1-1-1999

A Monte Carlo Comparison of Tests for Cointegration in Panel Data

Chihwa Kao
Syracuse University. Center for Policy Research, cdkao@maxwell.syr.edu

Center for Policy Research
Working Paper No. 3

A MONTE CARLO COMPARISON OF TESTS FOR COINTEGRATION IN PANEL DATA

Suzanne McCoskey and Chihwa Kao

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
Syracuse, New York 13244-1020

March 1999

$5.00

CENTER FOR POLICY RESEARCH - SPRING 1999

Timothy M. Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret M. Austin
Associate Director,
Budget and Administration

Douglas Holtz-Eakin
Chair, Professor of Economics
Associate Director, Center for Policy Research

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Stacy Dickert-Conlin.................................... Economics
William Duncombe.......................Public Administration
Thomas Dunn .... Economics
Vernon Greene ....Public Administration
David Greytak.. Economics
Leah Gutierrez.............................Public Administration
Christine Himes.. Sociology
Jacqueline Johnson...................................... Sociology
Bernard Jump..............................Public Administration
Duke Kao...Economics
Jeff Kubik..Economics
Madonna Harrington Meyer ....Sociology
Jerry Miner..Economics
Hisahiro Naito...Economics
Jan Ondrich ....Economics
John Palmer ....Public Administration
Grant Reeher.......................................Political Science
Jodi Sandfort ....Public Administration
Michael Wasylenko.....................................Economics
Douglas Wolf...............................Public Administration
Assata Zerai ....Sociology

RESEARCH ASSOCIATES

Andrew Houtenville...Economics
Diane Macunovich ....Economics

GRADUATE ASSOCIATES

Matthew Andrews........................Public Administration
Hakan Aykan................................Public Administration
Reagan Baughman ....Economics
Robert Bifulco...............................Public Administration
Don Bruce ....Economics
Christine Caffrey... Sociology
Noelle Chesley.............................Public Administration
Mohua Das..Economics
Deb Dougherty... Sociology
Seth Giertz...Economics
Tess Heintze ....Public Administration
Tim Hunter...................................Public Administration
Pam Herd ....Sociology
Garrett Ingoglia............................Public Administration
Kwangho Jung ....Public Administration
James Laditka..............................Public Administration
Donald Marples...Economics
Christine Muller...............................Economics/Biology
Katherin Ross.......................................Social Science
Shalini Sharma...Economics
Hava Tillipman...Sociology
Mehmet Serkan Tosun...............................Economics
Robert Weathers...Economics
Kristina Wolff...Sociology

STAFF

Martha Bonney.....................................Publications and Events Coordinator
Karen Cimilluca...................Librarian/Office Coordinator
Kati Foley........................... Administrative Assistant, LIS
Esther Gray ....Administrative Secretary
Detlef Jurkat .... Translator
Anne Metzger-Wormuth.............................Receptionist
Denise Paul......................Administrative Assistant, NTJ
Annie Pennella............Editorial Assistant, Gerontologist
Mary Santy... Secretary
Debbie Tafel............................Secretary to the Director
Ann Wicks...................... Administrative Secretary
Lobrenzo Wingo................. Computer Consultant

A Monte Carlo Comparison of Tests for Cointegration in Panel Data

Suzanne McCoskey
United States Naval Academy

Chihwa Kao*
Syracuse University

First Draft: August 1, 1997
This Draft: June 20, 1998

Abstract

This paper surveys recent developments and provides Monte Carlo comparison on various tests proposed for cointegration in panel data. In particular, tests for two panel models, varying intercepts and varying slopes and varying intercepts and common slopes, are presented from the literature with a total of seven tests being simulated. In all cases, results on empirical size and size-adjusted power are given.

Key Words and Phrases: Panel Cointegration.

JEL classifications: C23, C22.

1 Introduction

Evaluating the statistical properties of data along the time dimension has proven to be very different from analysis in the cross-section dimension. As economists have gained access to better data with more observations across time, understanding these properties has grown increasingly important. An area of particular concern in time series econometrics is the use of non-stationary data. With the desire to study the behavior of a cross-section of observations over time and the increased use of panel data, one new research area is examining the properties of non-stationary time series data in panel form. It is an intriguing question to ask: how exactly does this hybrid style of data combine the statistical elements of traditional cross-sectional analysis and time series analysis? In particular, what is the correct way to analyze non-stationarity, the spurious regression problem, and cointegration in panel data?

Two comprehensive overviews of the econometrics of panel data have been published, Hsiao (1986) and Baltagi (1995), yet neither of these books deal with the issues of non-stationarity and cointegration within panel data. Adding the cross-section dimension to the time dynamics offers a real advantage in the testing

*We thank participants of the Eighth International Conference on Panel Data, Göteborg 1998, for helpful comments. An electronic version of the paper in postscript format can be retrieved from http://web.syr.edu/~cdkao. Address correspondence to: Chihwa Kao, Center for Policy Research, 426 Eggers, Syracuse University, Syracuse, NY 13244-1020; e-mail: cdkao@maxwell.syr.edu.

for non-stationarity and cointegration. The hope of the econometrics of non-stationary panel data is to combine the best of both worlds: the method of dealing with non-stationary data from the time series and the increased data and power from the cross-section. The addition of the cross-section dimension, under certain assumptions, can act as repeated draws from the same distribution. Thus as the time and cross-section dimension increase, e.g., using the sequential limit theory or the joint limit theory of Phillips and Moon (1997), panel test statistics can be derived which converge in distribution to normally distributed random variables. Also within the testing framework, the addition of the cross-section dimension seemingly adds power to the tests.

The challenge in taking advantage of these properties is the difficulty in deriving the moments of the complex combinations of Brownian bridges and functionals of Brownian motion which often arise from the asymptotics in the time series literature. Several of the tests discussed in this paper use Monte Carlo simulations, as in the pure time series literature, to pin down these moments. Another difficulty which does not disappear in the panel setting is the difficulty in obtaining good estimates of long-run autocovariances. Finally, the panel setting offers a variety of models: common intercepts, common slopes, common intercepts and common slopes, differing intercepts and differing slopes; which have strong consequences for the estimation. In particular, asymptotics and estimation of common slopes is difficult. Also, the homogeneity or heterogeneity of the deterministic time structure of the cross-sectional observations needs to be considered.

Unit root tests in the literature test the stationarity of a given series. These tests can be adapted for residual-based cointegration tests by testing the series of estimated residuals for stationarity. There are unit root tests for panel data already in the literature such as Levin and Lin (1993), Im, Pesaran and Shin (1995) and Maddala and Wu (1996). Once again, as in the time series case, moving from the unit root tests to cointegration tests is complicated by the estimation. The cointegration tests which test the null hypothesis of no cointegration must take into consideration the so-called "spurious regression" problem. Tests based on the null hypothesis of cointegration must take into consideration efficient estimation of a cointegrated relationship. Further, the concept of "pooled" estimation is different from pooling the cross-section testing results. In the case of unit root testing, most tests treat each individual cross-section independently. In the case of cointegration, treating each cross-section independently may translate into allowing for varying slopes and varying intercepts. This has strong implications for the model.

This paper outlines and compares three recent studies which present panel data tests for cointegration: Kao (1997), Pedroni (1997) and McCoskey and Kao (1998). The first two articles present tests of the null of no cointegration and the last a test of the null of cointegration.

At least a brief mention should be given to the dynamics of the relationship between time series econometrics and economic theory. Most developments within the time series literature have been criticized as having more to do with a particular data set than economic theory in general. To be sure, there is no economic theory behind the techniques used to estimate lag orders for autoregressive representations, for example. Yet the cointegration literature offers a promising cross-over between economic theory and econometric techniques. The error-correction form, for example, captures short run deviations from the long run relationship between non-stationary variables. While this theory has not been totally resolved, it at least gives a practical motivation for theorists in all fields to familiarize themselves with applied time series techniques. *The Economic Journal* (Jan 1997) includes a discussion of the philosophy of modelling the long run using time series and cointegration results. Included in the journal are articles by Taylor and Dixon, Granger, Pesaran, and Harvey. The major issue mentioned for panel data in this discussion is the issue of how similar are the cross-sections in the panel and the difficulties of pooling heterogeneous cross-sections.

The paper is outlined as follows: Section 2 introduces tests of the null hypothesis of no cointegration in panel data with varying intercepts and common slopes. One test is currently in the literature and the other proposed for the first time here. Section 3 summarizes tests of the null hypothesis of no cointegration in panel data with varying intercepts and varying slopes. A total of four tests are presented. Section 4 presents a test of the null hypothesis of cointegration in panel data. This test assumes varying intercepts and varying slopes. Section 5 explains the Monte Carlo design for the comparison of the tests. Section 6 summarizes the results of the Monte Carlo experiment and Section 7 provides some concluding thoughts.

A word on notation used throughout the paper: integrals like $\int_0^1 W(s)ds$ as $\int W$ are used when there is no ambiguity over limits, $\Omega^{1/2}$ is defined as any matrix such that $\Omega = \left(\Omega^{1/2}\right)\left(\Omega^{1/2}\right)'$, $\xrightarrow{p}$ is used to denote convergence in probability, $\Rightarrow$ to denote weak convergence, $I(1)$ to signify a time series that is integrated of order one, and $BM\left(\Omega\right)$ to denote Brownian motion with covariance matrix Ω. All limits in this paper are taken as $T \to \infty$ and followed by $N \to \infty$ sequentially of Phillips and Moon (1997), except where otherwise noted.

2 Testing for Cointegration in Panels with the Null Hypothesis of No Cointegration: Varying Intercepts and Common Slopes

The first residual-based tests of cointegration in both the times series and panel data literature were based on the null hypothesis of no cointegration. These tests are based on the principle of deciding whether or not the error process of the regression equation is stationary. This section presents tests of the null hypothesis of no cointegration for panel data assuming common slopes. Because the tests are residual-based tests, obtaining good estimates of the residuals is the first necessary step in obtaining a good residual-based test. The asymptotic properties of the residual-based tests will depend on the asymptotics of the estimators. The tests are derived under the assumption of a spurious regression and are based on OLS estimation.

2.1 Kao (1997)

A well known result from the time series literature is that regressing a non-stationary variable on a vector of non-stationary variables may lead to spurious regression results. In Kao (1997) results are offered for the asymptotics of spurious regression within a panel data setting. The specification of the panel model allows for differing intercepts across cross-sections and common slopes. Further, the long-run variance covariance matrix is assumed the same for all cross-section observations.

Within the time series literature (e.g., Phillips, 1986) it has been shown that with a spurious regression: (a) the OLS estimator converges to a random variable; which implies (b) the OLS estimator is not consistent; and (c) the t-statistic diverges. The consequence of these properties is that a spurious regression would tend to show an apparently significant relationship even if the variables are generated independently. The results for least square dummy variable (LSDV) estimation of panel data are somewhat more encouraging. Kao showed that (a) the addition of the cross-section dimension allows that an appropriate normalization of the estimated parameter converges in distribution to a normal, mean zero, random variable; (b) even though the model is misspecified the LSDV estimator is consistent; and (c) the t-statistic still diverges.

These asymptotics on the spurious regression are crucial for testing the null of no cointegration. Under the null hypothesis of no cointegration the residuals required for the test need to be estimated, by construction, from a spurious regression. The residual based test is equivalent to testing for a unit root in the LSDV estimated residuals. Using the panel model, the Dickey-Fuller (DF) and augmented Dickey Fuller (ADF) test statistics, after appropriate normalizations will converge in distribution to random variables with normal distributions.

Kao presents two sets of specifications for the DF test statistics. The first set of test statistics depends directly on consistent estimation of long run parameters. The second set of test statistics does not.

