Forthcoming,
Oxford Bulletin of Economics and Statistics

Current Version: July 25, 1999

CRITICAL VALUES FOR COINTEGRATION TESTS IN HETEROGENEOUS PANELS WITH MULTIPLE REGRESSORS *

Peter Pedroni

Indiana University

mailing address: Economics, Indiana University
Bloomington, IN 47405
(812) 855-7925

email: PPEDRONI@INDIANA.EDU

Abstract: Asymptotic distributions and critical values are computed for several residual based tests of the null of no cointegration in panels for the case of multiple regressors, including regressions with individual specific fixed effects and time trends. The associated cointegrating vectors and the dynamics of the underlying error processes are permitted considerable heterogeneity across individual members of the panel.

Keywords: Time Series, Panel Cointegration, Critical Values. JEL Classifications: C22, C23.

* Acknowledgments: I thank Anindya Banerjee, David Canning and Peter Dunne for helpful comments and I thank Terri Ziacik for valuable research assistance. A computer program which implements these tests is available upon request from the author at ppedroni@indiana.edu.

CRITICAL VALUES FOR COINTEGRATION TESTS IN HETEROGENEOUS PANELS WITH MULTIPLE REGRESSORS

1. Introduction

In this paper we describe a method for testing the null of no cointegration in dynamic panels with multiple regressors and compute approximate critical values for these tests. Methods for nonstationary panels, including panel unit root and panel cointegration tests have been gaining increased acceptance in recent empirical research. To date, however, tests for the null of no cointegration in heterogeneous panels based on Pedroni (1995, 1997a) have been limited to simple bivariate examples, in large part due to the lack of critical values available for more complex multi-variate regressions. The purpose of this paper is to fill this gap by describing a method to implement tests for the null of no cointegration for the case with multiple regressors and to provide appropriate critical values for these cases. The tests allow for considerable heterogeneity among individual members of the panel, including heterogeneity in both the long run cointegrating vectors as well as heterogeneity in the dynamics associated with short run deviations from these cointegrating vectors.

1.1 Literature Review

Initial theoretical work on the nonstationary panel data focused on testing for unit roots in univariate panels. Early examples include Quah (1994), who studied the standard unit root null in panels with homogeneous dynamics, and Levin and Lin (1993) who studied unit root tests in panels with heterogeneous dynamics, fixed effects, and individual specific deterministic trends. These tests take the autoregressive root to be common under both the unit root null and the stationary alternative hypothesis. More recently, Im, Pesaran and Shin (1997) and Maddala and Wu (1999) suggest several panel unit root tests which also permit heterogeneity of the autoregressive root under the alternative hypothesis. Applications of panel unit root methods have included Bernard and Jones (1996), Coakley and Fuertes (1997), Evans and Karras (1996), Frankel and Rose (1996), Lee, Pesaran and Smith (1997), MacDonald (1996), O'Connell (1998), Oh (1996), Papell (1997), Wei and Parsely (1995), and Wu (1996).

However, many empirical questions involve multi-variate relationships and a researcher is interested to know whether or not a particular set of variables is cointegrated. Consequently Pedroni (1995, 1997a) studied the properties of spurious regression, and tests for the null of no cointegration in both homogeneous and heterogeneous panels. For the case with heterogeneous panels, Pedroni (1995, 1997a) provides asymptotic distributions for test statistics that are appropriate for various cases with heterogeneous dynamics, endogenous regressors, fixed effects, and individual specific deterministic trends. Pedroni (1997a) includes tests that are appropriate both for the case with common autoregressive roots under the alternative hypothesis as well as tests that permit heterogeneity of the autoregressive root under the alternative hypothesis in the spirit of Im, Pesaran and Shin (1997).

Applications of the panel cointegration tests developed in Pedroni (1995, 1997a) for the case with heterogeneous cointegrating vectors have included among others, Butler and Dueker (1999), Canzoneri, Cumby and Diba (1996), Chinn (1997), Chinn and Johnston (1996), Neusser and Kugler (1998), Obstfeld and Taylor (1996), Ong and Maxim (1997), Pedroni (1996b), and Taylor (1996). However, to date, these applications have been limited to cases in which the cointegrating regressions involved a single regressor. Many topics, on the other hand, involve applications in which more than a single regressor is likely to be required. Therefore, the purpose of this paper is to provide critical values that are appropriate for these situations based on the heterogeneous panel cointegration statistics developed in Pedroni (1995, 1997a).

Finally, the panel cointegration tests in this paper should be distinguished from those which are based on a maintained assumption of homogeneity of the cointegrating vectors among individual members of the panel. In addition to the heterogeneous case, Pedroni (1995, 1997a) also studied properties for the special case of homogeneous cointegrating vectors. Specifically, Pedroni (1995, 1997a) showed that for panels with homogeneous cointegrating vectors, an interesting special result holds such that residual based tests for the null of no cointegration have distributions that are asymptotically equivalent to raw panel unit root tests if and only if the regressors are exogenous. Kao (1999) further studied the special case in which cointegrating vectors are assumed to be homogeneous, but the asymptotic equivalency result is violated because of the endogeneity of regressors. An example of the application of these techniques for a test of the null of no cointegration in panels that are assumed to be homogeneous is the paper by Kao, Chiang and Chen (1999). By contrast, McCoskey and Kao (1999) examine the reversed null hypothesis of cointegration in their study of urbanization.

The remainder to the paper is structured as follows. Next, in section 2.1 we begin with a description of how the statistics can be used and how to construct them step by step. Then in section 2.2 we explain how the critical values can be computed for these statistics, and provide a table of adjustment terms that can be used to obtain appropriate critical values for various cases of interest with multiple regressors. Finally, section 3 offers a few concluding remarks.

