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: December 20, 1997

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 five 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

*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 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 consistent 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 (1995) and McCoskey and Kao (1997). 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 the tests of the null hypothesis of no cointegration in panel data both currently in the literature and proposed for the first time here; Section 3 summarizes the test of the null hypothesis of cointegration in panel data described in detail in Chapter 2 of the paper; Section 4 explains the Monte Carlo design for the comparison of the tests; Section 5 summarizes the results of the Monte Carlo experiment; Section 5 is the conclusion of the chapter.

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 $BM(\Omega)$ 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

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 currently available in the literature. 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 efficient 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

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

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

and

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

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:

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

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 $\rho = 1$.

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

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 $T \to \infty$:

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

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 efficient 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. Such test statistics are discussed in Pedroni (1995).

2.2 Pedroni (1995)

Another paper dealing with residual based tests in the presence of the spurious regression problem is the paper by Pedroni (1995). The results of the paper are very similar to the paper by Kao (1997) with two important differences: (a) Pedroni bases his residual tests not on the parametric DF and ADF tests, but rather on the non-parametric Phillips-Perron test for non-stationarity; and (b) Pedroni investigates a variety of different models with different panel specifications.

In his paper, Pedroni allows different assumptions on the homogeneity and heterogeneity of the panel data. In particular, Pedroni introduces a model which allows for differing intercepts and differing slopes and heterogenous long-run variance covariance matrices. The question of common or differing slopes is an important one. If slopes are allowed to differ across the cross-sections than estimation of the slopes will be equivalent to estimation in the strict time series case. If the slopes are restricted to a common value then the estimation will require more complex, pooled techniques.

Using a pooled unstandardized version of the Phillips and Perron statistic, Pedroni finds that models with common slopes and homogeneous errors produce statistics which converge in distribution to normal random variables. This result is analogous to the result from Kao (1997), except that Pedroni's results (see his Proposition 2) are based on the overly restrictive assumption that the regressors are strictly exogenous. When panels are allowed more heterogeneity, the statistics are more complex and depend on variance ratio statistics as well as Monte Carlo generated empirical moments. However, with the appropriate information and normalization, these statistics, too, will converge in distribution to normal random variables.

The first test from Pedroni is based on a pooled Phillips and Perron type test. The panel model assumes common slopes and common intercepts. This homogeneity has the interpretation that under the alternate hypothesis, the cointegrating vector for each cross-section is identical. The model is also general enough to allow for no intercept, $\alpha = 0$, and time trend, $\delta t \neq 0$, models for which Pedroni provides similar results. For convenience a specific case is presented.

$$y_{it} = \alpha + \beta x_{it} + e_{it},\ i = 1, ..., N,\ t = 1, ..., T \tag{11}$$

$$y_{it} = y_{it-1} + u_{it} \tag{12}$$

$$x_{it} = x_{it-1} + \varepsilon_{it}. \tag{13}$$

The test statistic then becomes:

$$\hat{\rho}_{NT} - 1 = \frac{\Sigma_{i=1}^{N} \Sigma_{t=2}^{T} (\hat{e}_{it-1\Delta} \hat{e}_{it-1} - \hat{\lambda}_i)}{(\Sigma_{i=1}^{N} \Sigma_{t=2}^{T} \hat{e}_{it-1}^2)} \tag{14}$$

where

$$\hat{\lambda}_i = \frac{1}{T} \sum_{s=1}^{k_i} w_{sk_i} \sum_{t=1}^{T} \widehat{\epsilon}_{it} \widehat{\epsilon}_{it-s},$$

and

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

Under the null hypothesis of no cointegration, e_{it} is a random walk and $\hat{\lambda}_i$ corrects for any correlation effects. $\hat{\lambda}_i$ acts as a scalar equivalent to Γ from the long-run variance covariance matrix. The estimation of $\hat{\lambda}_i$ is done with non-parametric kernel techniques with w_{sk_i} as an appropriately chosen Bartlett window. Thus, this correction is made in calculating the test statistic. The limiting distributions are given:

$$T\sqrt{N}(\hat{\rho}_{NT} - 1) \Rightarrow N(0, 2),$$

$$\sqrt{NT(T-1)}(\hat{\rho}_{NT} - 1) \Rightarrow N(0, 2),$$

and

$$t_{\hat{\rho}_{NT}} \Rightarrow N(0, 1).$$

As can be seen, these limiting distributions are independent of the nuisance parameters which were present in the parametric tests. However as hinted from above, these results may not be entirely correct. The asymptotics of the tests are based on the asymptotics of the spurious regression estimation and Pedroni seems to have made an overly restrictive assumption about the exogeneity between the regressor and error term in the spurious regression. Thus while $\hat{\lambda}_i$ corrects for correlation effects, there are no corrections made for possible weak exogeneity.