The DF type test from Kao follows the following model:

$$y_{it} = \alpha_i + \beta x_{it} + e_{it},\ i = 1, ..., N,\ t = 1, ..., T \tag{1}$$

$$y_{it} = y_{it-1} + u_{it} \tag{2}$$

$$x_{it} = x_{it-1} + \varepsilon_{it}. \tag{3}$$

As both y_{it} and x_{it} are random walks, it follows that under the null hypothesis of no cointegration, the residual series, e_{it} should be non-stationary. The model has varying intercepts across the cross-section observations, the fixed effects specification, and common slopes across i. With this model, the DF test can be calculated from the estimated residuals as:

$$\hat{e}_{it} = \rho \hat{e}_{it-1} + \nu_{it}, \tag{4}$$

where $\hat{e}_{it}$ is the estimated residual of (1).

To test the null hypothesis of a non-stationarity, the null can be written as $H_0 : \rho = 1$. The OLS estimate of ρ is given by:

$$\hat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{it} \hat{e}_{it-1}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{it}^2}. \tag{5}$$

Kao provides the following asymptotic results:

$$\sqrt{N}T(\hat{\rho} - 1) - \sqrt{N}\frac{\mu_{5T}}{\mu_{6T}} \Rightarrow N(0, 3 + \frac{7.2\sigma_v^4}{\sigma_{ov}^4}),$$

and

$$t_\rho - \frac{\sqrt{N}\mu_{5T}}{s\sqrt{\mu_{6T}}} \Rightarrow N(0, \frac{\sigma_{ov}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{ov}^2}),$$

where

$$\mu_{5T} = E[\frac{1}{T}\Sigma_{t=2}^{T} \hat{e}_{it-1} \Delta \hat{e}_{it-1}],$$

$$\mu_{6T} = E[\frac{1}{T^2}\Sigma_{t=2}^{T} \hat{e}_{it-1}^2],$$

$$\sigma_{0v}^2 = \sigma_{0u}^2 - \frac{\sigma_{0u\varepsilon}^2}{\sigma_{0\varepsilon}^2}$$

and

$$\sigma_v^2 = \sigma_u^2 - \frac{\sigma_{u\varepsilon}^2}{\sigma_\varepsilon^2}.$$

The limiting distributions have two very nice features: they are both asymptotically, normally distributed at mean zero. However, they also contain nuisance parameters in the distributions which are present because of possible long run weak exogeneity and serial correlation in the errors. As in most of the time series literature, good estimates of these long run parameters are necessary. If $w_{it} = (u_{it}, \varepsilon_{it})'$, estimates of these nuisance parameters would be based on the long-run variance covariance matrix of w_{it}. Note that in the special case with one regressor, Ω will be a 2×2 matrix, with X_{it} as a $T \times k$ matrix, Ω would be a $(k+1)\times(k+1)$ dimension matrix with $w_{it} = (u_{it}, \varepsilon_{it})'$.

Define

$$\Omega = \lim_{T\to\infty} \frac{1}{T} E(\sum_{t=1}^{T} w_{it})(\sum_{t=1}^{T} w_{it})' = \Sigma + \Gamma + \Gamma' = \begin{bmatrix} \sigma_{ou}^2 & \sigma_{ou\varepsilon} \\ \sigma_{ou\varepsilon} & \sigma_{o\varepsilon}^2 \end{bmatrix}, \tag{6}$$

$$\Gamma = \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T-1} \sum_{t=k+1}^{T} E(w_{it} w_{it-k}') = \begin{bmatrix} \Gamma_u & \Gamma_{u\varepsilon} \\ \Gamma_{u\varepsilon} & \Gamma_\varepsilon \end{bmatrix}, \tag{7}$$

and

$$\Sigma = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E(w_{it} w_{it}') = \begin{bmatrix} \sigma_u^2 & \sigma_{u\varepsilon} \\ \sigma_{u\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix}. \tag{8}$$

In this framework, Σ, can be thought of as the contemporaneous correlation and Γ as the correlation across time. A special case of this long run relationship is when there is strong exogeneity and no serial correlation. In that case, $\Gamma = 0$, and $\sigma_u^2 = \sigma_{ou}^2 = \sigma_v^2 = \sigma_{ov}^2$. This definition of the long run variance covariance matrix is used throughout the time series literature and is assumed for the entirety of the paper.

The second test from Kao is the ADF type of the regression:

$$\hat{e}_{it} = \rho \hat{e}_{it-1} + \sum_{j=1}^{p} \vartheta_j \Delta \hat{e}_{it-j} + v_{itp}. \tag{9}$$

The lags are added in the ADF specification to take care of possible autocorrelation and the number of lags, p, should be chosen such that the residual series, v_{itp}, is not serially correlated with past errors. In this case, the test statistic for the null hypothesis of no cointegration should be based on the t-statistic for $\rho = 1$.

$$t_{ADF} = (\hat{\rho} - 1) \frac{\left[\Sigma_{i=1}^{N} (e_i' Q_i e_i)\right]^{\frac{1}{2}}}{s_v}, \tag{10}$$

and

$$Q_i = I - X_{ip}(X_{ip}' X_{ip})^{-1} X_{ip}',$$

where X_{ip} is the matrix of observations on p regressors ($\Delta\hat{e}_{it-1}$, $\Delta\hat{e}_{it-2}$, ..., $\Delta\hat{e}_{it-p}$),

$$s_v^2 = \frac{1}{NT} \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \hat{v}_{itp}^2,$$

and e_i is a vector of $\hat{e}_{it-1}$.

Kao gives the asymptotic results for the ADF type test:

$$t_\rho - \frac{\sqrt{N}\mu_{7T}}{s\sqrt{\mu_{6T}}} \Rightarrow N(0, \frac{\sigma_{ov}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{ov}^2}),$$

where

$$\mu_{7T} = E\left[e_i' Q_i v_i\right]$$

and

$$\mu_{8T} = E[\frac{1}{T^2} e_i' Q_i e_i].$$

The corrected t-statistic has the same asymptotic distribution as the DF type test. Again, the limiting distribution is based on nuisance parameters. To summarize, Kao compares the following five tests through Monte Carlo simulation:

$$DF_\rho = \frac{\sqrt{N}T(\hat{\rho}-1)+3\sqrt{3}}{\sqrt{10.2}},$$

$$DF_t = \sqrt{1.25}t_\rho + \sqrt{1.875}N,$$

$$DF_\rho^* = \frac{\sqrt{N}T(\hat{\rho}-1)+\frac{3\sqrt{N}\hat{\sigma}_v^2}{\sigma_{ov}^2}}{\sqrt{3+\frac{7.2\hat{\sigma}_v^4}{\hat{\sigma}_{ov}^4}}},$$

$$DF_t^* = \frac{t_\rho + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{ov}}}{\sqrt{\frac{\hat{\sigma}_{ov}^2}{2\hat{\sigma}_v^2}+\frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{ov}^2}}},$$

and

$$ADF = \frac{t_{ADF} + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{ov}}}{\sqrt{\frac{\hat{\sigma}_{ov}^2}{2\hat{\sigma}_v^2}+\frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{ov}^2}}}.$$

We expect that DF_ρ^*, DF_t^* and ADF will converge to $N(0,1)$ in distribution. DF_ρ and DF_t are based on the results of assuming strong exogeneity of the regressor and error and no autocorrelation. These tests do not require estimates of the long-run variance-covariance matrix as the others do.

These statistics are derived from the asymptotic results of the paper and are based on convergence as $T \rightarrow \infty$:

$$\mu_{5T} \xrightarrow{p} -\frac{\sigma_v^2}{2}, \quad \mu_{7T} \xrightarrow{p} -\frac{\sigma_{ov}^2}{2}$$
$$\mu_{6T} \xrightarrow{p} \frac{\sigma_{ov}^2}{6}, \quad \mu_{8T} \xrightarrow{p} \frac{\sigma_{ov}^2}{6}.$$

The important empirical size results of the Monte Carlo tests for these residual based tests using one-sided standard normal critical values are the following: all of the tests show size distortions when T is small; the DF test statistics which utilize consistent estimation of long-run parameters outperform the other tests in terms of size distortion. The results for unadjusted power are: all tests have small power with small T and N; with T increased to at least 25, the DF statistics which use the long run estimates dominate even the ADF. When looking at the robustness of the tests across specifications for a moving average component, variance and cross-correlations, the distributions for the ADF and DF statistics with appropriate long run normalizations can be far from the standard normal distributions predicted by theory. Therefore an important conclusion of the paper is that the DF statistics which do not depend on the estimation of long-run parameters are much more robust to different specifications in the data. This result is due to the difficulty of obtaining good results for these long run estimates under different specifications in the sample sizes feasible for applied research work.

The residual-based tests presented in Kao depend on estimates of the long run variance-covariance matrix to correct for nuisance parameters once the limiting distributions have been found. Another approach to testing has been to adapt an approach where the variables are corrected for the long-run effects before the test statistics are calculated. These test statistics have the advantage that their limiting distributions are free of nuisance parameters.[1] In the next section we propose non-parametric corrections to the ADF t-statistic test proposed by Kao which will allow the limiting distribution to be free of nuisance parameters.

2.2 Corrected Panel ADF Estimator

The common slopes ADF test has been shown in Kao (1997) to have nuisance parameters in the limiting distribution. It is an unfortunate consequence of the additional cross section dimension that although it has a very desirable property of "smoothing" the limiting distribution into a normal distribution, the additional dimension adds the problem of nuisance parameters. In this section we propose corrections which can take advantage of the normal distribution, but also cleanse the limiting distribution.

Recall the Kao's panel ADF statistic has the following limiting distribution:

$$t_{ADF} - \frac{\sqrt{N}\mu_{7T}}{s_v\sqrt{\mu_{8T}}} \Rightarrow N(0, \frac{\sigma_{ov}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{ov}^2}),$$

which contains nuisance parameters. Using the similar approach from Kao (1997), some adjustments to the test statistic can be made to remove these nuisance parameters in the limiting distribution.

In Kao (1997) t_{ADF} can be written as

$$t_{ADF} = \frac{\sqrt{N}\xi_{7T}}{s_v\sqrt{\xi_{8T}}},$$

where ξ_{7T} and ξ_{8T} are defined in Kao (1997). The following adjustments can be made

[1] Pedroni (1997) proposes two types of tests of common slopes with these types of non-parametric corrections but with stronger restrictions than the tests in Kao. The first test restricts the intercept to be equal to zero for all the cross-sections. The second set of statistics is derived under the additional assumption that all regressors are strictly exogenous.

$$\frac{\sigma_v}{s_v}(\frac{\sqrt{N}(\xi_{7T}-\lambda)}{\sigma_{ov}\sqrt{\xi_{8T}}}+\frac{\sqrt{6N}}{2}) \Rightarrow N(0,\frac{4}{5})$$

which is free of nuisance parameters, where

$$\lambda = \frac{(\sigma_{ov}^2-\sigma_v^2)}{2}.$$

The rationale is as follows. Define $\xi_{7T}^{+} = \frac{1}{\sigma_{ov}^2}(\xi_{7T}-\lambda)$. Whereas

$$\xi_{7T} \Rightarrow -d(1)\frac{\sigma_{ov}^2}{2}[V_i^*(1)^2 - \frac{\sigma_v^2}{\sigma_{ov}^2}] + d(1)\sigma_{ov}^2 V_i^*(1)\int_0^1 V_i^*(r)dr,$$

It is clear that

$$\frac{\sigma_{ov}^2}{2}[V_i^*(1)^2 - \frac{\sigma_v^2}{\sigma_{ov}^2}] = \frac{\sigma_{ov}^2}{2}[V_i^*(1)^2 - 1] + \lambda;$$

thus subtracting λ and normalizing by σ_{ov}^2 results in $E[\xi_{7i}^{+}] = -d(1)\frac{1}{2} = \mu_{7T}^{+}$ and $Var[\xi_{7i}^{+}] = -d(1)\frac{1}{12} = \sigma_{7T}^{+2}$. Similarly, define $\xi_{8T}^{+} = \frac{\xi_{8T}}{\sigma_{ov}^2}$ such that $E[\xi_{8T}^{+}] = \frac{1}{6} = \mu_{8T}^{+}$ and $Var[\xi_{8T}^{+}] = \frac{1}{45} = \sigma_{8T}^{+2}$. Finally, as $s_v^2 \xrightarrow{p} d^2(1)\sigma^2$, the last step is to normalize s_v by dividing it by $\hat{\sigma}_v$.