2. Tests for the null of no cointegration in heterogeneous panels with multiple regressors

In the conventional time series case, cointegration refers to the idea that for a set of variables that are individually integrated of order one, some linear combination of these variables can be described as stationary. The vector of slope coefficients that render this combination stationary is referred to as the cointegrating vector. It is well known that this vector is generally not unique, and the question of how many cointegrating relationships exist among a certain set of variables is also an important question in many cases. In this study, we do not address issues of normalization or questions regarding the particular number of cointegrating relationships, but instead focus on reporting critical values for the case where we are interested in the simple null hypothesis of no cointegration versus cointegration.

Consequently, it is important to keep in mind that we are implicitly assuming for example that the researcher has in mind a particular normalization among the variables which is deemed sensible and is simply interested in knowing whether or not the variables are cointegrated. In this case, it is also well known that conventional tests often tend to suffer from unacceptably low power when applied to series of only moderate length, and the idea of pooling the data across individual members of a panel is intended to address this issue by making available considerably more information regarding the cointegration hypothesis. Thus, in effect, panel cointegration techniques are intended to allow researchers to selectively pool information regarding common long run relationships from across the panel while allowing the associated short run dynamics and fixed effects to be heterogeneous across different members of the panel.

In this case, one can think of such a panel cointegration test as being one in which the null hypothesis is taken to be that for each member of panel the variables of interest are not cointegrated and the alternative hypothesis is taken to be that for each member of the panel there exists a single cointegrating vector, although this cointegrating vector need not be the same for each member. Indeed, an important feature of these tests is that they allow not only the dynamics and fixed effects to differ across members of the panel, but also that they allow the cointegrating vector to differ across members under the alternative hypothesis. We consider this to be a valuable feature of the test, since in practice the cointegrating vectors are often likely not to be strictly homogeneous in such panels. In such cases, incorrectly imposing homogeneity of the cointegrating vectors in the regression would imply that the null of no cointegration may not be rejected despite the fact the variables are actually cointegrated.[1]

2.1 How to compute the test statistics

We now turn to a discussion of how to construct and implement tests for the null of no cointegration in such panels. Our first step is to compute the regression residuals from the hypothesized cointegrating regression. In the most general case, this may take the form

$$y_{i,t} = á_i + ä_i t + â_{1i} x_{1i,t} + â_{2i} x_{2i,t} + \dots + â_{Mi} x_{Mi,t} + e_{i,t}$$

$$\text{for} \quad t = 1,\dots,T \quad ; \quad i = 1,\dots,N \quad ; \quad m = 1,\dots,M \qquad \textbf{(1)}$$

where T refers to the number of observations over time, N refers to the number of individual members in the panel, and M refers to the number of regression variables. Since there are N different members of the panel, we can think of N different equations, each of which has M regressors. Notice that the slope coefficients $â_{1i}$, $â_{2i}$, ... , $â_{Mi}$ are permitted to vary across individual members of the panel. The parameter $á_i$ is the member specific intercept, or fixed effects parameter which of course is also allowed to vary across individual members. In addition, for some applications, we may also wish to include deterministic time trends which are specific to individual members of the panel, and are captured by the term $ä_i t$, although it will also often be the case that we choose to omit these $ä_i t$ terms.

Whether we include these member specific fixed effects or member specific time trends will in general affect the asymptotic distributions and the corresponding critical values just as in the conventional

[1] For further discussion, see also Pedroni (1998), which develops a test for whether or not individual members converge to a common cointegrating vector in such panels under the maintained assumption that at least some (possibly heterogeneous) cointegrating relationship exists for each member.

time series case. The reason for this is because in the presence of unit roots, sample averages such as $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{i,t}$ taken over the time series dimension, T, do not converge to population means as T grows large, but instead diverge at the rate $\sqrt{T}$. Therefore, when one constructs the panel statistic, with appropriate standardization with respect to N and T to ensure a stable distribution under the null hypothesis, the effect of estimating the sample means is not eliminated asymptotically, no matter how large N or T become. Similar arguments apply to the case of detrending. Consequently, as we will see in the next section, critical values have been tabulated for either of the cases depending on whether or not these have been included.

It is also important to note that in practice, for many applications we may also wish to include a set of common time dummies that are intended to capture disturbances which may be shared across the different members of the panel so that the remaining disturbances can be taken to be independent across individual members, as we discuss in the next section. A frequently used method for controlling for such common time effects in panels is to demean the data over the cross section dimension by subtracting off sample averages such as $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{i,t}$, taken over the N dimension. In contrast to the example of a time series mean, since the cross section mean is taken over the N dimension it is not affected by unit root asymptotics, and instead converges to the population mean as N grows large so that its estimation does not affect the asymptotic distribution. However, it should be pointed out that when regression coefficients are allowed to vary over individual members of the panel, as in (1) above, then demeaning the data over the cross section dimension in this fashion is no longer equivalent to estimating the time dummies directly in the panel regression. This can have the consequence of introducing data dependencies into the estimated residuals so that the asymptotic distributions are no longer nuisance parameter free. In small samples therefore, the more direct approach of estimating time dummies may be preferable in practice whenever it is computationally feasible. Finally, it should also be noted that in circumstances where the time series dimension substantially exceeds the cross section dimension, another possibility is to employ a GLS correction, as in Pedroni (1997b)