The second set of statistics from the Pedroni paper are based on a heterogeneous panel model which allows the cointegrating vectors under the alternative hypothesis to vary across the cross-section:

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it},\ i = 1, ..., N,\ t = 1, ..., T \tag{15}$$

where y_{it} and x_{it} are again random walks.

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

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

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

These corrections are now "complete" in the sense that $\hat{L}_{11i}$ and $\hat{\lambda}_i$ take into account the full information from the long-run variance-covariance matrix and for cross-correlation. 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.

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

$$Z_{t_{\hat{\rho}_{NT}}} \Rightarrow \frac{\frac{1}{\sqrt{N}}\Sigma_{i=1}^{N}\int Q_i dQ_i}{\sqrt{(\frac{1}{N}\Sigma_{i=1}^{N}\int Q_i^2)(1+\frac{1}{N}\Sigma_{i=1}^{N}\tilde{\beta}_i^2)}}, \tag{16}$$

as $T \to \infty$. Hence Pedroni claimed 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

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

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 (1995). 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? A more straightforward test could be based on the overall average across the cross-sectional units of a non-parametric test statistic. 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.

Finally, Pedroni discusses an application of his test, pooling cross-country results on the purchasing power parity hypothesis. The model is written as the log of nominal bilateral US dollar exchange rate for a given country in a given year equal to an intercept term constant across time, varying across country added to a coefficient multiplied by the log price differential between the given country and the US for a given year added to an error term for the given country and given year. In tests on individual series, the cointegration tests of the PPP hypothesis have been unable to reject the null of no PPP in post 1973 data. Data over long periods of time have tended to reject the result, although many economists are particularly interested in post-1973, Bretton Woods, data. By pooling the data in a panel model allowing for differing slopes and intercepts, Pedroni is able to reject the null hypothesis for a variety of models at the 10% level or better.

2.3 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.

$$y_{it} = \alpha_i + x_{it}'\beta_i + e_{it}, \; i = 1, ..., N, \; t = 1, ..., T. \tag{17}$$

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. 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:

$$\hat{e}_{it} = \rho_i \hat{e}_{it-1} + \sum_{j=1}^{p} \vartheta_{ij} \Delta \hat{e}_{it-j} + v_{itp}, \tag{18}$$

where $\hat{e}_{it}$ are OLS residuals from (17). An equivalent way to write equation (18) 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:

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

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

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

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

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

and

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

Consider equation (16), Pedroni 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. The limiting distribution of $\bar{t}_{ADF}$,

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

as the Pedroni statistic, is free of nuisance parameters.

Define

$$E[\int RdS] = \mu_{Adf}, \quad Var[\int RdS] = \sigma^2_{Adf},$$

it can be shown that:

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

The moments μ_{Adf} and σ^2_{Adf} can be found through simulation. 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. A description of how these moments were simulated along with appropriate values of 2 to 5 regressors is provided in the Appendix.

2.4 Average Phillips Z_t Statistic for Varying Slopes

The tests from Pedroni outlined above, are based on the idea that non-parametric correction is desirable for residual-based tests of cointegration. 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:

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

and

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

These terms are used to calculate the final statistic:

$$\hat{Z}_{it} = (\sum_{t=2}^{T} \hat{e}_{it-1}^2)^{\frac{1}{2}}(\hat{\alpha} - 1)/s_{iTl} - \frac{1}{2}(s_{iTl}^2 - s_{iv}^2)\frac{1}{s_{iTl}(\frac{1}{T^2}\sum_{t=2}^{T}\hat{e}_{it-1}^2)^{\frac{1}{2}}}. \tag{19}$$

Phillips and Ouliaris 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:

$$\sqrt{N}(\bar{Z}_t - \mu_{Adf}) \Rightarrow N(0, \sigma_{Adf}^2). \tag{20}$$

2.5 Corrected Panel ADF Estimator

The common slopes ADF test has been shown in Kao (1997) to have nuisance parameters in the limiting distribution. Here corrections are proposed to correct for this problem. Ultimately, the corrections will be similar to those used by Pedroni (1995) in his heterogeneous test.

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 approch 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

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

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

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

Banerjee et. al. show 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 $s_v \to d^2(1)\sigma^2$, 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

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

These corrections are similar in spirit to those used by Pedroni although here the corrections are done on the ADF statistic.