The term $d(1)$ is introduced with possible autocorrelation allowed for in the ADF test. From Phillips and Ouliaris (1990) this concept is explained with the following assumptions. Suppose ϖ_{it}, an error process, can be represented $\varpi_{it} = \Sigma_{j=-\infty}^{j=\infty} d_j w_{t-j}$ with the condition that $\Sigma_{j=-\infty}^{j=\infty} \|d_j\| < \infty$, then $d(1) = \Sigma_{j=-\infty}^{j=\infty} d_j$.

Using the logic from Appendix E in Kao (1997), the adjustments have the following effects:

$$(\frac{\sqrt{N}\xi_{7T}^{+}}{\frac{s_v}{\sigma_v}\sigma_{ov}\sqrt{\xi_{8T}^{+}}} - \frac{\sqrt{N}\mu_{7T}^{+}}{\frac{s_v}{\sigma_v}\sqrt{\mu_{8T}^{+}}}) \Rightarrow N(0,\frac{\sigma_{7T}^{+2}}{\mu_{8T}^{+}}+\frac{1}{4}\frac{\mu_{7T}^{2+}\sigma_{8T}^{+2}}{\mu_{8T}^{+3}})$$

which becomes

$$\frac{\sigma_v}{s_v}(t_{ADF} - \frac{\sqrt{N}\lambda}{\sigma_{ov}\sqrt{\xi_{8T}}}+\frac{\sqrt{6N}}{2}) \Rightarrow N(0,\frac{4}{5}).$$

3 Testing for Cointegration in Panels with the Null Hypothesis of No Cointegration: Varying Intercepts and Varying Slopes

In the last section, we discussed tests of the null hypothesis of no cointegration which assumes common slopes across the cross section. The links between the cointegrating vector and economic theory are immediate. Therefore, such an assumption of common slopes assumes a form of homogeneity in the relationship of the variables (allowing for heterogeneity only in the intercepts.) In this section we relax that assumption and allow intercepts and slopes to vary across the cross-sectional observations. Thus under the H_a each cross section can have a unique cointegrating vector.

3.1 Average Augmented Dickey-Fuller Test for Varying Slopes

Kao (1997) proposes an ADF test for common slopes and varying intercepts. Here we propose an ADF test for varying slopes and varying intercepts.

$$y_{it} \quad = \quad \alpha_i + x_{it}^{'}\beta_i \quad + \quad e_{it}, \ i = 1, ..., N, \ t = 1, ..., T. \tag{11}$$

Each cross-section regression is estimated individually and the pooling from the panel is done in the final step where the panel test statistic is based on some average of the individual cross-section statistics. Each cross-section is allowed its individual cointegrating vector. Each test is constructed such that the cross-sections are assumed independent of each other and heteroskedasticity across the cross-sections is allowed. Im, Pesaran and Shin (1995) present a panel data unit root test based on the average of the ADF statistic of the cross-sections.[2] Using an analogous approach this style of test statistic can be used to test for cointegration. Recall that the ADF test can be constructed as:

$$\hat{e}_{it} = \rho_i \hat{e}_{it-1} + \sum_{j=1}^{p} \vartheta_{ij} \Delta \hat{e}_{it-j} + v_{itp}, \tag{12}$$

where $\hat{e}_{it}$ are OLS residuals from (11). An equivalent way to write equation (12) is given in Phillips and Ouliaris (1990):

$$\Delta \hat{e}_{it} = \rho_i \hat{e}_{it-1} + \sum_{j=1}^{p} \vartheta_{ij} \Delta \hat{e}_{it-j} + v_{itp}.$$

The null hypothesis is written as $H_0 : \rho_i = 0$ and the t-statistic for each i constructed:

$$t_{iADF} = \frac{(\widehat{u}'_{-1} Q_{xp} \widehat{u}_{-1})^{\frac{1}{2}} \hat{\rho}_i}{s_v},$$

where X_p is the matrix of observations on the p regressors $(\Delta \hat{u}_{t-1}, ..., \Delta \hat{u}_{t-p})$, $\hat{u}_{-1}$ is the vector of observations of $\hat{u}_{t-1}$, $Q_{X_p} = I - X_p(X_p' X_p)^{-1} X_p'$ and $s_v^2 = \frac{1}{T} \Sigma_{t=1}^{T} \hat{v}_{tp}^2$.

Phillips and Ouliaris show that the ADF converges to a functional of Brownian motion.

$$t_{iADF} \Rightarrow \frac{\int Q_i dQ_i}{\sqrt{(\int Q_i^2)(\varkappa' \varkappa)}} = \int R dS$$

where

[2] In his paper, Pedroni (1997) also identifies the possibility of extending the logic of Im, Pesaran and Shin from panel unit root tests to panel cointegration tests. Although he does not present the ADF-t statistic in the body of his paper, he does examine some of the properties of the test in his Monte Carlo simulations.

$$R(r) = \int_0^1 \frac{Q(r)}{\sqrt{(\int_0^1 Q^2)}} d\left(\frac{Q(r)}{\sqrt{(\varkappa'\varkappa)}}\right),$$

$$Q(r) = \frac{W_1(r) - (\int W_1 W_2')}{\int W_2 W_2'} W_2(r),$$

$$\int WW' = \begin{bmatrix} f_{11} & f_{21}' \\ f_{21} & F_{22} \end{bmatrix},$$

and

$$\varkappa = (1, 1 - \frac{f_{21}'}{F_{22}}).$$

Finally,

$$\bar{t}_{ADF} = \frac{1}{N}\sum_{i=1}^{N} t_{iADF}.$$

Define $E[\int RdS] = \mu_{Adf}$ and $Var[\int RdS] = \sigma^2_{Adf}$. It can be shown using the logic from Phillips and Moon (1997) that:

$$\sqrt{N}(\bar{t}_{ADF} - \mu_{Adf}) \Rightarrow N(0, \sigma^2_{Adf}).$$

As Phillips and Ouliaris note, the limiting distribution of the ADF test statistic is free of nuisance parameters and depends only in the number of regressors. They provide cut-off tail values for the test in the time series case. It is a mere extension of the logic to then simulate the moments μ_{Adf} and σ^2_{Adf}. Using RDNS procedure in Gauss with 50,000 replications the moments were found to be $\mu_{Adf} = -2.026$ and $\sigma_{Adf} = .8200$ in the case of one regressor. Appropriate values for the mean and standard deviation for 1 to 5 regressors is provided in the Appendix. Tail values are also provided for comparison with the Phillips and Ouliaris results.

3.2 Average Phillips Z_t Statistic for Varying Slopes

As Phillips and Ouliaris (1990) show, corrections for autocorrelation and contemporaneous correlation can either be performed through differencing and the ADF-t statistic method or through non-parametric corrections. In accordance with this idea, another test can be considered which is based on the average, across the cross-sections, of the Phillips Z_t statistics. This statistic is by definition, for the varying intercepts and varying slopes model.

Phillips and Ouliaris (1990) provide exact details on how to calculate the Phillips Z_t test. The first step, as in the ADF test, is to calculate the estimated residuals from the original regression equation using OLS. Then using the estimated residuals, $\hat{e}_{it}$, perform the following regression:

$$\hat{e}_{it} = \alpha_i \hat{e}_{it-1} + v_{it}.$$

Note that this is similar to the ADF test, although here without the lagged terms, the v_{it} may have some effects from then cross-correlation and autocorrelation.

Define:

$$s_{iv}^2 = \frac{1}{T} \sum_{t=1}^{T} \widehat{v}_{it}^2$$

and

$$s_{iTl}^2 = \frac{1}{T} \sum_{i=1it}^{T} \widehat{v}_{it}^2 + \frac{2}{T} \sum_{s=1}^{l} w_{sl} \sum_{t=s+1}^{T} \widehat{v}_{it} \widehat{v}_{it-s}.$$

These terms are used to calculate the final statistic:

$$\hat{Z}_{it} = \frac{(\hat{\alpha} - 1)}{\frac{s_{iTl}}{(\sum_{t=2}^{T} \hat{e}_{it-1}^2)^{\frac{1}{2}}}} - \frac{\frac{1}{2}(s_{iTl}^2 - s_{iv}^2)}{s_{iTl}(\frac{1}{T^2} \sum_{t=2}^{T} \hat{e}_{it-1}^2)^{\frac{1}{2}}}. \tag{13}$$

Phillips and Ouliaris (1990) show that this t-statistic converges in distribution to the same functional of Brownian motion as the ADF t-statistic. Thus, for the purposes here, it is convenient to note that a test based on the Z_t test uses the same simulated moments as the ADF test-statistic given above.

The average of the cross-section $\hat{Z}_{it}$ statistics can be defined as $\bar{Z}_t$

$$\bar{Z}_t = \frac{1}{N} \sum_{i=1}^{N} \hat{Z}_{it}$$

and it can be shown that:

$$\sqrt{N}(\bar{Z}_t - \mu_{Adf}) \Rightarrow N(0, \sigma_{Adf}^2). \tag{14}$$

In his paper, Pedroni (1997) also considers a version of the average of the $\hat{Z}_{it}$ statistic. His test is constructed as follows:

$$\sum_{i=1}^{N} \frac{\sum_{t=1}^{T}(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i)}{(\sum_{t=1}^{T} \hat{L}_{11ie}^{-2}\hat{e}_{it-1}^{2})^{\frac{1}{2}}} \tag{15}$$

where

$$\hat{\lambda}_i = \frac{1}{2}(s_{iTl}^2 - s_{iv}^2)$$

and

$$\hat{L}_{11ie}^{-2} = \sigma_{0ui}^2 - \frac{\sigma_{0u\varepsilon i}^2}{\sigma_{0\varepsilon i}^2}$$

in the scalar case based on the estimates for $\hat{\Omega}_i$ similarly to the one outlined in the previous section.

The two tests are conceptually quite similar[3] with the exception that the form from Phillips and Ouliaris only uses the kernel estimate of $\hat{v}_{it}$ for the non-parametric corrections while Pedroni (1997) also uses $\hat{\Omega}_i$. Intuitively, the form by Phillips and Ouliaris (1990) is clear, a t-statistic on $\hat{\alpha}$ is adjusted with a non-parametric component. To be consistent with the original article, we follow the format from Phillips and Ouliaris (1990). Although it should be noted that Pedroni's simulated moments for the case with an intercept are 2.03 for the mean and $\sqrt{0.66}$ for the standard deviation-values very close to our own. Pedroni (1997) also provides moments for a model without intercept and one including a time trend.

3.3 Pedroni (1997)

As shown in the previous section, Pedroni (1997) also proposes several tests for the null hypothesis of cointegration in panel data. His tests for heterogeneous slopes and intercepts fall into two categories. The first set, as discussed above, involves averaging test statistics for cointegration in the time series across cross-sections. The second set groups the statistics such that instead of averaging across statistics, the averaging is done in pieces so that the limiting distributions are based on limits of piecewise numerator and denominator terms.

The first set of statistics as discussed includes a form of the average of the Phillips and Ouliaris (1990) Z_t statistic. Pedroni also proposes averaging the $Z_{\hat{\alpha}-1}$ statistic. His form for this statistic is

$$\tilde{Z}_{\hat{\alpha}-1} = \sum_{i=1}^{N} \frac{\sum_{t=1}^{T}(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i)}{(\sum_{t=1}^{T} \hat{e}_{it-1}^{2})} \tag{16}$$

[3] In fact Z_t can be written as $\frac{\sum \hat{e}_{it-1}v_{it} - T\lambda}{(\sum \hat{e}_{it-1}^2)^{\frac{1}{2}} s_{itL}}$. Under the null hypothesis, v_{it} is equivalent to $\Delta\hat{e}_{it}$.