We are now ready to use the results of the multi-variate panel regression (1) to construct our tests. Here we have a choice of which type of statistic we wish to construct. Specifically, Pedroni (1997a) derives the asymptotic distributions and explores the small sample performances of seven different statistics. Of these seven statistics, four are based on pooling along what is commonly referred to as the within-dimension, and three are based on pooling along what is commonly referred to as the between-dimension. Specifically, the within-dimension statistics are constructed by summing both the numerator and the denominator terms over the N dimension separately, whereas the between-dimension statistics are constructed by first dividing the numerator by the denominator prior to summing over the N dimension. Thus, the former are based on estimators that effectively pool the autoregressive coefficient across different members for the unit root tests on the estimated residuals, while the latter are based on estimators that simply average the individually estimated coefficients for each member *i*. A consequence of this distinction arises in terms of the autoregressive coefficient, $\tilde{a}_i$, of the estimated residuals under the alternative hypothesis of cointegration. For the within-dimension statistics the test for the null of no cointegration is implemented as a residual based test of the null hypothesis $H_o : \tilde{a}_i = 1$ *for all i*, versus the alternative hypothesis $H_1 : \tilde{a}_i = \tilde{a} < 1$ *for all i*, so that it presumes a common value for $\tilde{a}_i = \tilde{a}$. By contrast, for the between-dimension statistics the null of no cointegration is implemented as a residual based test of the null hypothesis $H_o : \tilde{a}_i = 1$ *for all i*, versus the alternative hypothesis $H_1 : \tilde{a}_i < 1$ *for all i*, so that it does not presume a common value for $\tilde{a}_i = \tilde{a}$ under the alternative hypothesis. Thus, the between-dimension based statistics allow one to model an additional source of potential heterogeneity across individual members of the panel.

Following the terminology in Pedroni (1997a), we will refer to the within-dimension based statistics simply as panel cointegration statistics, and the between-dimension based statistics as group mean panel cointegration statistics. The first of the simple panel cointegration statistics is a type of nonparametric variance ratio statistic. The second is a panel version of a nonparametric statistic that is analogous to the familiar Phillips and Perron rho statistic.[2] The third statistic is also nonparametric and is

---

[2] The name “rho” derives from the fact that the statistic is based on an estimate of the autoregressive parameter, which is often denoted with the Greek letter ñ. In order to employ notation compatible with the conventions adopted throughout this journal issue, we have used the letter $\tilde{a}$ for the autoregressive parameter here. However, in keeping with earlier versions of this paper, and with the terminology used in Pedroni (1995, 1997a), we refer to the statistics in name as rho-statistics.

Table I. Panel Cointegration Statistics

1. Panel v-Statistic:

$$T^2 N^{3/2} Z_{\hat{\nu}_{N,T}} \equiv T^2 N^{3/2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{2} \right)^{-1}$$

2. Panel ρ-Statistic:

$$T\sqrt{N} Z_{\hat{\rho}_{N,T}-1} \equiv T\sqrt{N} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{2} \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \left(\hat{e}_{i,t-1} \Delta \hat{e}_{i,t} - \hat{\lambda}_i \right)$$

3. Panel t-Statistic:
(nonparametric)

$$Z_{t_{N,T}} \equiv \left(\tilde{\sigma}_{N,T}^{2} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{2} \right)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \left(\hat{e}_{i,t-1} \Delta \hat{e}_{i,t} - \hat{\lambda}_i \right)$$

4. Panel t-Statistic:
(parametric)

$$Z_{t_{N,T}}^{*} \equiv \left(\tilde{s}_{N,T}^{*2} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{*2} \right)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{*} \Delta \hat{e}_{i,t}^{*}$$

5. Group ρ-Statistic:

$$TN^{-1/2} \tilde{Z}_{\hat{\rho}_{N,T}-1} \equiv TN^{-1/2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1} \sum_{t=1}^{T} \left(\hat{e}_{i,t-1} \Delta \hat{e}_{i,t} - \hat{\lambda}_i \right)$$

6. Group t-Statistic:
(nonparametric)

$$N^{-1/2} \tilde{Z}_{t_{N,T}} \equiv N^{-1/2} \sum_{i=1}^{N} \left(\hat{\sigma}_i^{2} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{2} \right)^{-1/2} \sum_{t=1}^{T} \left(\hat{e}_{i,t-1} \Delta \hat{e}_{i,t} - \hat{\lambda}_i \right)$$

7. Group t-Statistic:
(parametric)

$$N^{-1/2} \tilde{Z}_{t_{N,T}}^{*} \equiv N^{-1/2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} \hat{s}_i^{*2} \hat{e}_{i,t-1}^{*2} \right)^{-1/2} \sum_{t=1}^{T} \hat{e}_{i,t-1}^{*} \Delta \hat{e}_{i,t}^{*}$$

where $\hat{\lambda}_i = \frac{1}{T} \sum_{s=1}^{k_i} \left(1 - \frac{s}{k_i+1} \right) \sum_{t=s+1}^{T} \hat{\mu}_{i,t} \hat{\mu}_{i,t-s}$, $\hat{s}_i^2 \equiv \frac{1}{T} \sum_{t=1}^{T} \hat{\mu}_{i,t}^2$, $\hat{\sigma}_i^2 = \hat{s}_i^2 + 2\hat{\lambda}_i$, $\tilde{\sigma}_{N,T}^2 \equiv \frac{1}{N} \sum_{i=1}^{N} \hat{L}_{11i}^{-2} \hat{\sigma}_i^2$

$\hat{s}_i^{*2} \equiv \frac{1}{T} \sum_{t=1}^{T} \hat{\mu}_{i,t}^{*2}$, $\tilde{s}_{N,T}^{*2} \equiv \frac{1}{N} \sum_{i=1}^{N} \hat{s}_i^{*2}$, $\hat{L}_{11i}^{2} = \frac{1}{T} \sum_{t=1}^{T} \hat{\eta}_{i,t}^2 + \frac{2}{T} \sum_{s=1}^{k_i} \left(1 - \frac{s}{k_i+1} \right) \sum_{t=s+1}^{T} \hat{\eta}_{i,t} \hat{\eta}_{i,t-s}$

and where the residuals $\hat{\mu}_{i,t}$, $\hat{\mu}_{i,t}^{*}$ *and* $\hat{\eta}_{i,t}$ *are obtained from the following regressions:*

$$\hat{e}_{i,t} = \hat{\gamma}_i \hat{e}_{i,t-1} + \hat{u}_{i,t} \,, \quad \hat{e}_{i,t} = \hat{\gamma}_i \hat{e}_{i,t-1} + \sum_{k=1}^{K_i} \hat{\gamma}_{i,k} \Delta \hat{e}_{i,t-k} + \hat{u}_{i,t}^{*} \,, \quad \Delta y_{i,t} = \sum_{m=1}^{M} \hat{b}_{mi} \Delta x_{mi,t} + \hat{\eta}_{i,t}$$