2.6 Comments

The articles by Kao and Pedroni present two 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 Perron test, which corrects 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 and this paper proposes three additional tests: two for varying slopes and one for common slopes. 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 is a complicated affair and the estimation procedure relies on non-parametric kernel methods.

Ω can be estimated by

$$\widehat{\Omega}=\left\{\frac{1}{T}\sum_{t=1}^{T}\hat{\xi}_i\widehat{\xi}_i^{'}+\frac{1}{T}\sum_{\tau=1}^{l}\varpi_{\tau l}\sum_{t=\tau+1}^{T}\left(\widehat{\xi}_i\widehat{\xi}_{it-\tau}^{'}+\widehat{\xi}_{it-\tau}\widehat{\xi}_{it}^{'}\right)\right\}, \tag{21}$$

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 $\tau>l$.

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

3.1 McCoskey and Kao (1997)

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 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 (1997). 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:

$$y_{it} = \alpha_i + x_{it}^{'}\beta_i + e_{it},\ i = 1, ..., N,\ t = 1, ..., T, \tag{22}$$

$$x_{it} = x_{it-1} + \varepsilon_{it} \tag{23}$$

$$e_{it} = \gamma_{it} + u_{it}, \tag{24}$$

and

$$\gamma_{it} = \gamma_{it-1} + \theta u_{it}. \tag{25}$$

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

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

$$\overline{LM} = \frac{\frac{1}{N}\Sigma_{i=1}^{N}\frac{1}{T^2}\Sigma_{t=1}^{T}S_{it}^{+2}}{s^{+^2}}, \tag{26}$$

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 two test is:

$$\sqrt{N}(\overline{LM} - \mu_v) \Rightarrow N(0, \sigma_v^2), \tag{27}$$

where μ_v and σ_v^2 are defined in McCoskey and Kao (1997). 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 and is suggested for use.

3.2 Comments

The asymptotics of the panel tests take advantage of the mathematics of triangular arrays 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.

4 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 three tests are considered: the $\bar{t}_{ADF}$ statistic, the $\overline{Z_t}$ statistic and $\overline{LM}$. The first two statistics are constructed under the null of no cointegration and the last under the null of cointegration. 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

4.1 The Random Number Generator

The random number generator used for this study is URN22 from Karian and Dudewicz (1991). This is a multiplicative congruential method of the following form:

$$x_{i+1} = (x_i * 69069 + 1) \ mod \ 2^{32}$$

The original seed is set to 32007779.

The uniform random numbers are transformed into standard normal random variables using the Box-Muller transformation (Karian and Dudewicz p. 160). The standard normal random variables are used to construct a random walk variable for x_{it} in the regression equation and the residuals. The program automatically saves the last number generated to use as the beginning seed in the next replication.

The RNDNS command within GAUSS uses the fast acceptance-rejection algorithm proposed by Kindermand and Ramage (1976) for generating standard normal variables from uniform random numbers. The uniform random numbers are generated according to the multiplicative-congruential method discussed in Kennedy and Gentle (1980). This method can be written in the form $x_i = (ax_{i-1} + c)(\mathrm{mod}\, m)$. In GAUSS the default values are $c = 0$; $m = 2^{31} - 1$; and $a = 397204094$. This particular method performed well in spectral testing as done by Hoaglin (1976) but was outperformed by a number of other generating techniques when tested by Karian and Dudewicz (1991). In particular, URN 22 outperformed this routine.

4.2 Experimental Design

4.2.1 Varying Slopes and Varying Intercepts

$$y_{it} \quad = \quad \alpha_i + x_{it}^{'}\beta_i \quad + \quad e_{it},\ i = 1, ..., N,\ t = 1, ..., T. \tag{28}$$

The three tests compared are the: $\bar{t}_{ADF}$, $\bar{Z}_t$, and $\overline{LM}$. 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, $\rho = 1$. Under the null, the error term is a random walk. The study includes the following possible values for ρ :

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

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 $\theta = 0$. 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 θ :

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

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

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

4.2.2 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, $k = A$ or B for a given N and T.

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

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

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):

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

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.

4.2.3 Common Slopes and Varying Intercepts

$$
y_{it} = \alpha_i + x_{it}^{'}\beta_i + e_{it},\ i = 1, ..., N,\ t = 1, ..., T\ . \tag{29}
$$

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.

4.3 Test Statistics

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

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

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

and

$$LM^* = \frac{\sqrt{N}(\overline{LM} + 2.026)}{.104403}.$$

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

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

4.4 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:

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

and

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

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.