This form directly corresponds to the statistic proposed by Phillips and Ouliaris. [4]

Pedroni shows that

$$\sqrt{N}(T\tilde{Z}_{\hat{\alpha}-1} + 9.05) \Rightarrow N(0, 35.98)$$

For his second set of statistics[5], Pedroni defines the following as panel variance ratio statistics:

$$Z_{\hat{v}_{NT}} = \frac{1}{(\Sigma_{i=1}^{N}\Sigma_{t=2}^{T}\hat{L}_{11i}^{-2}\hat{e}_{it-1}^{2})} \;,$$

$$Z_{\hat{\rho}_{NT}} = \frac{\Sigma_{i=1}^{N}\Sigma_{t=2}^{T}\hat{L}_{11i}^{-2}(\hat{e}_{it-1}\hat{e}_{it}-\hat{\lambda}_{i})}{(\Sigma_{i=1}^{N}\Sigma_{t=2}^{T}\hat{L}_{11i}^{-2}\hat{e}_{it-1}^{2})} \;,$$

and

$$Z_{t_{\hat{\rho}_{NT}}} = \frac{\Sigma_{i=1}^{N}\Sigma_{t=2}^{T}\hat{L}_{11i}^{-2}(\hat{e}_{it-1\Delta}\hat{e}_{it}-\hat{\lambda}_{i})}{\sqrt{\tilde{\sigma}_{NT}^{2}(\Sigma_{i=1}^{N}\Sigma_{t=2}^{T}\hat{L}_{11i}^{-2}\hat{e}_{it-1}^{2})}} \;,$$

where

$$\tilde{\sigma}_{NT} = \frac{1}{N}\sum_{i=1}^{N}(\frac{\hat{\sigma}_i}{\hat{L}_{11i}})^2,$$

$$\hat{\lambda}_i = \frac{1}{2}(\hat{\sigma}_i^2 - \hat{s}_i^2),$$

and

$$\hat{s}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{it}^2.$$

Using consistent estimates of Ω_i, the long-run variance-covariance matrix given in the previous section, define $\hat{L}_i$ to be the lower triangular Cholesky composition of $\hat{\Omega}_i$ such that in the scalar case $\hat{L}_{22i} = \hat{\sigma}_{\varepsilon i}$ and $\hat{L}_{11i} = \hat{\sigma}_{ui}^2 - \frac{\hat{\sigma}_{u\varepsilon i}^2}{\hat{\sigma}_{\varepsilon i}^2}$, the long-run conditional variance.

Let θ and ψ signify the mean and variance for $\tau = (\int Q^2, \int QdQ, \tilde{\beta}^2)$, a vector of functionals of Brownian motion.

Define

$$\tilde{\beta} = \frac{\int VW}{\int W^2},$$

$$Q = V - \tilde{\beta}W,$$

and $\psi_{(i)}$, $i = 1, 2, 3$ refers to i x i upper sub-matrix of ψ.

[4] The format for this statistic from Phillips and Ouliaris is given as $\hat{Z}_\alpha = T(\hat{\alpha}-1) - (\frac{1}{2})(s_{Tl}^2 - s_k^2)(\frac{\Sigma_2^T \hat{u}_{t-1}^2}{T^2})^{-1}$, the panel version can be written as $\frac{\sqrt{N}(\bar{Z}_\alpha + 9.05)}{\sqrt{35.98}}$

[5] Again with these statistics, Pedroni (1997) uses corrections which do not follow directly from Phillips and Ouliaris (1990). In the simulations, therefore we adapt Pedroni's $Z_{t_{\hat{\rho}_{NT}}}$ by eliminating the $\hat{L}_{11i}^{-2}$ terms. With these corrections the performance of the test improves dramtically.

Pedroni suggested the following results:

$$TN^{\frac{3}{2}}Z_{\hat{v}_{NT}} \Rightarrow \frac{1}{\frac{1}{N}\Sigma_{i=1}^{N}Q_i^2},$$

$$T\sqrt{N}(Z_{\hat{\rho}_{NT}}-1) \Rightarrow \frac{\frac{1}{\sqrt{N}}\Sigma_{i=1}^{N}\int Q_i dQ_i}{\frac{1}{N}\Sigma_{i=1}^{N}\int Q_i^2},$$

and

$$Z_{t_{\hat{\rho}_{NT}}} \Rightarrow \frac{\frac{1}{\sqrt{N}}\Sigma_{i=1}^{N}\int Q_i dQ_i}{\sqrt{(\frac{1}{N}\Sigma_{i=1}^{N}\int Q_i^2)(1+\frac{1}{N}\Sigma_{i=1}^{N}\tilde{\beta}_i^2)}}, \tag{17}$$

as $T \to \infty$. Hence Pedroni shows that:

$$TN^{\frac{3}{2}}Z_{\hat{v}_{NT}} - \frac{\sqrt{N}}{\theta_1} \Rightarrow N(0, \phi'_{(1)}\psi_{(1)}\phi_{(1)}),$$

$$T\sqrt{N}(Z_{\hat{\rho}_{NT}}-1) - \frac{\sqrt{N}\theta_2}{\theta_1} \Rightarrow N(0, \phi'_{(2)}\psi_{(2)}\phi_{(2)}),$$

and

$$Z_{t_{\hat{\rho}_{NT}}} - \frac{\theta_2\sqrt{N}}{\sqrt{\theta_1(1+\theta_3)}} \Rightarrow N(0, \phi'_{(3)}\psi_{(3)}\phi_{(3)}),$$

where

$$\phi_{(1)} = -\frac{1}{\theta_1^2},$$

$$\phi'_{(2)} = (\frac{1}{\theta_1}, \frac{\theta_2}{\theta_1^2}),$$

and

$$\phi'_{(3)} = (\frac{1}{\sqrt{\theta_1(1+\theta_3)}}, -\frac{1}{2}\frac{\theta_2}{\theta_1^{\frac{3}{2}}\sqrt{1+\theta_3}}, -\frac{1}{2}\frac{\theta_2}{\sqrt{\theta_1}(1+\theta_3)^{\frac{3}{2}}}).$$

Pedroni thus bases his test on the average on the numerator and denominator respectively, rather than the average for the statistic as a whole and this allows him to decompose the limiting ratio into the separate θs.

Because the model is based on varying slopes and intercepts, each cross-section is estimated individually. It is clear that the asymptotic distributions, then, are based on the means of functionals of Brownian motion accounting for the independence across the cross-sectional observations. The limiting distributions are free of nuisance parameters as the moments of the functionals of Brownian motion are independent of the data and can be found by Monte Carlo simulation. Thus, with the Monte Carlo results the asymptotic distributions can be written as:

$$TN^{\frac{3}{2}}Z_{\hat{v}_{NT}} - 8.62\sqrt{N} \Rightarrow N(0, 60.75),$$

$$T\sqrt{N}(Z_{\hat{\rho}_{NT}} - 1) + 6.02 \Rightarrow N(0, 31.27),$$

and

$$Z_{t_{\hat{\rho}_{NT}}} + 1.73\sqrt{N} \Rightarrow N(0, 0.93).$$

Note that these distributions apply to the model including an intercept and not including a time trend. Asymptotic results for other model specifications can be found in Pedroni (1997). The intuition on these tests with varying slopes is not straightforward. The convergence in distribution is based on individual convergence of the numerator and denominator terms. What is the intuition of rejection of the null hypothesis? Using the average of the overall test statistic allows more ease in interpretation: rejection of the null hypothesis means that enough of the individual cross-sections have statistics "far away" from the means predicted by theory were they to be generated under the null.

3.4 Comments

The articles by Kao and Pedroni present important methods for testing cointegration in panel data under the null of no cointegration. They mirror the development in the time series literature in that they present a parametric approach, such as the ADF approach, and a non-parametric approach, such as the Phillips and Ouliaris test statistics, which correct the data non-parametrically as the test statistics are calculated thus arriving at distributions free of nuisance parameters. The tests as they appear in the original articles can be improved upon. Most of the tests proposed necessarily depend on consistent estimates of the long-run variance covariance matrix of the residuals of the random walk processes. The question becomes when is the best time to correct for the possible presence of autocorrelation and weak exogeneity of the error terms.

The estimation of Ω_i is a complicated affair and the estimation procedure relies on non-parametric kernel methods.

Ω_i can be estimated by

$$\widehat{\Omega}_i = \left\{ \frac{1}{T} \sum_{t=1}^{T} \hat{\xi}_i \widehat{\xi}_i^{'} + \frac{1}{T} \sum_{\tau=1}^{l} \varpi_{\tau l} \sum_{t=\tau+1}^{T} \left(\widehat{\xi}_i \widehat{\xi}_{it-\tau}^{'} + \widehat{\xi}_{it-\tau} \widehat{\xi}_{it}^{'} \right) \right\}, \tag{18}$$

where $\varpi_{\tau l}$ is a weight function or a kernel. Usual kernels are truncated by the bandwidth parameter l so that $\varpi_{\tau l} = 0$ for $\tau > l$.

4 Testing for Cointegration in Panels with the Null Hypothesis of Cointegration

4.1 McCoskey and Kao (1998)

In this section, a panel test of the null hypothesis of cointegration is presented. Tests of this null hypothesis were first introduced in the times series literature as a response to some critiques of the null hypothesis of no cointegration. For example, testing the null of cointegration rather than the null of no cointegration could be very appealing in applications where cointegration is predicted a priori by economic theory. Also, failure to reject the null of no cointegration could be caused, in many cases, by the low power of the test and not by the true underlying nature of the data.

The residual-based test for null of cointegration in panel data proposed by McCoskey and Kao (1998) is an extension of the Lagrange multiplier (LM) test and locally best invariant (LBI) test for an MA unit root in the time series literature. This test is also discussed in McCoskey and Kao (1998). Cointegration tests of the null of cointegration in the time series case have been proposed by Harris and Inder (1994) and Shin (1994). Under the null, the asymptotics no longer depend on the asymptotic properties of estimating spurious regression, rather the asymptotics of the estimation of a cointegrated relationship are needed. For models which allow the cointegrating vector to change across the cross-sectional observations, the asymptotics depend merely on the time series results as each cross-section is estimated independently. For models with common slopes, the estimation is done jointly and therefore the asymptotic theory is based on the joint estimation of a cointegrated relationship in panel data.

For the residual based test of the null of cointegration, it is necessary to use an efficient estimation technique of cointegrated variables. In the time series literature a variety of methods have been shown to be efficient asymptotically. These include the fully modified (FM) estimator of Phillips and Hansen (1990) and the dynamic least squares (DOLS) estimator as proposed by Saikkonen (1991) and Stock and Watson (1993). For panel data, Kao and Chiang (1997) show that both the FM and DOLS methods can produce estimators which are asymptotically normally distributed with zero means.

The model presented allows for varying slopes and intercepts:

$$y_{it} = \alpha_i + x_{it}^{'}\beta_i + e_{it}, \; i = 1, ..., N, \; t = 1, ..., T, \tag{19}$$

$$x_{it} = x_{it-1} + \varepsilon_{it} \tag{20}$$

$$e_{it} = \gamma_{it} + u_{it}, \tag{21}$$

and

$$\gamma_{it} = \gamma_{it-1} + \theta u_{it}. \tag{22}$$

The null of hypothesis of cointegration is equivalent to $\theta = 0$.

The test statistic proposed by McCoskey and Kao is the following:

$$\overline{LM} = \frac{\frac{1}{N}\Sigma_{i=1}^{N}\frac{1}{T^2}\Sigma_{t=1}^{T}S_{it}^{+2}}{s^{+^2}}, \tag{23}$$

where S_{it} is partial sum process of the residuals,

$$S_{it}^{+} = \sum_{j=1}^{t}\widehat{e}_{ij}$$

with

$$s^{+2} = \frac{1}{NT}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}\hat{e}_{it}^{+2}.$$

($\hat{\varpi}_{1.2}^{2}$ is defined as a consistent estimator of σ_v^2, the long-run conditional variance under the H_0 and is used in place of s^{+^2} if the residuals are estimated using the FM estimator.) The FM estimator non-parametrically corrects for the possible serial correlation and weakly exogenous regressors in a cointegrated regression. The DOLS estimator uses lagged and future differences of x_{it} to correct for these effects.