Notes: See text for further discussion of notation and procedures for implementation. All statistics are from Pedroni (1997a).

analogous to the Phillips and Perron t-statistic. Finally, the fourth of the simple panel cointegration statistics is a parametric statistic which is analogous to the familiar augmented Dickey-Fuller t-statistic. Just as in the conventional time series case, each of these statistics is shown to have a comparative advantage in terms of small sample size and power properties depending on the underlying data generating process, and Pedroni (1997a) presents extensive Monte Carlo simulation evidence for these patterns for the bivariate regression case. The other three panel cointegration statistics are based on a group mean approach. The first of these is analogous to the Phillips and Perron rho statistic, and the last two are analogous to the Phillips and Perron t-statistic and the augmented Dickey Fuller t-statistic respectively. Again, the comparative advantage of each of these statistics will depend on the underlying data generating process, and the reader is referred to Pedroni (1997a) for a detailed analysis based on the bivariate regression case.

Table I gives the precise form for each of these seven statistics as taken from Pedroni (1997a). In each case, the statistics can be constructed using the residuals of the cointegrating regression described in (1) above in combination with various nuisance parameter estimators which can be obtained from these. For many of the statistics, the nuisance parameter estimator $\hat{L}_{11i}^2$ is used. This nuisance parameter corresponds to the member specific long run conditional variance for the residuals. If we define $\Omega_i = \lim_{T\to\infty} E[T^{-1}(\sum_{t=1}^T \Delta z_{i,t})(\sum_{t=1}^T \Delta z_{i,t})']$ to be the long run covariance matrix for the partitioned vector of differenced unit root series $\Delta z_{i,t} = (\Delta y_{i,t}, \Delta x_{i,t}')'$, then based on Pedroni (1997a), we see that the multi-variate version of this estimator $\hat{L}_{11i}^2$ is given as $\hat{L}_{11i}^2 = \hat{\Omega}_{11i} - \hat{\Omega}_{21i}\hat{\Omega}_{22i}^{-1}\hat{\Omega}_{21i}'$, where $\hat{\Omega}_i$ is any consistent estimator of Ω_i, such as the Newey-West (1987) estimator. Notice that when $\Delta x_{i,t}$ is in an *M*-dimensional vector series, this implies that while $\hat{\Omega}_{11i}$ remains scalar, $\hat{\Omega}_{22i}$ is now an *MxM* dimensioned matrix, and $\hat{\Omega}_{21i}$ is a *1xM* dimensional vector. Obviously then , $\hat{L}_{11i}^2$ remains scalar regardless of the dimension of the vector $\Delta x_{i,t}$.

Indeed, closer inspection of the expression for $\hat{L}_{11i}^2$ reveals that it can also be interpreted as a conditional asymptotic variance based on the projection of $\Delta y_{i,t}$ onto $\Delta x_{i,t}$. Consequently, one relatively simple way to estimate $\hat{L}_{11i}^2$ is to regress $\Delta y_{i,t}$ onto the vector $\Delta x_{i,t}$ and then compute the asymptotic variance of the residuals of this regression, using for example the Newey-West estimator. The remaining nuisance parameter estimators presented in Table I are the same as in Pedroni (1997a) and are given by $\tilde{\sigma}^2_{N,T} \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{L}^{-2}_{11i}\hat{\sigma}^2_i$ and $\hat{\lambda}_i = \frac{1}{2}(\hat{\sigma}^2_i - \hat{s}^2_i)$, for which $\hat{s}^2_i$ and $\hat{\sigma}^2_i$ are the individual contemporaneous and long run variances respectively of the residuals $\hat{u}_{i,t}$ of the autoregression $\hat{u}_{i,t} = \hat{e}_{i,t} - \hat{\gamma}_i \hat{e}_{i,t-1}$. Finally, we note that the parametric versions of the t-statistics refer to the ADF based statistics, and so the estimator $\hat{s}^*_i$ refers to the standard contemporaneous variance of the residuals from the ADF regression, and $\tilde{s}^{*2}_{N,T} \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{s}^{*2}_i$ is simply the contemporaneous panel variance estimator.

In summary, we can compute any one of the desired statistics by performing the following steps:

1. Estimate the panel cointegration regression (1), making sure to include any desired intercepts, time trends, or common time dummies in the regression and collect the residuals $\hat{e}_{i,t}$ for later use.

2. Difference the original series for each member, and compute the residuals for the differenced regression $\Delta y_{i,t} = b_{1i}\Delta x_{1i,t} + b_{2i}\Delta x_{2i,t} + \ldots + b_{Mi}\Delta x_{Mi,t} + \eta_{i,t}$.

3. Calculate $\hat{L}^2_{11i}$ as the long run variance of $\hat{\eta}_{i,t}$ using any kernel estimator, such as the Newey-West (1987) estimator.