5 Results

5.1 Varying Slopes and Varying Intercepts

The first consideration of the three tests is the empirical size. Considering Table 1, both tests ADF^* and Z^* should be asymptotically standard normal (after the appropriate normalizations derived in the text) and thus the mass of their distributions should be symmetric, centered on zero, with 5% of the density either less than -1.645 or greater than 1.645. From the Table 1, ADF^* actually seems to approximate the standard normal distribution reasonably well. For small T, the distribution is shifted to the right somewhat, although as T grows this problem is less severe. Consider the change from $T = 15$ and $N = 100$ with 5% and 95% values of -1.1850 and 2.8383 respectively to $T = 100$ and $N = 100$ with values -1.6427 and 1.7243. For Z^*, however the standard normal distribution does not seem to be well approximated. For small T, the distribution is heavily skewed toward the left and for large T the distribution is heavily skewed to the right. Again consider the change from $T = 15$ and $N = 100$ with 5% and 95% values of -9.6859 and -2.7297 to $T = 100$ and $N = 100$ with values of 2.4866 and 5.8014.

This result is supported in Table 3 with the results for DGP A. A rejection rate of 5% is predicted by theory for ADF^* and Z^*. For ADF^* the actual rejection rate for all T and N is contained within the interval $[.0229, .0782]$ whose border values are both reached when $T = 15$. For T $\geq$ 50 all of the rejection rates fall between $[.0379, .0520]$. The results for Z^* are dramatically different. For T $\leq$ 25, the test overrejects the null, reaching a maximum rejection rate of .9877 when $T = 15$ and $N = 100$. For $T > 25$, the test underrejects for all cases except when $T = 50$ and $N = 100$.

Table 2 shows the distribution of LM^* under its null hypothesis. Given the appropriate normalization used in constructing LM^*, this statistic should be asymptotically standard normal with 5% of its density less than -1.645 and 5% greater than 1.645. Once again the results show that the critical values for LM^* do a fair job of approximating the those of the standard normal distribution. As T increases from 15 to 100, holding $N = 100$, the left and right 5% critical values change from -5.9154 and 0.4331 to -2.3677 and 1.3730, respectively. In Table 3 for LM^* and DGP-B, this result is further supported as the rejection rates all fall within the interval $[.0157, .0740]$. Overall, it seems that while ADF^* performs the best with regard to size, both ADF^* and LM^* do a reasonable job of approximating standard normal random variables for large T and N. On the other hand, Z^* shows huge size distortions which range from overrejection of the null hypothesis with small T to underrejection of the null hypothesis with large T. In all cases, when the tests are used on the DGP contrary to their own null, they reject the null of the DGP with a rejection rate close to 1.

Table 4 shows a comparison of the three tests in their ability to correctly reject the null hypothesis of DGP-A. Given the two step procedure of the experiment, the ability of the tests to correctly reject the null hypothesis should be directly comparable. Theory suggests that the ability to reject the null hypothesis should increase with larger T and N and increase with smaller ρ. When a test has a significantly greater power than the other two when compared pair-wise, that value is marked with a '*'. The first obvious result is that Z^* performs the best in almost all cases where a best test can be identified. This should not be so surprising considering the size distortions of this test. In only two cases LM^* dominates and in no cases does ADF^* dominate. As ADF^* and Z^* were derived under this DGP it should not be surprising that one of these tests is the most powerful. However, what is interesting is that in almost all cases the second most powerful test is not ADF^* but rather LM^*. All tests have a power which approaches 1 as N and T increase and in those cases where $T = 100$, $N \geq 15$, and $\rho \leq 0.85$ all of the tests are statistically equivalent.

Table 5 compares the performance of the tests with DGP-B. In this case the ability to reject should again increase with larger T and N and increase with larger θ. For $T < 25$, Z^* again dominates although for $T > 25$ LM^* is best, when a best test can be determined. In this case, only LM^* is derived under the null of this DGP and so it is encouraging that the test does, in certain cases, outperform the other tests although Z^* still performs well with regards to power. ADF^* again lags behind in its power, in particular for small θ.