The asymptotic result for the test is:

$$\sqrt{N}(\overline{LM} - \mu_v) \Rightarrow N(0, \sigma_v^2), \tag{24}$$

where $\mu_v = .1162$ and $\sigma_v^2 = .0109$ and are defined in McCoskey and Kao (1998). The constants μ_v and σ_v^2 are moments of a complex functional of Brownian motion, which depend only on the number of regressors and can be found through Monte Carlo simulation.

The limiting distribution of $\overline{LM}$ is then free of nuisance parameters and robust to heteroskedasticity. For the Monte Carlo results, the fully modified estimation method is used.

4.2 Comments

The asymptotics of the panel tests take advantage of the sequential limit theory which allows for indices across the two dimensions of T and N. For the panel LM test an additional dimension is added to create the partial sums of the residuals. The fact that the model here allows for varying intercepts means that each cross-section is actually estimated individually, thus the additional dimension is manageable in the asymptotics.

5 The Monte Carlo Design

The ultimate goal of this Monte Carlo study is to compare the size and power of different residual based tests for cointegration for two models: varying slopes and varying intercepts and common slopes and varying intercepts. For the common slopes model, the two tests to be considered are both derived under the null of no cointegration, so the comparison is quite straightforward. The ADF_t from Kao and the corrected ADF_t are compared. However, for the varying slopes model a total of five tests are considered: ADF^*, PO_t^*, PO_α^*, APG^* and LM^*. The first four statistics are constructed under the null of no cointegration and the last under the null of cointegration. We first compare the four tests of the null of no cointegration for size and power and then select two of these tests to compare with LM^*. To compare these three tests, the study follows Harris and Inder (1994) who suggest testing the ability of the tests to properly identify the underlying nature of the data using two different Data Generating Processes. Thus, each null hypothesis is represented in the final experiment. This entails a two-step procedure outlined below. The simulations were performed in GAUSS using the package COINT 2.0.

5.1 Experimental Design

5.1.1 Varying Slopes and Varying Intercepts

$$y_{it} = \alpha_i + x_{it}^{'}\beta_i + e_{it},\ i = 1, ..., N,\ t = 1, ..., T. \tag{25}$$

As discussed in Section 6.1, after the preliminary comparisons of size and power of all four tests of the null hypothesis of cointegration, three tests, ADF^*, APG^*, and LM^*, are selected for comparison across both the null of cointegration and null of no cointegration. To compare the tests with varying intercepts, the study considers the two different data generating processes (DGP):

DGP-A, null of no cointegration:

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it},$$

and

$$e_{it} = \rho e_{it-1} + v_{it}.$$

Under the null hypothesis of no cointegration, $\rho = 1$. The study includes the following possible values for $\rho :$

$$\rho \in \{1, 0.95, 0.85, 0.75\}.$$

DGP-B, null of cointegration:

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it},$$

and

$$e_{it} = \theta \sum_{k=1}^{t} v_{ik} + v_{it}.$$

Under the null hypothesis of cointegration $\theta = 0$, i.e., under the null the error term does not remember past errors and collapses to a standard normal random variable. The study includes the following values for $\theta :$

$$\theta \in \{0, 0.05, 0.15, 0.25\}$$

Notes for both DGP-A and DGP-B we assume that v_{it} is distributed $N(0,1)$ and

$$x_{it} = x_{it-1} + \varepsilon_{it},$$

where ε_{it} is distributed $N(0, \sigma_i^2)$.

In the study, other parameters are also considered to test the flexibility across DGPs. In particular and in accordance with Phillips and Loretan (1991), the study looks at parameters allowing for a moving average component in the error term and weak exogeneity. Only the special cases where $\rho = 0.75$ or $\theta = .25$ and $N = T = 50$ are considered.

Define π, the moving average component (autocorrelation) in v_{it}, $\pi \in \{-0.8, 0, 0.8\}$

$$v_{it} = v_{it}^* + \pi v_{it-1}^*$$

and δ, cross-correlation in v_{it}^* and ε_{it}, $\delta \in$\{-0.5, 0, 0.5\}, with

$$\begin{pmatrix} v_{it}^* \\ \varepsilon_{it} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \delta\sigma_i \\ \delta\sigma_i & \sigma_i^2 \end{bmatrix} \right).$$

Unless otherwise specified, the study considers the following dimensions for N and T: $N \in \{1, 15, 25, 50, 100\}$ and $T \in \{15, 25, 50, 100\}$ and the number of replications for each dimension is $10,000$. The choice of N and T for this experiment underlines an important point: this test is not really appropriate with severely unbalanced data sets, for example extremely large N and small T.

α_i, β_i and σ_i are generated using the default uniform random number generator in GAUSS, i.e., $\alpha_i \sim U[0, 10]$, $\beta_i \sim U[0, 2]$, and $\sigma_i \sim U[0.5, 1.5]$.

5.1.2 Two Stages of the Experiment

Because the tests are not all derived under the same null hypothesis, it is difficult to compare their performance directly. A two stage procedure is used here to make sure the results are comparable. The first stage is to compute 5% and 95% critical values under the null of each DGP, A and B. These critical values are used to set the probability of rejecting the null of the particular DGP to 0.05 for all three tests. For the tests of the null hypothesis of no cointegration, with DGP-A, this 5% is simply the value which leads to size equal to 0.05. For the third test of the null of cointegration, this 5% critical value with DGP-A does not, strictly speaking, relate to the size of the test but rather is simply a probability of rejection. With DGP-B the logic is just the reverse. The 5% critical value of the test of the null hypothesis is the value which leads to a size of 0.05 whereas for the other two tests it does not have this exact interpretation. The size is directly related to the null hypothesis of the test not the DGP of the experiment. To summarize:

DGP-A (generated under the null of no cointegration), choose critical values such that

$$\begin{array}{cccc} \text{Test 1} & \text{Test 2} & \text{Test 3} & \\ \Pr(\text{reject null}) = & \Pr(\text{reject null}) = & \Pr(\text{reject null}) = & 0.05 \;\cdot \\ \text{Size} & \text{Size} & & \end{array}$$

DGP-B (generated under the null of cointegration), choose critical values such that

$$\begin{array}{cccc} \text{Test 1} & \text{Test 2} & \text{Test 3} & \\ \Pr(\text{reject null}) = & \Pr(\text{reject null}) = & \Pr(\text{reject null}) = & 0.05 \;\cdot \\ & & \text{Size} & \end{array}$$

The critical values are given Tables 1 and 2. The empirical rejection rates, using a one-sided $N(0,1)$ distribution are given in Table 3. These critical values are constructed to insure the tails of all three tests have equivalent density.

The second stage is to calculate rejection rates for the alternative values of the parameters for DGPs A and B. In other words, the second step is used to calculate the ability of the tests to properly reject the null hypothesis of the DGP based on critical values found in the first stage. The intuition behind these two steps is analogous to finding the power of a test after adjusting the critical values for the size of the test. Again, the strict concepts of size and power must be used with caution in this experiment as the null hypothesis used in the DGP is not necessarily the null hypothesis used to construct the test.

Let $c_j^k(N,T)$ be the critical value calculated in Stage 1 for test $j = 1,2,3$ and DGP-k, $k = A$ or B for a given N and T.

DGP-A given a specific $\rho < 1$ (data generated under the alternative of cointegration):

$$\begin{array}{ccc} \text{Test 1} & \text{Test 2} & \text{Test 3} \\ \Pr\left(\text{test1} < c_1^A(N,T)\right) & \Pr\left(\text{test2} < c_2^A(N,T)\right) & \Pr\left(\text{test3} < c_3^A(N,T)\right) \\ \text{Power (Size=0.05)} & \text{Power (Size=0.05)} & \end{array}.$$

In all cases this is the probability of correctly rejecting the null hypothesis of no cointegration.

DGP-B given a specific $\theta > 0$ (data generated under the alternative of no cointegration):

$$\begin{array}{ccc} \text{Test 1} & \text{Test 2} & \text{Test 3} \\ \Pr\left(\text{test1} > c_1^B(N,T)\right) & \Pr\left(\text{test2} > c_2^B(N,T)\right) & \Pr\left(\text{test3} > c_3^B(N,T)\right) \\ & & \text{Power (Size=0.05)} \end{array}.$$

These probabilities represent the probability of the test correctly rejecting the null hypothesis of cointegration. These comparisons of the tests' ability to properly reject the null hypothesis of DGPs A and B are given in Tables 4 and 5.

5.1.3 Common Slopes and Varying Intercepts

In this case, both tests are derived under the null of no cointegration so the following specification is used:

$$y_{it} = \alpha_i + \beta x_{it} + e_{it}$$

$$e_{it} = \rho e_{it-1} + v_{it}$$

Under the null hypothesis of no cointegration, $\rho = 1$. The study includes the following possible values for ρ :

$$\rho \in \{1, 0.95, 0.85, 0.75\}.$$

α_i and σ_i are generated using the default uniform random number generator in GAUSS:

$$\alpha_i \sim U[0, 10]$$

and

$$\sigma_i \sim U[0.5, 1.5]$$

Unlike in the previous section, β is assumed constant across the cross-sections and is set equal to 2. The other assumptions are the same as in the previous model.

5.2 Test Statistics

Results from the following forms of the tests are reported. Define the following standardized statistics for varying slopes and varying intercepts:

$$ADF^* = \frac{\sqrt{N}(\bar{t}_{ADF} + 2.026)}{.82},$$

$$PO_t^* = \frac{\sqrt{N}(\bar{Z}_t + 2.026)}{.82},$$

$$PO_\alpha^* = \frac{\sqrt{N}(\bar{Z}_\alpha + 9.05)}{\sqrt{35.98}},$$

$$APG^* = \left(\frac{\Sigma_{i=1}^N \Sigma_{t=2}^T (\hat{e}_{it-1\Delta}\hat{e}_{it} - \hat{\lambda}_i)}{\sqrt{\bar{\sigma}_{NT}^2 (\Sigma_{i=1}^N \Sigma_{t=2}^T \hat{e}_{it-1}^2)}} + 1.73\sqrt{N}\right)/\sqrt{.93},$$

and

$$LM^* = \frac{\sqrt{N}(\overline{LM} - .1162)}{\sqrt{.0109}},$$

where

$$\bar{\sigma}_{NT} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}_i)^2.$$

Define Kao's standardized ADF, ADF_K, and a biased corrected ADF, ADF_{SM}, for common slopes and varying intercepts as:

$$ADF_K = \frac{t_{ADF} + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{ov}}}{\sqrt{\frac{\hat{\sigma}_{ov}^2}{2\hat{\sigma}_v^2} + \frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{ov}^2}}}$$

and

$$ADF_{SM} = \sqrt{\frac{5}{4}}\left(\frac{\hat{\sigma}_v}{\hat{\sigma}_{ov}}\left(t_{ADF} - \frac{\sqrt{N}\lambda}{s_v\sqrt{\xi_{8T}}}\right) + \frac{\sqrt{1.5N}\hat{\sigma}_v}{s_v}\right).$$

5.3 Interpreting the results

Ultimately the goal of simulations with varying slopes is to see how well these three tests can distinguish between the true character of the DGP and its alternative. For each of these tests under the two different DGPs, the following probability is desired:

$$\Pr_{DGP-H_A}(\text{Rejecting the } DGP|H_0),$$

i.e., the probability of rejecting the null of the DGP when the alternative is true.

Call this probability rej_j^k, the rejection rate of test j under DGP k :

$$rej_j^A = rej_j^A(N, T, \rho)$$

and

$$rej_j^B = rej_j^B(N, T, \theta).$$

Each of these individual experiments can be considered as the sum of Bernoulli random variables where

$$X_i = \begin{cases} 1 & \text{if reject} \\ 0 & \text{otherwise} \end{cases}$$

and the pdf is given:

$$p_X(X) = \begin{cases} p^x(1-p)^{1-x} \ x = 0, 1, ... \\ \\ 0 \text{ otherwise} \end{cases}$$

with $E(X) = p$ and $Var(X) = p(1-p)$.

