44.04
r=2	10.25	3.33	4.06	4.81	5.76	10.36	16.73	18.85	20.64	23.31

Notes:
When the eigenvalues are ordered in ascending order and the hypothesized CI rank is r, the trace statistic is: $\lambda_{trace} = -T \sum_{i=1}^{4-r} \ln(1-\hat{\lambda}_i)$. It is an upper tail test.

Table 7

A. Bootstrap Distributions of the CC-based Trace Statistic

Hypothesized CI Rank	Empirical Values	Bootstrapped Percentiles								
		1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
r=0	94.84	38.97	43.01	46.40	50.88	69.18	88.67	93.77	98.45	104.40
r=1	51.31	13.52	15.51	17.59	20.32	32.91	48.53	53.19	56.86	61.25
r=2	25.28	2.68	3.22	3.84	4.69	9.80	19.38	22.53	25.41	27.98
r=3	1.08	0.24	0.30	0.38	0.49	1.28	3.25	4.17	4.99	6.41

Notes:

When eigenvalues are ordered in ascending order and the hypothesized CI rank is r, the CC-based trace statistic is defined as $\lambda_{trace} = T \sum_{i=r+1}^{4} (1 - \hat{\lambda}_i)$.

B. Bootstrap Distributions of Eigenvalues

Estimated Eigenvalues	Empirical Values	Bootstrapped Percentiles								
		1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
$\hat{\lambda}_1$	.2621	.0005	.0039	.0118	.0273	.1283	.2677	.3120	.3482	.3860
$\hat{\lambda}_2$	.5589	.1392	.1791	.2140	.2588	.4298	.5959	.6384	.6736	.7075
$\hat{\lambda}_3$	.5897	.5480	.6075	.6547	.7078	.8618	.9403	.9533	.9616	.9689
$\hat{\lambda}_4$	.9817	.8929	.9142	.9311	.9451	.9788	.9920	.9939	.9951	.9962