Table 6 shows the critical values for DGP-B when autocorrelation and weak exogeneity values are specified. Theory shows that these affects should be corrected for asymptotically. Using the case where $\delta = 0$ and $\pi = 0$ as a benchmark (reported in detail in Table 1.1), it can be seen allowing these parameters to be $\neq 0$, seriously skews the distributions in finite samples. For ADF^* and Z^* allowing a negative moving average,

$\pi = -0.8$ causes the most serious changes. For ADF^* the critical values move from -1.5437 and 1.8678 to -20.2313 and -16.2978 when allowing $\pi = -0.8$ and keeping $\delta = 0$. For Z^* this change moves the values from 0.1459 and 3.78988 to -41.9818 and -35.2251. The empirical size results from DGP A in Table 8 support this observation. In all cases when there is a negative moving average, the empirical size for ADF^* and Z^* is equal to .9999 far from the empirical values .0420 and .0006 reported in Table 1. Including a positive moving average has a contradictory effect on the tests. For ADF^*, a positive moving average causes the empirical size to increase, for example from .0420 to .5470 when $\delta = 0$; for Z^* a positive moving average causes the empirical size to decrease, for example from .0006 to .0000. These patterns are observed across all values for δ. Table 7 shows the effect of these parameters on the distribution of LM^* under the null. In this case, a positive moving average has the largest effect on the distribution, heavily skewing the distribution to the right. For example holding $\delta = 0$ and increasing π from 0 to 0.8 changes the critical values from -2.6244 and 1.3693 to 32.0239 and 71.1100. In Table 8 the size results for LM^* and DGP-B show that a positive moving average, for all δ, causes serious overrejection of the null while all other combinations cause an empirical size close to 0.

The size-adjusted power for all tests is close to 1. These results are entirely consistent with the results in Tables 4 and 5. Changing δ and π has no serious impact on the power.

5.2 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-the adjustments 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 ρ =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.

6 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 tests from which to choose. Two of the tests are constructed under the null hypothesis of cointegration, ADF^* and Z^*. These tests are based on the ADF test and Z-test, first proposed in the time series case. 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 this test, LM^*, may in fact be the best to use. Although one of the tests of the null of no cointegration, Z^*, outperformed this test in terms of power, it had serious size distortions. The other test, ADF^*, had slightly better size results but was outperformed by the LM^* in terms of power.

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.

7 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., J. Dolado, J. W.. Galbraith, 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," Manuscript, Center for Policy Research, Syracuse University.

[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. (1997), "A Residual-Based Test of the Null of Cointegration in Panel Data," *Econometric Reviews*, forthcoming.

[18] Pedroni, P. (1995), "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: Critical Tail Values for DGP-A

	ADF^*		Z^*		LM^*	
	.05	.95	.05	.95	.05	.95
			T=15			
N=1	-1.9994	1.9921	-4.5032	1.8051	-1.1091	93.6244
N=15	-1.7377	2.3097	-6.0667	0.8082	.6704	695.7660
N=25	-1.5913	2.3799	-6.7103	0.1654	6.8815	928.4529
N=50	-1.4249	2.5299	-7.8779	-1.0480	28.8310	1226.3689
N=100	-1.1850	2.8383	-9.6859	-2.7297	77.2354	1435.5586
			T=25			
N=1	-1.8039	1.8774	-2.8089	1.8502	-0.9603	2072.2263
N=15	-1.6362	1.9783	-2.6779	2.1218	123.3981	11572.642
N=25	-1.5702	2.0768	-2.6614	2.0878	301.9243	14525.383
N=50	-1.4993	2.1560	-2.7173	2.0824	821.5418	17932.356
N=100	-1.3563	2.3019	-2.8331	1.9502	1969.8660	22108.834
			T=50			
N=1	-1.6623	1.8015	-1.6726	1.8851	4.8682	66675.878
N=15	-1.6320	1.8054	-0.8023	2.9020	4273.0974	321416.64
N=25	-1.5972	1.8449	-0.4700	3.2014	9690.2244	388028.85
N=50	-1.5437	1.8678	0.1459	3.7988	26697.279	459568.43
N=100	-1.5321	1.9239	0.9781	4.6703	61335.386	564145.59
			T=100			
N=1	-1.6165	1.7408	-1.3274	1.9015	129.7267	1438815.0
N=15	-1.6710	1.6960	-0.0428	3.2002	95401.811	6665604.8
N=25	-1.6511	1.7177	.4383	3.6767	222133.5	7977185.0
N=50	-1.6326	1.7131	1.3326	4.5706	586352.98	9799089.0
N=100	-1.6427	1.7243	2.4866	5.8014	1316189.6	11673094

Notes:
(a) ADF^* and Z^* are derived under the null of no cointegration.
(b) Theory predicts that ADF^* and Z^* should be asymptotically standard normal.