In this experiment, each p_j^A and p_j^B is estimated by finding the mean of these Bernoulli random variables:

$$rej_j^A = \frac{\Sigma_{i=1}^{10,000} X_{ji}^A}{10,000}$$

Given that each experiment is iid we obtain:

$$E(rej_j^A) = \quad \tfrac{1}{10,000} * 10,000 * E(X_{ji}^A) = \quad p_j^A$$

and

$$Var(rej_j^A) = \quad \tfrac{1}{(10,000)^2} * 10,000 * Var(X_{ji}^A) = \quad \tfrac{p_j^A(1-p_j^A)}{10,000} \quad .$$

Thus the standard error for each rejection rate is equal to

$$\frac{\sqrt{p_j^A(1-p_j^A)}}{\sqrt{10,000}}.$$

The standard error reaches a maximum at the rejection rate of 0.5 with a standard deviation of 0.005. For a rejection rate of 0.99, the standard deviation would be 0.000995.

How about comparing rej_1^A and rej_2^A? Comparing the rejection rates across two tests for the same DGP (i.e., to answer the question: which is better at correctly rejecting?) is equivalent to evaluating the significance of the difference of two random variables:

$$Var(rej_1^A - rej_2^A) = Var(rej_1^A) + Var(rej_2^A) - 2Cov(rej_1^A, rej_2^A)$$

where the covariance is given by

$$\sum_{i=1}^{10{,}000} \frac{(X_{1i}^A - rej_1^A)(Y_{2i}^A - rej_2^A)}{10,000}.$$

The intuition here is that the covariance is measuring whether the tests will reject for the same data or not. The standard error of comparison is given by:

$$\sqrt{Var(rej_1^A) + Var(rej_2^A) - 2Cov(rej_1^A, rej_2^A)}.$$

This reaches a maximum when the tests are assumed independent and each have a rejection rate of 0.5. In that case the standard deviation for comparison would be .007071. In the results, a test is considered "significantly better" if the difference between the two rejection rates is at least as large as two times the standard deviation of comparison.

The above discussion also applies to the simulations with common slopes although the performance with only one DGP is compared.

6 Results

6.1 Varying Slopes and Varying Intercepts

In Table 1 we show a preliminary comparison of the four tests of the null hypothesis of no cointegration in terms of empirical size. Of the four tests, ADF^* seemingly performs the best. Theory would predict that size should converge to .05 for all tests. In fact ADF^* has a small range across all N and T. The maximum rejections rate reported is .0782 and the minimum .0229. Both of these occur when $T = 15$. When $T = 100$, the range narrows considerably with a size of .0470 when $N = 1$ to .0518 when $N = 15$. PO_t^* has a strong tendency to over-reject when $T \leq 25$ and tendency to under-reject for T$\geq$ 50. This is especially true for large N. For example, when $T = 15$ and $N = 100$, PO_t^* will reject the null hypothesis almost 28% of the time. When $T = 100$ and $N = 100$, PO_t^* will only reject the null .19% of the time. PO_α^*, to the contrary, underrejects the null in all cases expect when $T = 100$ and $N = 1$. This is especially severe for small T. For example, when $T = 15$ and $N \geq 15, PO_\alpha^*$ never rejects the null hypothesis. APG^*, like PO_t^*, has a tendency to overreject for small T and underreject for large T, although in neither cases is the problem as severe. The range of size values for PO_t^* is $[.0019, .2789]$ while for APG^* the range is $[.0035, .1371]$. For APG^*, under-rejection is especially a problem when $T \geq 25$ and $N \geq 25$. These results underline the intuition of panel data that relative sizes of the T and N dimension can significantly impact the characteristics of the test.

In Table 2 we compare results for the power of the test to correctly reject the null hypothesis of no cointegration. In this preliminary comparison, only results for $\rho = 0.95$ or $\rho = 0.75$ are reported. A '*' is used to indicate when a test performs significantly better than any of the other three when compared pairwise. At first glance it is clear the APG^* performs well with regard to power. In all cases when a most powerful test can be determined, it is APG^*. All tests show that decreasing the value of ρ increases the power of the test. It is also clear that increasing the T dimension increases the power of the test more than increasing the cross-section dimension. This result is especially important in applications where researchers may have more ability to increase the cross-sections in their data rather than find more time series data. For example, holding T constant at 25, increasing N from 15 to 100 has the following impact when $\rho = 0.95$, the power of ADF^* increases from .1040 to .3153; the power of PO_t^* increases from .1572 to .5621; the power of PO_α^* increases from .1445 to .5036; and the power of APG^* increases from .1709 to .7289. However, when N is held constant at 25 and T increases from 15 to 100 ($\rho = 0.95$), the power of ADF^* increases from .0934 to .8019; the power of PO_t^* increases from .1291 to .9636; the power of PO_α^* increases from .1218 to .9314; and the power of APG^* from .1510 to .9826. Another strong result is how using a panel rather than strict time series can increase power dramatically. In the strict time series case ($N = 1$), when $\rho = 0.95$ (i.e. data is almost non-stationary) the power of the four tests ranges from .0918 to .1251. However, increasing N from 1 to 15 shifts the range to a minimum at .5695 and maximum at .8610. Thus, the difficulty in correctly rejecting the null for "near non-stationary data" is alleviated.

Given the results from Tables 1 and 2, we choose ADF^* and APG^* as the tests to compare with LM^* for the second phase of the Monte Carlo experiments. ADF^* performed the best with regard to empirical size and APG^* with regard to power.

As the results on empirical size for ADF^* and APG^* have already been discussed, we now turn our attention to size results for LM^*. From Table 5 we observe that, in general, holding T fixed and increasing N decreases the size of the test. In the case of $T < 50$, this causes the test to under-reject the null hypothesis for large N. For $T > 50$, in all cases the test underrejects the null with a range in size of $[.0533, .0917]$, the minimum being reached when $T = 50$ and $N = 100$ and the maximum when $T = 100$ and $N = 15$. Again we see the importance of the relative size of the dimensions of the panel.

Turning our attention to the ability of the three tests to correctly reject the null hypothesis of DGP-A shown in Table 6, we see some surprising results. Intuition seems to suggest that ADF^* and APG^* should outperform LM^* as they are tests derived under this null. However, LM^* clearly outperforms the other two tests for the cases where $\rho = 0.95$ and $\rho = 0.85$. APG^* outperforms the other two when $\rho = 0.75$ for all the cases when $N = 1$, the strict time series case, and when $T \leq 25$ and $N = 15$.

The dominance of LM^* in the cases of ρ close to 1 is significant as these errors which show cointegration are very close to being non-stationary, and it is these "nearly non-stationary" errors which can give researchers the most difficulty. In some cases the power of LM^* is very far away from the other two tests. For example, when $\rho = 0.95$, $T = 50$ and $N = 25$, LM^* correctly rejects the null hypothesis 91.57% of the time, APG^* properly rejects the null only 57.20% of the time and ADF^* only 27.92% of the time. All of the tests show the nicely behave properties that power increases as ρ decreases, and N and T increase.

Considering Table 7 and the tests' ability to correctly reject the null hypothesis of DGP-B, we see that again LM^* dominates. In fact in all cases when a most powerful test could be determined, LM^* was the most powerful test. However, this results is not quite so surprising as LM^* is the only test of the three derived under the null hypothesis of cointegration. For both Tables 6 and 7 we see the benefits of using panel data rather than the simple time series. In Table 7, when $T = 15$ and $\theta = 0.15$, increasing N from 1 to 15 increases the power from .1066 to .2872; when $T = 25$ an identical increase changes power from .1705 to .6375; when $T = 50$ power increases from .3943 to .9937; and for $T = 100$ power increases from .7224 to .9999

6.2 Common Slopes and Varying Intercepts

Results for the 5% percent critical tails and empirical size for the two tests with common slopes is given in Table 9. Again, theory for these two one-sided tests would predict values close to -1.645 for a rejection rate of 5%. For $T \leq 25$, in almost all cases (with the one exception of $T = 25$ and $N = 15$) the size of the ADF_{SM} test is closer to 0.05 than the size of the ADF_K test. For $T > 25$, the results are not so clear. When $N = 15$ or 25, the size of the ADF_K test is closer to 0.05 than the size for the ADF_{SM}. It is interesting to note that the ADF_K test seems more sensitive to questions of balance in the dimensions of T and N. For both $T = 50$ and $T = 100$, increasing the N dimension from 50 to 100 results in an increase in empirical size away from the 5% level. For the ADF_{SM} test, in all cases increasing the N dimension results in smaller empirical sizes-moving closer to 0.05.

The size-adjusted powers for both tests are given in Table 10. Theory once again predicts that power should increase as T and N increase and increase as ρ decreases. In this case the results are encouraging in which the adjustments made by the ADF_{SM} increase the power dramatically, particularly for small T. In fact, when possible to determine the more powerful test, the ADF_K test outperformed the ADF_{SM} test only when the cross-section dimension is limited to 1. Both tests have powers which increase toward 1 as T and N increase and for $T > 15$, decreases in ρ away from the null value of 1, causes the power to increase. The difference in the power between the two tests is most dramatic for small T. For example when $T = 25$ and $N = 25$ the power of the ADF_K test is .1774, .3009 and .3804 for $\rho = 0.95, 0.85$ and 0.75 respectively while the power for the ADF_{SM} test is .5180, .9896. and .9999.

The performance of the two tests with different values for π and δ is interesting. Once again, theory shows that, asymptotically, these affects should have no impact on the distribution. Results are given in Table 11. As in the Monte Carlo experiment for varying intercepts, the results are quite different for the two tests. The effect of the moving average, π, seems again to dominate. For the ADF test, a positive moving average has a greater impact on the size of the test. Consider the case for $\delta = 0$, for $\pi = -0.8$ the empirical size decreases from .0954 to .0084 in contrast to when $\pi = 0.8$ which results in a jump to an empirical size of .4073. The sign on δ, the weak exogeneity parameter has little effect. For $\delta = -0.5$ and $\pi = 0.8$, the empirical size is .6417 while the size for $\delta = 0.5$ and $\pi = 0.8$ is .6399. For the ADF_{SM} test, a negative moving average increases the empirical size to .9999 for all values of δ. A positive moving average value causes the empirical size to decrease dramatically to values of .0001 or .0000. It is interesting to note that the ADF_K test seems to work better in the presence of a negative moving average while the ADF_{SM} test seems to work better in the presence of a positive moving average. The size adjusted power of both tests is equal to .9999 for all cases.

7 Conclusion

The development of non-stationary econometrics in the time series literature allowed for a deeper understanding of the statistics of "long-run steady state" relationships. These relationships were identified as cointegrated relationships among non-stationary variables. Extending these results to panel data offers the new challenge of how to combine results on cross-sectional data combined with the time series. This chapter evaluates tests for cointegration in panel data. Which test among these is best?

The first step in selecting a test is to understand clearly the nature of the long run relationship to be tested. In many applications, theory determines how homogeneous the long run relationships should be across the cross sections. The most important factor is whether or not there exists a common slope coefficient across the cross sections. There is also the empirical consideration of how appropriate it is to pool the data. Ideally, a test would exist which could test this property of the data.

If the theory suggests that the cross sections need not have a common slope, then this chapter has presented three final tests from which to choose. Two of the tests are constructed under the null hypothesis of no cointegration, ADF^* and APG^*. These tests are based on the ADF test and Pedroni's pooled tests. The third test is based on the null hypothesis of cointegration which is based on the LM test from the time series literature, LM^*. Of these three tests for varying slopes, which is best?

The test of the null hypothesis was originally proposed in response to the low power of the tests of the null of no cointegration, especially in the time series case. Further, in those cases where economic theory predicted a long run steady state relationship, it seemed that a test of the null of cointegration rather than the null of no cointegration would be appropriate. The results from the Monte Carlo study here shows that LM^* does outperform the other two tests. In both experiments, LM^* was seen to be more powerful, especially for cases when the parameters generated were very close to values under the null.