4. Using the residuals $\hat{e}_{i,t}$ of the original cointegrating regression, estimate the appropriate autoregression, choosing either of the following forms (a) or (b):

(a) For the nonparametric statistics (all except #4 and #7 of Table I) estimate $\hat{e}_{i,t} = \hat{\gamma}_i \hat{e}_{i,t-1} + \hat{u}_{i,t}$, and use the residuals to compute the long run variance of $\hat{u}_{i,t}$, denoted $\hat{\sigma}^2_i$. The term $\hat{\lambda}_i$ can then be computed as $\hat{\lambda}_i = \frac{1}{2}(\hat{\sigma}^2_i - \hat{s}^2_i)$ where $\hat{s}^2_i$ is just the simple variance of $\hat{u}_{i,t}$. Notice that these are the same as the usual correction terms that enter into the conventional single equation Phillips-Perron tests.

(b) For the parametric statistics (#4 and #7) estimate $\hat{e}_{i,t} = \hat{\gamma}_i \hat{e}_{i,t-1} + \sum_{k=1}^{K_i} \hat{\gamma}_{i,k}\Delta\hat{e}_{i,t-k} + \hat{u}^*_{i,t}$ and use the residuals to compute the simple variance of $\hat{u}^*_{i,t}$, denoted $\hat{s}^{*2}_i$.

5. Using each of these parts, construct any one of the statistics in Table I, and then apply the appropriate mean and variance adjustment terms discussed in section 2.2 and reported in Table II.

Notice furthermore, that for the group mean statistics (# 5 through #7 in Table I), step 2 and 3 are not required, since these statistics do not require an estimate of $\hat{L}_{11i}^{2}$. Furthermore, since the expressions inside the summation over the index *i* are equivalent to the expressions for the conventional single equation statistics, one can also compute these group mean statistics simply as the sum of the corresponding individual member ñ-statistics or t-statistics. Thus, statistic #4 can be viewed as most closely analogous to the Levin and Lin panel unit root statistic applied to the estimated residuals of a cointegrating regression and statistic #7 can be viewed as most closely analogous to the Im, Pesaran and Shin group mean unit root statistic applied to the estimated residuals of a cointegrating regression. Finally, we point out that truncation values for the number of lagged differences in the parametric ADF regression are permitted to vary by individual member, and can be determined individually for each member using standard step down procedures. Likewise, truncations for the kernel estimators which are used are also permitted to vary by individual member, and may also be determined individually by any preferred data determined selection scheme, such as for example the one suggested by Newey and West (1994) for their kernel estimator.

2.2 Critical values

In this section we discuss how the critical values for these tests are computed and present tables that can be used to obtain appropriate values. The asymptotic distributions and critical values that we present here are appropriate as approximations under the following assumptions regarding the data as applied to a panel regression of the form given in (1) above. First, we require that the standard conditions are present in the data for the individual members of the panel that would enable one to apply cointegration techniques in the conventional time series case. Formally, the requirements are simply that the multi-variate error process for the differenced data meet the criteria required for a functional central limit theorem to apply. These conditions represent relatively weak restrictions on the data, and are part of the reason that cointegration techniques have appealed to empirical researchers. These restrictions allow for a wide range of different forms of temporal dependence in the differenced data, and include for example the entire class of ARMA processes for Ä$z_{i,t}$. Furthermore, there is no requirement for these serial correlation properties to be the same for different members of the panel. This means that the differenced data, and thus also the residuals of an autoregression for the data, are permitted to have complex and differing dynamic properties among different members of the panel. We note also that there is no requirement for exogeneity of the regressors, as these dynamics are jointly determined for the entire vector of variables for any one member of the panel.

In addition to this assumption regarding the time series properties of the data, we also require an assumption regarding the cross member panel-wide properties of the data. Specifically, we require the common panel data assumption that, conditional on any shared aggregate disturbances that might be captured by the inclusion of common time dummies, the remaining idiosyncratic error terms are independent across individual members of the panel. Formally, we can summarize each of these assumptions in the following concise statement adapted from Pedroni (1995, 1997a) regarding the error process $\varepsilon_{i,t}$, where $z_{i,t} = z_{i,t-1} + \varepsilon_{i,t}$, so that $\Delta z_{i,t} = \varepsilon_{i,t}$, and which is taken to be conditional on any common aggregate disturbances.

Assumptions: *The process* $\hat{\varepsilon}_{i,t}^{\prime} \equiv (\hat{\varepsilon}_{i,t}^{y}, \hat{\varepsilon}_{i,t}^{x\,\prime})$ *satisfies* $\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \hat{\varepsilon}_{it} \Rightarrow B_i(\Omega_i)$, *for each member i as* $T \to \infty$, *where* $B_i(\Omega_i)$ *is vector Brownian motion with asymptotic covariance* Ω_i *with* $\Omega_{22i} > 0$, *such that the* $B_i(\Omega_i)$ *are taken to be defined on the same probability space for all i, and* $E[\hat{\varepsilon}_{i,t}\hat{\varepsilon}_{j,s}^{\prime}] = 0$ *for all* $i \neq j$ *for all s, t.*

Notice also, that we have restricted Ω_{22i} here to be positive definite, which guarantees that the vector of regressors are not mutually cointegrated so that we are considering a single cointegrating vector under the alternative hypothesis, although this cointegrating vector need not be the same among different members.

Under these assumptions, Pedroni (1995, 1997a) shows that following an appropriate standardization, each of the statistics described in Table I here will be distributed as standard normal when both the time series and cross sectional dimensions of the panel grow large. Furthermore, Pedroni (1995, 1997a) shows that the particular standardizations required to obtain asymptotic normality for each of the statistics depend only on the moments of certain underlying Brownian motion functionals. Specifically, these Brownian motion functionals correspond to the continuous time functionals to which the individual member statistics of the panel cointegration converge as T grows large prior to summing over the N dimension. When these are summed over the N dimension, the result is that by virtue of conditional independence across the i members, the appropriately standardized sums of Brownian motion functionals converge to normal distributions as N grows large.