Table 2: Critical Tail Values for DGP-B

	ADF^*		Z^*		LM^*	
	.05	.95	.05	.95	.05	.95
			T=15			
N=1	-3.4193	0.4025	-9.3341	-1.1133	-1.1043	2.2669
N=15	-6.8738	-2.8324	-21.2169	-12.2681	-3.2382	2.5278
N=25	-8.2807	-4.2391	-25.9449	-17.0904	-3.8215	2.0858
N=50	-10.7820	-6.7398	-34.8334	-25.7286	-4.7402	1.5095
N=100	-14.4006	-10.3834	-47.1406	-38.0089	-5.9154	0.4331
			T=25			
N=1	-4.1307	-0.5586	-8.8107	-2.2460	-1.0596	2.0034
N=15	-10.1258	-6.4900	-22.8753	-15.9055	-2.6370	2.0252
N=25	-12.5117	-8.9496	-28.4773	-21.4661	-3.0255	1.6991
N=50	-16.9732	-13.3324	-38.6082	-31.6693	-3.6244	1.2457
N=100	-23.2049	-19.5680	-53.1609	-46.0509	-4.4334	0.5101
			T=50			
N=1	-5.6605	-2.2999	-10.1848	-4.5697	-0.9816	1.9680
N=15	-16.6417	-13.2351	-30.0891	-24.2488	-2.0756	1.8768
N=25	-20.9452	-17.5617	-38.0617	-32.0797	-2.2924	1.7359
N=50	-28.9441	-25.5184	-52.4632	-46.5230	-2.6244	1.3693
N=100	-40.1773	-36.8097	-72.8668	-66.9527	-3.1285	1.0117
			T=100			
N=1	-8.0143	-4.7625	-13.1823	-7.8894	-0.9202	1.8987
N=15	-26.0655	-22.8093	-42.5413	-37.2137	-1.7699	1.8385
N=25	-33.2090	-29.9421	-54.2189	-48.7420	-1.9018	1.7132
N=50	-46.2659	-43.0323	-75.3773	-70.0356	-2.1247	1.5703
N=100	-64.7422	-61.5001	-105.4764	-100.0805	-2.3677	1.3730

Notes:
(a) Z^* is derived under the null of cointegration.
(b) Theory predicts that Z^* should be asymptotically standard normal.

Table 3: Empirical Rejection Rates

	ADF^*		Z^*		LM^*	
	DGP-A	DGP-B	DGP-A	DGP-B	DGP-A	DGP-B
			T=15			
N=1	.0782	.3166	.2358	.8824	.2811	.0643
N=15	.0577	.9978	.6041	.9999	.9321	.0740
N=25	.0456	.9999	.7380	.9999	.9859	.0609
N=50	.0349	.9999	.9048	.9999	.9997	.0469
N=100	.0229	.9999	.9877	.9999	.9999	.0238
			T=25			
N=1	.0656	.6545	.1296	.9862	.6859	.0638
N=15	.0493	.9999	.1536	.9999	.9999	.0645
N=25	.0430	.9999	.1528	.9999	.9999	.0521
N=50	.0392	.9999	.1641	.9999	.9999	.0345
N=100	.0271	.9999	.1881	.9999	.9999	.0157
			T=50			
N=1	.0520	.9930	.0525	.9999	.9789	.0636
N=15	.0492	.9999	.0107	.9999	.9999	.0627
N=25	.0457	.9999	.0045	.9999	.9999	.0550
N=50	.0420	.9999	.0006	.9999	.9999	.0361
N=100	.0379	.9999	.0002	.9999	.9999	.0225
			T=100			
N=1	.0470	.9999	.0288	.9999	.9998	.0620
N=15	.0518	.9999	.0008	.9999	.9999	.0619
N=25	.0505	.9999	.0002	.9999	.9999	.0537
N=50	.0486	.9999	.0000	.9999	.9999	.0452
N=100	.0499	.9999	.0000	.9999	.9999	.0346

Notes:
(a) Empirical critical value for ADF^* and Z^* is -1.645 and for Test 3, 1.645.