Of the two reasons for the introduction of the test of the null hypothesis of cointegration, low power and attractiveness of the null, the introduction of the cross-section dimension of the panel solves one: all of the tests show decent power when used with panel data. For those applications where the null of cointegration is more logical than the null of no cointegration, this study, at a minimum, concludes that using LM^* does not compromise the ability of the researcher of determining the underlying nature of the data.

If the theory suggests that the cross sections should be restricted to a common slope, this chapter presents two tests from which to choose, an ADF-type test and a test based on non-parametric adjustment of the ADF-type test. Both of these tests are constructed under the null hypothesis of no cointegration. In this case, unless specific information is known about the presence of a negative moving average in the errors, then the test based on the non-parametric adjustment should be used, the ADF_{SM} test.

8 Appendix

The following are critical values, including mean and standard deviation, for ADF and Z_t which are used for $\bar{t}_{ADF}$ and $\bar{Z}_t$.

k	*mean*	*std*	10%	5%	1%
1	−2.0261	.8200	−3.0383	−3.3329	−3.9197
2	−2.4687	.8000	−3.4695	−3.7576	−4.3290
3	−2.8535	.7800	−3.8319	−4.1212	−4.6750
4	−3.1758	.7668	−4.1500	−4.4344	−4.9978
5	−3.4816	.7583	−4.4584	−4.7451	−5.2998

The values were calculated in GAUSS using 50,000 replications. Since asymptotic theory tells us that the ADF test should asymptotically be identical in distribution to a Dickey-Fuller test with iid errors, a Dickey-Fuller test on non-stationary residuals was simulated.

Values for the individual 10%, 5%, and 1% levels are provided as a comparison to the values given in Phillips and Ouliaris (1990).

References

[1] Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

[2] Baltagi, B. H. (1995), *Econometric Analysis of Panel Data*, John Wiley & Sons, Inc., New York.

[3] Banerjee, A., Dolado, J., Galbraith, J. W. , and Hendry, D. F.. (1994), *Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press, New York.

[4] Granger, C. W. J. (1997), "On Modelling the Long Run in Applied Economics," *The Economic Journal*, 107, 169-177.

[5] Harris, D., and Inder, B. (1994), "A Test of the Null Hypothesis of Cointegration," in *Nonstationary Time Series Analysis and Cointegration*, Hargreaves, Colin P. Ed., Oxford University Press, New York.

[6] Harvey, A. (1997), "Trends, Cycles and Autoregressions," *The Economic Journal*, 107, 192-201.

[7] Hoaglin, D.C. (1976), "Theoretical Properties of Congruential Random-Number Generators: an Empirical View," *Memorandum NS-340*, Department of Statistics, Harvard University, Cambridge Massachusetts.

[8] Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press, Cambridge.

[9] Im, K., Pesaran, M. H., and Shin Y. (1995), "Testing for Unit Roots in Heterogenous Panels," Manuscript, Department of Applied Economics, University of Cambridge.

[10] Kao, C. (1997), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics*, forthcoming.

[11] Kao, C., and Chiang, M-H. (1997), "On the Estimation and Inference of a Cointegrated Regression in Panel Data," submitted to *Econometric Theory.*

[12] Karian, Z. A., and Dudewicz, E. J. (1991), *Modern Statistical, Systems, and GPSS Simulation: The First Course*, W. H.. Freeman and Company, New York.

[13] Kennedy, W. J., and Gentle, J. E. (1980), *Statistical Computing*, Marcel Dekker, Inc., New York.

[14] Kinderman, A. J., and Ramage, J. G. (1976), "Computer Generation of Normal Random Numbers," *Journal of the American Statistical Association*, Vol. 71, Number 356, 893-896.

[15] Levin, A., and Lin, C.-F. (1993), "Unit Root Tests in Panel Data: New Results," Discussion Paper, Department of Economics, UC-San Diego.

[16] Maddala, G. S., and Wu, S. (1996), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test: Evidence From Simulations and the Bootstrap," Manuscript, Department of Economics, Ohio State University.

[17] McCoskey, S., and Kao, C. (1998), "A Residual-Based Test of the Null of Cointegration in Panel Data," *Econometric Reviews*, 17, 57-84.

[18] Pedroni, P. (1997), "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis," Manuscript, Department of Economics, Indiana University.

[19] Pesaran, M. H. (1997), "The Role of Economic Theory in Modelling the Long Run," *The Economic Journal*, 107, 178-191.

[20] Phillips, P. C. B. (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, 311-340.

[21] Phillips, P. C. B., and Loretan, M. (1991), "Estimating Long-Run Economic Equilibria," *Review of Economic Studies*, 58, 407-436.

[22] Phillips, P. C. B., and Moon, H. (1997), "Linear Regression Limit Theory for Nonstationary Panel Data," Cowles Foundation Discussion Paper.

[23] Phillips P. C. B., and Ouliaris, S. (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58, 165-193.

[24] Phillips, P. C. B., and Perron, P. (1988), "Testing for a Unit Root in Time Series Regressions," *Biometrika*, 75, 335-346.

[25] Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7, 1-21.

[26] Shin, Y. (1994), "A Residual Based Test of the Null of Cointegration Against the Alternative of No Cointegration," *Econometric Theory*, 10, 91-115.

[27] Stock, J., and Watson, M. (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783-820.

[28] Taylor, A. M. R., and Dixon, H. D. (1997), "Controversy: On Modelling the Long Run in Applied Economics," *The Economic Journal*, 107, 165-168.

Table 1: Preliminary Comparison of Empirical Size-DGP A

		ADF*	PO_t^*	PO_α^*	APG*
T=15					
	N=1	.0782	.0887	.0063	.1045
	N=15	.0577	.1339	.0000	.1014
	N=25	.0456	.1557	.0000	.1054
	N=50	.0349	.1945	.0000	.1149
	N=100	.0229	.2789	.0000	.1371
T=25					
	N=1	.0656	.0690	.0301	.0784
	N=15	.0493	.0587	.0039	.0524
	N=25	.0430	.0593	.0017	.0467
	N=50	.0392	.0563	.0004	.0379
	N=100	.0271	.0559	.0003	.0311
T=50					
	N=1	.0520	.0431	.0444	.0487
	N=15	.0492	.0257	.0198	.0278
	N=25	.0457	.0204	.0171	.0208
	N=50	.0420	.0147	.0105	.0150
	N=100	.0379	.0083	.0055	.0079
T=100					
	N=1	.0470	.0353	.0593	.0404
	N=15	.0518	.0156	.0356	.0222
	N=25	.0505	.0106	.0338	.0148
	N=50	.0486	.0049	.0266	.0077
	N=100	.0499	.0019	.0224	.0035

Notes:
(a) Size based on one-sided test with critical value equal to -1.645.

Table 2: Preliminary Comparison of Power-DGP A

	ADF*		PO_t^*		PO_α^*		APG*	
ρ	0.95	0.75	0.95	0.75	0.95	0.75	0.95	0.75
				T=15				
N=1	.0506	.0631	.0574	.1025	.0556	.0984	.0573	.1036
N=15	.0738	.2412	.0993	.5631	.1016	.5678	.1196*	.6904*
N=25	.0934	.3804	.1291	.7782	.1218	.7648	.1510*	.8896*
N=50	.1164	.6141	.1818	.9682	.1698	.9649	.2379*	.9953*
N=100	.1746	.8859	.2910	.9995	.2566	.9992	.3981*	.9999
				T=25				
N=1	.0541	.1013	.0600	.1566	.0597	.1548	.0611	.1620
N=15	.1040	.6736	.1572	.9549	.1445	.9408	.1709	.9813*
N=25	.1367	.8666	.2086	.9962	.1955	.9939	.2527*	.9997
N=50	.1897	.9913	.3469	.9999	.3068	.9999	.4365*	.9999
N=100	.3153	.9999	.5621	.9999	.5036	.9999	.7289*	.9999
				T=50				
N=1	.0607	.2671	.0780	.4342	.0780	.4292	.0778	.4333
N=15	.1970	.9998	.3443	.9999	.3118	.9999	.3786*	.9999
N=25	.2797	.9999	.5184	.9999	.4547	.9999	.5720*	.9999
N=50	.4882	.9999	.7849	.9999	.7087	.9999	.8653*	.9999
N=100	.7567	.9999	.9963	.9999	.9282	.9999	.9903	.9999
				T=100				
N=1	.0918	.7898	.1237	.9398	.1251	.9420	.1231	.9391
N=15	.5695	.9999	.8287	.9999	.7612	.9999	.8610*	.9999
N=25	.8019	.9999	.9636	.9999	.9314	.9999	.9826*	.9999
N=50	.9782	.9999	.9996	.9999	.9977	.9999	.9999	.9999
N=100	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Notes:
(a) * indicates the power is "significantly greater" when compared with any of the other three tests.

Table 3: Critical Tail Values for DGP-A

	ADF^*		APG^*		LM^*	
	.05	.95	.05	.95	.05	.95
			T=15			
N=1	-1.9994	1.9921	-2.1033	1.1027	-1.0357	280.4795
N=15	-1.7377	2.3097	-2.0477	1.4099	32.0205	846.2267
N=25	-1.5913	2.3799	-2.0931	1.4207	67.8861	.986.4166
N=50	-1.4249	2.5299	-2.1508	1.3783	154.9916	1176.9469
N=100	-1.1850	2.8383	-2.2683	1.2876	(c)	
			T=25			
N=1	-1.8039	1.8774	-1.878	1.1559	-0.2937	2344.6887
N=15	-1.6362	1.9783	-1.6736	1.6555	303.8285	6359.1715
N=25	-1.5702	2.0768	-1.6179	1.6926	618.700	7273.6785
N=50	-1.4993	2.1560	-1.5287	1.7606	1337.6544	8528.3492
N=100	-1.3563	2.3019	-1.4160	1.8755	2612.4676	10215.484
			T=50			
N=1	-1.6623	1.8015	-1.6371	1.1605	16.6210	29780.201
N=15	-1.6320	1.8054	-1.3997	1.7719	4144.7633	73931.707
N=25	-1.5972	1.8449	-1.2919	1.8912	8079.4509	83643.576
N=50	-1.5437	1.8678	-1.1139	2.0745	17270.721	117259.26
N=100	-1.5321	1.9239	-0.9029	2.3417	32972.924	117259.26
			T=100			
N=1	-1.6165	1.7408	-1.5679	1.1806	212.8913	293334.36
N=15	-1.6710	1.6960	-1.2801	1.8369	43127.502	732973.87
N=25	-1.6511	1.7177	-1.1576	1.9878	83414.588	819136.24
N=50	-1.6326	1.7131	-0.9082	2.2406	176527.25	961931.94
N=100	-1.6427	1.7243	-0.5960	2.5427	329494.67	1167615.0

Notes:
(a) ADF^* and APG^* are derived under the null of no cointegration.
(b) Theory predicts that ADF^* and APG^* should be asymptotically standard normal.
(c) The results for T=15 and N=100 were unobtainable using the fm procedure in COINT 2.0.