Consequently is important to know the moments of the underlying Brownian motion functionals. In particular, Pedroni (1995, 1997a) shows that the asymptotic distributions for each of the 4 pooled panel cointegration statistics require standardizations based on the moments of the vector of Brownian motion functionals given by $Õ' \equiv (\int Q^2, \int QdQ, \tilde{â}^2)$ where $\tilde{â}$ is defined as $\tilde{â} \equiv \left(\int WW'\right)^{-1} \int WV$ and Q is defined as $Q \equiv V - \tilde{â}'W$, where V and W are mutually independent standard Brownian motion. With multiple regressors, we simply allow W to be an m-dimensional vector of standard Brownian motion, so that $\tilde{â}$ also becomes m dimensional and in this case we define $\tilde{â}^2 \equiv \tilde{â}'\tilde{â}$. Correspondingly, if we use $È$ and $ø$ to refer respectively to the vector of means and covariance matrix of these functionals, then Pedroni (1995, 1997a) shows that the asymptotic distributions for the pooled panel cointegration statistics will be as follows:

Panel v-Statistic: $T^2 N^{3/2} Z_{\hat{v}_{N,T}} - È_1^{-1}\sqrt{N} \;\; => \;\; N(0, ö'_{(1)} ø_{(1)} ö_{(1)})$

Panel ñ-Statistic: $T\sqrt{N} Z_{\hat{ñ}_{N,T}-1} - È_2 È_1^{-1}\sqrt{N} \;\; => \;\; N(0, ö'_{(2)} ø_{(2)} ö_{(2)})$

Panel t-Statistics: $Z_{t_{N,T}} - È_2 (È_1(1+È_3))^{-1/2}\sqrt{N} \;\; => \;\; N(0, ö'_{(3)} ø_{(3)} ö_{(3)})$

In this notation, $ø_{(j)}$, $j = 1,2,3$ refers to the j x j upper sub-matrix of the covariance matrix $ø$, and $ö_{(j)}$, $j = 1,2,3$ refers to the vectors $ö'_{(1)} = -È_1^{-2}$, $ö'_{(2)} = -(È_1^{-1}, È_2 È_1^{-2})$ and likewise $ö'_{(3)} = (È_1^{-1/2}(1+È_3)^{-1/2}, -\frac{1}{2}È_2 È_1^{-3/2}(1+È_3)^{-1/2}, -\frac{1}{2}È_2 È_1^{-1/2}(1+È_3)^{-3/2})$, which are composed of the elements of the mean vector $È$. Following Pedroni (1995, 1997a), the asymptotic distribution for the panel t-statistic is reported only once, since the parametric and nonparametric versions have the same asymptotic distribution.

Similarly, Pedroni (1997a) shows that the asymptotic distributions for each of the three group mean panel cointegration statistics require standardizations based on the moments of the vector of

Brownian motion functionals given by $\tilde{Õ}' \equiv \left((\int Q^2)^{-1} \int QdQ \, , \; ((1 - \tilde{â}^2) \int Q^2)^{-1/2} \int QdQ \right)$ where $\tilde{â}$ and Q are as previously defined and we have used a tilde here to distinguish $\tilde{Õ}$ from $Õ$. Similarly, use $\tilde{Ẽ}$, $\tilde{ø}$ with tildes here to refer respectively to the vector of means and covariance matrix of the vector functional. In this case, Pedroni (1997a) shows that the asymptotic distributions for the pooled panel cointegration statistics will be as follows:

Group ñ-Statistic: $$TN^{-1/2} \tilde{Z}_{ñ_{N,T}-1} - \tilde{Ẽ}_1 \sqrt{N} \;\; => \;\; N(0, \tilde{ø}_{1,1})$$

Group t-Statistics: $$N^{-1/2} \tilde{Z}_{t_{N,T}} - \tilde{Ẽ}_2 \sqrt{N} \;\; => \;\; N(0, \tilde{ø}_{2,2})$$

In this notation we refer directly to the elements of $\tilde{Ẽ}$ and $\tilde{ø}$, and again, following Pedroni (1997a), the asymptotic t-statistic is reported only once, since parametric and nonparametric versions have the same asymptotic distribution.

The result for both the pooled panel cointegration statistics and the group mean panel cointegration statistic indicates that each of the standardized statistics converges to a normal distribution whose moments depend on a total of 13 different elements of $È$, $ø$, $\tilde{Ẽ}$ and $\tilde{ø}$. The result is quite general and applies to whether or not the functionals are defined in terms of vectors or scalars, and whether the functionals are demeaned or detrended. In Pedroni (1995, 1997a) they are computed by Monte Carlo simulation for the case in which the panel of regressors $x_{i,t}$ is composed of a scalar time series, whereas the purpose of this study is to provide critical values for the case in which $x_{i,t}$ is composed of a vector of time series so that we can accommodate the case in which we have multiple regressors. Table III of the appendix reports simulated estimates of each of these 13 key moments for each case, depending on the number of regressors and whether or not constants and trends have been included. Specifically, these moments are obtained on the basis of 100,000 draws of *m* independent random walks of length T=1000.