Table 4: Power to Reject: DGP-A

	ADF^*			Z^*			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	.0564	.0743*	.0898*	.0512	.0456	.0437
N=15	.0738	.1337	.2412	.0996	.2643*	.4721*	.0927	.1602	.2330
N=25	.0934	.2048	.3804	.1343*	.3913*	.6899*	.1157	.2457	.3771
N=50	.1164	.3149	.6141	.1914*	.6470*	.9372*	.1732	.4222	.6495
N=100	.1746	.5310	.8859	.3071*	.9005*	.9985*	.2410	.6356	.8595
					T=25				
N=1	.0541	.0746	.1013	.0628*	.0965*	.1427*	.0577	.0695	.0797
N=15	.1040	.3089	.6736	.1962	.6552*	.9538*	.2065*	.5655	.8625
N=25	.1367	.4525	.8666	.2820	.8607*	.9976*	.2977*	.7678	.9680
N=50	.1897	.7268	.9913	.4659	.9916*	.9999	.4605	.9393	.9969
N=100	.3153	.9483	.9999	.7309*	.9999*	.9999	.6859	.9920	.9998
					T=50				
N=1	.0607	.1277	.2671	.0945*	.2329*	.4574*	.0782	.1659	.3133
N=15	.1970	.8945	.9998	.5502	.9987*	.9999	.5535	.9917	.9999
N=25	.2797	.9881	.9999	.7558*	.9999	.9999	.7290	.9993	.9999
N=50	.4882	.9999	.9999	.9600*	.9999	.9999	.9274	.9999	.9999
N=100	.7567	.9999	.9999	.9996	.9999	.9999	.9899	.9999	.9999
					T=100				
N=1	.0918	.3857	.7898	.1705*	.6286*	.9580*	.1467	.5251	.8761
N=15	.5695	.9999	.9999	.9776*	.9999	.9999	.9463	.9999	.9999
N=25	.8019	.9999	.9999	.9990*	.9999	.9999	.9892	.9999	.9999
N=50	.9782	.9999	.9999	.9999	.9999	.9999	.9995	.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.

Table 5: Power to Reject: DGP-B

	ADF^*			Z^*			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	.0601	.1094	.1906*	.0623	.0975	.1527
N=15	.0539	.0933	.1787	.0931*	.3148*	.6678*	.0801	.2424	.5240
N=25	.0560	.1126	.2383	.1111*	.4357*	.8436*	.0949	.3243	.7032
N=50	.0573	.1368	.3510	.1339*	.6446*	.9731*	.1124	.4788	.9069
N=100	.0638	.2026	.5527	.1773*	.8680*	.9994*	.1504	.7359	.9946
					T=25				
N=1	.0553	.0960	.1555	.0690	.1705	.3404*	.0702	.1675	.3123
N=15	.0619	.2168	.5436	.1262*	.6342*	.9735*	.1171	.6028	.9587
N=25	.0704	.3053	.7349	.1550*	.8114*	.9964	.1428	.7906	.9951
N=50	.0791	.4598	.9270	.2288*	.9669*	.9999	.2100	.9572	.9999
N=100	.0886	.6780	.9964	.3242*	.9993	.9999	.3125	.9992	.9999
					T=50				
N=1	.0676	.2150	.4152	.0932	.3734	.6939*	.1079*	.3838*	.6596
N=15	.1081	.7904	.9959	.2460	.9896	.9999	.3159*	.9940*	.9999
N=25	.1315	.9291	.9999	.3316	.9997	.9999	.4232*	.9999	.9999
N=50	.1863	.9948	.9999	.5180	.9999	.9999	.6563*	.9999	.9999
N=100	.2801	.9999	.9999	.7679	.9999	.9999	.8816*	.9999	.9999
					T=100				
N=1	.1042	.5100	.8032	.1499	.7124	.9564*	.2259*	.7204	.9288
N=15	.3280	.9999	.9999	.6210	.9999	.9999	.8617*	.9999	.9999
N=25	.4952	.9999	.9999	.7867	.9999	.9999	.9646*	.9999	.9999
N=50	.7098	.9999	.9999	.9636	.9999	.9999	.9989*	.9999	.9999
N=100	.9141	.9999	.9999	.9992	.9999	.9999	.9999	.9999	.9999

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

Table 6: Critical Tail Values for DGP-A with Different Parameter Values

	ADF^*		Z^*		LM^*	
	.05	.95	.05	.95	.05	.95
			$\delta = -0.5$			
$\pi = -0.8$	-20.9590	-17.0997	-45.8708	-38.9944	1.8812	39.6096
$\pi = 0.0$	-1.5869	1.8767	-2.5482	1.6997	11189.623	202965.42
$\pi = 0.8$	-3.0901	0.6049	6.9244	9.4277	358425.41	67776272.3
			$\delta = 0.0$			
$\pi = -0.8$	-20.2313	-16.2978	-41.9818	-35.2251	15.6162	105.0690
$\pi = 0.0$	-1.5437	1.8678	0.1459	3.78988	26697.279	459568.43
$\pi = 0.8$	-3.6441	0.0469	8.5413	10.7641	855233.03	15366379
			$\delta = 0.5$			
$\pi = -0.8$	-20.8811	-17.0931	-45.8382	-38.9254	2.2450	40.5881
$\pi = 0.0$	-1.5391	1.8658	-2.4949	1.7081	11447.613	200539.06
$\pi = 0.8$	-3.0244	0.5792	6.9635	9.4475	366944.11	6717846.5

Notes:
(a) ADF^* and Z^* are derived under the null of no cointegration.
(b) Theory predicts that ADF^* and Z^* should be asymptotically standard normal.
(c) N=T=50.
(d) $\rho = .75$.
(e) Table 1.1 critical values are the special case where $\delta = 0$ and $\pi = 0$.