Table 4: Critical Tail Values for DGP-B

	ADF^*		APG^*		LM^*	
	.05	.95	.05	.95	.05	.95
			T=15			
N=1	-3.4193	0.4025	-5.1601	-1.0647	-1.0861	1.5511
N=15	-6.8738	-2.8324	-12.9827	-8.8675	-3.2473	1.2969
N=25	-8.2807	-4.2391	-16.0953	-12.0475	-3.8934	0.6806
N=50	-10.7820	-6.7398	-21.9040	-17.7879	-4.9465	-0.0879
N=100	(c)					
			T=25			
N=1	-4.1307	-0.5586	-5.9456	-2.1471	-1.0595	2.1618
N=15	-10.1258	-6.4900	-16.7929	-12.9665	-2.7912	2.0879
N=25	-12.5117	-8.9496	-21.1169	-17.2411	-3.2033	1.7417
N=50	-16.9732	-13.3324	-28.9876	-25.1793	-3.9026	1.2486
N=100	-23.2049	-19.5680	-40.2533	-36.3506	-4.7454	0.5269
			T=50			
N=1	-5.6605	-2.2999	-7.7609	-4.2157	-1.0129	2.2434
N=15	-16.6418	-13.2351	-24.4750	-20.8528	-2.1795	2.3093
N=25	-20.9452	-17.5617	-31.0713	-27.4119	-2.4221	2.2149
N=50	-28.9441	-25.5184	-43.1480	-39.4899	-2,7129	1.9091
N=100	-40.1773	-36.8097	-60.2685	-56.5836	-3.1108	1.7284
			T=100			
N=1	-8.0143	-4.7625	-10.5901	-7.1763	-0.9541	2.2391
N=15	-26.0655	-22.8093	-35.7923	-32.2859	-1.8140	2.2426
N=25	-33.2090	-29.9421	-45.7346	-42.2381	-1.9298	2.1714
N=50	-46.2659	-43.0329	-63.9175	-60.3727	-2.0677	2.1417
N=100	-64.7422	-61.5001	-89.6306	-86.1329	-2.2664	1.9821

Notes:
(a) LM^* is derived under the null of cointegration.
(b) Theory predicts that LM^* should be asymptotically standard normal.
(c) The results for T=15 and N=100 were unobtainable using the fm procedure in COINT 2.0.

Table 5: Empirical Rejection Rates

	ADF^*	APG^*	LM^*
	DGP-A	DGP-A	DGP-B
		T=15	
N=1	.0782	.1045	.0475
N=15	.0577	.1014	.0405
N=25	.0456	.1054	.0266
N=50	.0349	.1149	.0147
N=100	.0229	.1371	(b)
		T=25	
N=1	.0656	.0784	.0638
N=15	.0493	.0524	.0666
N=25	.0430	.0467	.0544
N=50	.0392	.0379	.0371
N=100	.0271	.0311	.0183
		T=50	
N=1	.0520	.0487	.0735
N=15	.0492	.0278	.0883
N=25	.0457	.0208	.0816
N=50	.0420	.0150	.0655
N=100	.0379	.0079	.0533
		T=100	
N=1	.0470	.0404	.0736
N=15	.0518	.0222	.0917
N=25	.0505	.0148	.0868
N=50	.0486	.0077	.0880
N=100	.0499	.0035	.0734

Notes:
(a) Empirical critical value for DGP-A is -1.645 and for DGP-B, 1.645.
(b) Results for N=100 and T=15 were unobtainable for the fm routine of COINT 2.0.

Table 6: Power to Reject: DGP-A

	ADF^*			APG^*			LM^*		
ρ	0.95	0.85	0.75	0.95	0.85	0.75	0.95	0.85	0.75
					T=15				
N=1	.0506	.0565	.0631	.0573	.0781	.1036*	.0573	.0665	.0771
N=15	.0738	.1337	.2412	.1196	.3571	.6904*	.1403*	.3647	.6030
N=25	.0934	.2048	.3804	.1510	.5303	.8896*	.1910*	.5454*	.8089
N=50	.1164	.3149	.6141	.2379	.8255*	.9953	.2988*	.7962	.9961
N=100	.1746	.5310	.8859	.3981	.9795	.9999	(b)		
					T=25				
N=1	.0541	.0746	.1013	.0611	.0987	.1620*	.0652	.0921	.1357
N=15	.1040	.3089	.6736	.1709	.6706	.9813*	.2397*	.7170*	.9566
N=25	.1367	.4525	.8666	.2527	.8769	.9997	.3757*	.9062*	.9973
N=50	.1897	.7268	.9913	.4395	.9941	.9999	.5978*	.9929	.9999
N=100	.3153	.9483	.9999	.7280	.9999	.9999	.8605*	.9998	.9999
					T=50				
N=1	.0607	.1277	.2671	.0778	.1980	.4333*	.1081*	.2047	.3923
N=15	.1970	.8945	.9998	.4084	.9980	.9999	.6652*	.9969	.9999
N=25	.2797	.9881	.9999	.5720	.9999	.9999	.9157*	.9999	.9999
N=50	.4882	.9999	.9999	.8653	.9999	.9999	.9999*	.9999	.9999
N=100	.7567	.9999	.9999	.9903	.9999	.9999	.9999	.9999	.9999
					T=100				
N=1	.0918	.3857	.7898	.1231	.5484	.9391*	.2533*	.5382	.8817
N=15	.5695	.9999	.9999	.8610	.9999	.9999	.9999*	.9999	.9999
N=25	.8019	.9999	.9999	.9826	.9999	.9999	.9999*	.9999	.9999
N=50	.9782	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
N=100	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Notes:
(a) * indicates the rejection rate is "significantly greater" when compared with either of the other two tests.
(b) Results for N=100 and T=15 were unobtainable for the fm routine in COINT 2.0.

Table 7: Power to Reject: DGP-B

	*ADF**			*APG**			*LM**		
θ	0.05	0.15	0.25	0.05	0.15	0.25	0.05	0.15	0.25
					T=15				
N=1	.0516	.0638	.0854	.0562	.0646	.0897	.0619	.1066*	.1797*
N=15	.0539	.0933	.1787	.0558	.1111	.2504	.0815*	.2872*	.6460*
N=25	.0560	.1126	.2383	.0587	.1368	.3459	.0951*	.3989*	.8290*
N=50	.0573	.1368	.3510	.0597	.1789	.5181	.1141*	.5798*	.9717*
N=100	(b)								
					T=25				
N=1	.0553	.0960	.1550	.0537	.0886	.1555	.0680	.1705*	.3219*
N=15	.0619	.2168	.5436	.0637	.2325	.6364	.1146*	.6375*	.9676*
N=25	.0704	.3053	.7349	.0660	.3157	.7980	.1423*	.8158*	.9970*
N=50	.0781	.4598	.9270	.0790	.4993	.9665	.1868*	.9659*	.9999*
N=100	.0886	.6780	.9964	.0902	.7220	.9996	.2639*	.9995*	.9999
					T=50				
N=1	.0676	.2150	.4152	.0654	.1947	.3941	.1083*	.3943*	.6566*
N=15	.1081	.7904	.9959	.1005	.7653	.9964	.3163*	.9937*	.9999
N=25	.1315	.9291	.9999	.1201	.9027	.9999	.4190*	.9999*	.9999
N=50	.1863	.9948	.9999	.1614	.9941	.9999	.6462*	.9999	.9999
N=100	.2801	.9999	.9999	.2325	.9999	.9999	.8564*	.9999	.9999
					T=100				
N=1	.1042	.5100	.8032	.0962	.4692	.7796	.2254*	.7224*	.9262*
N=15	.3280	.9999	.9999	.2613	.9999	.9999	.8635*	.9999	.9999
N=25	.4592	.9999	.9999	.3814	.9999	.9999	.9674*	.9999	.9999
N=50	.7100	.9999	.9999	.9999	.9999	.5686	.9998*	.9999	.9999
N=100	.9141	.9999	.9999	.8141	.9999	.9999	.9999	.9999	.9999

Note:
(a) * indicates the rejection rate is "significantly greater" when compared with either of the other two tests.
(b) Results for T=15 and N=100 were unobtainable for the fm routine in COINT 2.0.

Table 8: 5 Percent Tail and Empirical Size for Common Slopes Model

		ADF_K		ADF_{SM}	
		Tail	Size	Tail	Size
T=15					
	N=1	-2.2251	.1249	-1.8561	.0720
	N=15	-2.2924	.1605	-2.2872	.1192
	N=25	-2.4511	.1967	-2.2230	.1152
	N=50	-2.7184	.2866	-2.1832	.1059
	N=100	-3.1045	.4386	-2.0541	.0931
T=25					
	N=1	-2.1751	.1363	-1.9857	.0986
	N=15	-2.0853	.1177	-2.2041	.1184
	N=25	-2.1239	.1200	-2.0696	.1011
	N=50	-2.2963	.1532	-2.0278	.0907
	N=100	-2.4573	.2095	-2.0278	.0838
T=50					
	N=1	-2.2422	.1660	-2.2358	.1557
	N=15	-1.9666	.0971	-2.0935	.1058
	N=25	-1.9879	.0975	-2.0023	.0905
	N=50	-1.9721	.0958	-1.8866	.0760
	N=100	-2.1021	.1159	-1.8393	.0734
T=100					
	N=1	-2.3050	.1863	-2.3723	.2051
	N=15	-1.8988	.0840	-1.9824	.0970
	N=25	-1.8973	.0808	-1.9344	.0843
	N=50	-1.8872	.0797	-1.8594	.0733
	N=100	-1.8849	.0811	-1.7826	.0685

Notes:
(a) Size based on one-sided test with critical value equal to -1.645.

Table 9: Size Adjusted Power for Common Slopes Model

	ADF_K			ADF_{SM}		
ρ	0.95	0.85	0.75	0.95	0.85	0.75
			T=15			
N=1	.0508*	.0439*	.0361*	.0353	.0204	.0124
N=15	.0941	.0796	.0418	.2479*	.7204*	.9400*
N=25	.1116	.0947	.0436	.3768*	.9169*	.9965*
N=50	.1477	.1278	.0415	.6137*	.9963*	.9999*
N=100	.2117	.1913	.0508	.8874*	.9999*	.9999*
			T=25			
N=1	.0497*	.0459*	.0424*	.0383	.0239	.0135
N=15	.1387	.2067	.2477	.3319*	.8928*	.9982*
N=25	.1774	.3009	.3804	.5180*	.9896*	.9999*
N=50	.2780	.5129	.6363	.7995*	.9999*	.9999*
N=100	.4824	.8068	.9142	.9795*	.9999*	.9999*
			T=50			
N=1	.0518*	.0797*	.1016*	.0314	.0266	.0185
N=15	.2988	.8879	.9929	.5582*	.9999*	.9999*
N=25	.4484	.9844	.9999	.7971*	.9999*	.9999
N=50	.7588	.9998	.9999	.9999*	.9999	.9999
N=100	.9540	.9999	.9999	.9999*	.9999	.9999
			T=100			
N=1	.0841*	.2480*	.4285*	.0427	.0637	.0467
N=15	.8373	.9999	.9999	.9594*	.9999	.9999
N=25	.9697	.9999	.9999	.9978*	.9999	.9999
N=50	.9999	.9999	.9999	.9999	.9999	.9999
N=100	.9999	.9999	.9999	.9999	.9999	.9999

Notes:
(a) * indicates the power is "significantly greater" when compared with either of the other two tests.

Table 10: Empirical Size and Power with Different Parameter Values

	ADF_K		ADF_{SM}	
	Size	Power	Size	Power
		$\delta = -0.5$		
$\pi = -0.8$	.1921	.9999	.9999	.9999
$\pi = 0.0$	.0976	.9999	.0817	.9999
$\pi = 0.8$	.6417	.9999	.0001	.9999
		$\delta = 0.0$		
$\pi = -0.8$	.0084	.9999	.9999	.9999
$\pi = 0.0$	.0954	.9999	.0763	.9999
$\pi = 0.8$	.4073	.9999	.0000	.9999
		$\delta = 0.5$		
$\pi = -0.8$	.1852	.9999	.9999	.9999
$\pi = 0.0$	.1010	.9999	.0794	.9999
$\pi = 0.8$	.6399	.9999	.0000	.9999

Notes:
(a) N=T=50.
(b) $\rho = .75$.

Center for Policy Research Working Paper Series

Number	Author(s)	Title
1	Kao and Emerson	On the Estimation of a Linear Time Trend Regression with a One-Way Error Component Model in the Presence of Serially Correlated Errors
2	Kao and Chiang	On the Estimation and Inference of a Cointegrated Regression in Panel Data
3	McCoskey and Kao	A Monte Carlo Comparison of Tests for Cointegration in Panel Data