Next, using these simulated moments, it is possible to construct approximations for the asymptotic distributions, and consequently to compute approximate critical values for each of the cases. Notice in particular that the asymptotic distributions for each of the seven panel and group mean statistics

Table II. Adjustment Terms for Panel Cointegration Tests

	Panel V		Panel Rho		Panel t		Group Rho		Group t	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Standard case.										
m=2	6.982	81.145	-6.388	64.288	-1.662	1.559	-9.889	41.943	-1.992	0.649
m=3	10.402	140.804	-10.191	89.962	-2.156	1.286	-13.865	57.801	-2.440	0.600
m=4	14.254	182.450	-14.136	103.176	-2.571	1.028	-17.834	72.097	-2.819	0.567
m=5	18.198	217.784	-18.042	120.787	-2.926	0.928	-21.805	88.611	-3.151	0.559
m=6	22.169	256.530	-21.985	132.499	-3.244	0.820	-25.750	103.371	-3.450	0.544
m=7	26.120	277.429	-25.889	143.561	-3.533	0.750	-29.627	117.059	-3.723	0.530
Heterogeneous intercepts included.										
m=2	11.754	104.546	-9.495	57.610	-2.177	0.964	-12.938	51.49	-2.453	0.618
m=3	15.197	151.094	-13.256	81.772	-2.576	0.923	-16.888	67.123	-2.827	0.585
m=4	18.910	190.661	-17.163	99.331	-2.930	0.843	-20.841	81.835	-3.157	0.560
m=5	22.715	231.864	-21.013	119.546	-3.241	0.800	-24.775	98.278	-3.452	0.553
m=6	26.603	270.451	-24.944	134.341	-3.531	0.750	-28.720	113.131	-3.726	0.542
m=7	30.457	293.431	-28.795	144.615	-3.795	0.685	-32.538	126.059	-3.976	0.525
Heterogeneous deterministic trends and intercepts included.										
m=2	21.162	160.249	-14.011	64.219	-2.648	0.690	-17.359	66.387	-2.872	0.555
m=3	24.556	198.167	-17.600	83.815	-2.967	0.686	-21.116	81.832	-3.179	0.548
m=4	28.046	239.425	-21.287	103.905	-3.262	0.688	-24.930	97.362	-3.464	0.543
m=5	31.738	276.997	-25.130	124.613	-3.545	0.686	-28.849	113.145	-3.737	0.538
m=6	35.537	310.982	-28.981	138.227	-3.806	0.654	-32.716	127.989	-3.986	0.530
m=7	39.231	348.217	-32.756	154.378	-4.047	0.638	-36.494	140.756	-4.217	0.518

Notes: *The value m refers to the number of regressors, excluding any constant or deterministic trend terms. See text for definitions of the statistics and adjustment terms. All adjustment terms are computed on the basis of the simulated moments from Table III.*

can be expressed in the form

$$\frac{\varkappa_{N,T} - \mu\sqrt{N}}{\sqrt{\nu}} \quad => \quad N(0,1) \tag{2}$$

where $\varkappa_{N,T}$ is the appropriately standardized (with respect to the dimensions N and T) form for each of the statistics as described in Table I, and the values for μ and v are functions of the moments of the underlying Brownian motion functionals. Consequently, in order to compute appropriate critical values for each of the tests, we simply use the moments presented in Table III of the appendix to construct the corresponding values for µ and í for each case, depending on the number of regressors and whether or not we have included constants or trends in the regression. These values for µ and í have been tabulated and are reported in Table II and are referred to as the mean and variance adjustment terms respectively.

Thus, to test the null of no cointegration, one simply computes the value of the statistic so that it is in the form of (2) above based on the values of µ and í from Table II and compares these to the appropriate tails of the normal distribution. Under the alternative hypothesis, the panel variance statistic diverges to positive infinity, and consequently the right tail of the normal distribution is used to reject the null hypothesis. Consequently, for the panel variance statistic, large positive values imply that the null of no cointegration is rejected. For each of the other six test statistics, these diverge to negative infinity under the alternative hypothesis, and consequently the left tail of the normal distribution is used to reject the null hypothesis. Thus, for any of these latter tests, large negative values imply that the null of no cointegration is rejected.

3. Concluding Remarks

The purpose of this paper has been to compute and report critical values appropriate for the application of the panel cointegration statistics first developed in Pedroni (1995, 1997a) to test for the null of no cointegration in the case in which multiple regressors are included. We have assumed throughout that the researcher has in mind a particular normalization among variables which is deemed sensible and is simply interested in knowing whether or not the variables are cointegrated. In this context, the procedures investigated in this paper can be thought of as extensions of the traditional Engle and Granger (1987) two step residual based methods for the tests of the null of no cointegration, as applied to heterogeneous panels. Specifically, the statistics as applied to panels have been constructed in such a manner as to permit as much member specific heterogeneity as possible, not only in terms of member specific fixed effects and time trends, but also in terms of member specific cointegrating vectors and member specific dynamics. In this way, the results of this paper further compliment the initial work that has been done on raw panel unit root tests. An interesting and promising area for further research will be to examine the extent to which panel methods in the spirit of this line of research can also be used to improve ones ability to make inferences about the specific rank or number of cointegrating relationships in panels with multiple regressors in the context of maximum likelihood based methods.

REFERENCES

Bernard, A. and C. Jones (1996) "Productivity Across Industries and Countries; Time Series Theory and Evidence," *Review of Economic Statistics*, 78, 135-146.

Butler, A. and M. Dueker (1999) “Does Foreign Innovation Affect Domestic Wage Inequality?” *Journal of International Economics*, 47, 61-89.

Canzoneri M., R. Cumby and B. Diba (1996) "Relative Labor Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel of OECD Countries," NBER Working Paper No. 5676.

Chinn, M. (1997) "Sectoral productivity, government spending and real exchange rates : empirical evidence for OECD countries," NBER Working Paper No. 6017.

Chinn, M. and L. Johnston (1996) "Real Exchange Rate Levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries," NBER Working Paper No. 5709.

Coakley, J. and A. Fuertes (1997) “New Panel Unit Root Tests of PPP,” *Economics Letters*, 57, 17-22.

Engle R. and C. Granger (1987) “Co-integration and Error Correction: Representation, Estimation, and Testing,” *Econometrica*, 55, 251-76.

Evans, P. and G. Karras (1996) "Convergence Revisited," *Journal of Monetary Economics*, 37, 249-65.