Table 7: Critical Tail Values for DGP-B with Different Parameter Values

	ADF^*		Z^*		LM^*	
	.05	.95	.05	.95	.05	.95
			$\delta = -0.5$			
$\pi = -0.8$	-54.6160	-50.2815	-82.4793	-74.0113	-5.6814	-4.4792
$\pi = 0.0$	-28.0004	-24.6423	-51.5815	-45.8210	-5.6781	-3.9497
$\pi = 0.8$	-28.6830	-24.8357	-19.4054	-16.3693	8.8526	25.6634
			$\delta = 0.0$			
$\pi = -0.8$	-55.2317	-50.7890	-83.2971	-74.6613	-2.6665	0.1860
$\pi = 0.0$	-28.9441	-25.5184	-52.4632	-46.5230	-2.6244	1.3693
$\pi = 0.8$	-29.7631	-25.7923	-19.4745	-16.4779	32.0239	71.1100
			$\delta = 0.5$			
$\pi = -0.8$	-54.5773	-50.2465	-82.5288	-74.0214	-5.6861	-4.4770
$\pi = 0.0$	-27.9845	-24.6335	-51.6027	-45.7909	-5.6829	-3.9874
$\pi = 0.8$	-29.6632	-24.8378	-19.3873	-16.3945	8.8277	25.5352

Notes:
(a) LM^* is derived under the null of no cointegration.
(b) Theory predicts that LM^* should be asymptotically standard normal.
(c) N=T=50.
(d) $\theta = .25$.
(e) Table 2 critical values are the special case where $\delta = 0$ and $\pi = 0$.

Table 8: Empirical Size and Power with Different Parameter Values

	ADF^*		Z^*		LM^*	
	Size	Power	Size	Power	Size	Power
DGP-A						
			$\delta = -0.5$			
$\pi = -0.8$	.9999	.9999	.9999	.9838	.9569	.9999
$\pi = 0.0$	.0440	.9999	.1636	.9999	.9999	.9999
$\pi = 0.8$	.3647	.9999	.0000	.9999	.9999	.9999
			$\delta = 0.0$			
$\pi = -0.8$	.9999	.9999	.9999	.9999	.9999	.9999
$\pi = 0.0$	.0420	.9999	.0006	.9999	.9999	.9999
$\pi = 0.8$	.5470	.9999	.0000	.9999	.9999	.9999
			$\delta = 0.5$			
$\pi = -0.8$	.9999	.9999	.9999	.9830	.9647	.9999
$\pi = 0.0$	.0408	.9999	.1583	.9999	.9999	.9999
$\pi = 0.8$	.3562	.9999	.0000	.9999	.9999	.9999
DGP-B						
			$\delta = -0.5$			
$\pi = -0.8$	.9999	.8502	.9999	.9999	.0000	.9999
$\pi = 0.0$	.9999	.9999	.9999	.9999	.0000	.9999
$\pi = 0.8$	.9999	.9999	.9999	.9999	.9999	.9999
			$\delta = 0.0$			
$\pi = -0.8$	.9999	.8006	.9999	.9999	.0020	.9999
$\pi = 0.0$	.9999	.9999	.9999	.9999	.0361	.9999
$\pi = 0.8$	.9999	.9999	.9999	.9999	.9999	.9999
			$\delta = 0.5$			
$\pi = -0.8$	.9999	.8443	.9999	.9999	.0000	.9999
$\pi = 0.0$	.9999	.9999	.9999	.9999	.0000	.9999
$\pi = 0.8$	.9999	.9999	.9999	.9999	.9999	.9999

Notes:
(a) LM^* is derived under the null of no cointegration.
(b) Theory predicts that LM^* should be asymptotically standard normal.
(c) N=T=50.
(d) For DGP- A, $\rho = .75$; for DGP- B $\theta = .25$.

Table 9: 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 10: 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 11: 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) $\rho = .75$.