Frankel, J. and A. Rose (1996) “A panel project on purchasing power parity: Mean reversion within and between countries, *Journal of international economics*, 40, 209-224.

Im, K., H. Pesaran and Y. Shin (1997) "Testing for Unit Roots in Heterogeneous Panels, Discussion Paper, University of Cambridge, June.

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

Kao, C., Chiang, M. and B. Chen (1999) “International R&D Spillovers: An Application of Estimation and Inference in Panel Cointegration,” forthcoming in this issue.

Lee, K., H. Pesaran and R. Smith (1997) "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Growth Model," *Journal of Applied Econometrics*, 12, 357-92.

Levin, A. and C. Lin (1993) "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," University of California, San Diego Discussion Paper, December.

MacDonald, R. (1996) “Panel Unit Root Tests and Real Exchange Rates,” *Economics Letters*, 50, 7-11.

Maddala, G.S. and S. Wu (1999) “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test,” forthcoming in this issue.

McCoskey, S. and Kao, C. (1999) “Testing the Stability of a Production Function with Urbanization as a Shift Factor: An Application of Non-Stationary Panel Data Techniques,” forthcoming in this issue.

Neusser, K. and M. Kugler (1998) “Manufacturing Growth and Financial Development: Evidence from OECD Countries,” *Review of Economics and Statistics*, 80, 638-46.

Newey, W. and K. West (1987) "A Simple Positive Semi-Definite, Heteroscadasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 50, 703-8.

___________________ (1994) “Autocovariance Lag Selection in Covariance Matrix Estimation,” *Review of Economic Studies*, 61, 631-653.

Obstfeld M. and A. Taylor (1996) "International Capital-Market Integration over the Long Run: The Great Depression as a Watershed," manuscript, U.C. Berkeley.

O’Connell, P. (1998) “The Overvaluation of Purchasing Power Parity,” *Journal of International Economics*, 44, 1-19.

Oh, K. (1994) "Purchasing Power Parity and Unit Root Tests Using Panel Data," *Journal of International Money and Finance*, 15, 405-418.

Ong, S. E. and C. Maxam (1997) "A Heterogenous Panel Cointegration- Error Correction Approach to Modeling Commerical Mortgage Backed Security Prices," *Journal of Property Finance*, 8, 317-335.

Papell, D. (1997) "Searching for Stationarity: Purchasing Power Parity Under the Current Float," *Journal of International Economics*, 43, 313-32.

Pedroni, P. (1993) “Panel Cointegration,” Chapter 2 in PANEL COINTEGRATION, ENDOGENOUS GROWTH AND BUSINESS CYCLES IN OPEN ECONOMIES, Columbia University Dissertation, UMI Publishers, Ann Arbor, Michigan.

________ (1995) “Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, With an Application to the PPP Hypothesis,” Indiana University Working Papers In Economics, No. 95-013, June.

________ (1996a) “Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity,” Indiana University Working Papers In Economics, No. 96-020, June.

________ (1996b) “Endogenous Growth, Human Capital, and Cointegration for Multi-Country Panels,” working paper, Indiana University, November.

________ (1997a) “Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, With an Application to the PPP Hypothesis: New Results,” working paper, Indiana University, April.

________ (1997b) "On the Role of Cross Sectional Dependency in Dynamic Panel Unit Root and Panel Cointegration Exchange Rate Studies," working paper, Indiana University, November.

________. (1998) "Testing for Convergence to Common Steady States in Nonstationary Heterogeneous Panels," working paper, Indiana University, November.

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

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

Quah, D. (1994) "Exploiting Cross-Section Variation for Unit Root Inference in Dynamic Data," *Economics Letters*, 44, 9-19.

Taylor, A. (1996). "International Capital Mobility in History: Purchasing Power Parity in the Long-Run." NBER Working Paper No. 5742.

Wei, S. and D. Parsley (1995) "Purchasing Power Dis-Parity During the Floating Rate Period: Exchange Rate Volatility, Trade Barriers and Other Culprits," NBER working paper No. 5032.

Wu, Y. (1996) "Are Real Exchange Rates Nonstationary? Evidence from a Panel-Data Test," *Journal of Money Credit and Banking*, 28, 54-63.

TABLE III. SIMULATED MOMENTS

	è$_1$	è$_2$	è$_3$	ø$_{11}$	ø$_{22}$	ø$_{33}$	ø$_{12}$	ø$_{13}$	ø$_{23}$	$\tilde{\text{è}}_1$	$\tilde{\text{è}}_2$	$\tilde{\text{ø}}_1$	$\tilde{\text{ø}}_2$
Standard													
m=2	0.143	-0.915	2.116	0.034	0.782	3.569	-0.067	0.205	-1.572	-9.889	-1.992	41.943	0.649
m=3	0.096	-0.980	2.148	0.012	0.669	3.000	-0.053	0.130	-1.378	-13.865	-2.440	57.801	0.600
m=4	0.070	-0.992	2.120	0.004	0.476	2.093	-0.030	0.069	-0.980	-17.834	-2.819	72.097	0.567
m=5	0.055	-0.991	2.090	0.002	0.350	1.505	-0.018	0.039	-0.716	-21.805	-3.151	88.611	0.559
m=6	0.045	-0.992	2.071	0.001	0.284	1.214	-0.012	0.026	-0.581	-25.750	-3.450	103.371	0.544
m=7	0.038	-0.991	2.055	0.001	0.214	0.902	-0.008	0.012	-0.436	-29.627	-3.723	117.059	0.530
Demeaned													
m=2	0.085	-0.808	1.618	0.005	0.218	0.722	-0.016	0.031	-0.361	-12.938	-2.453	51.490	0.618
m=3	0.066	-0.872	1.743	0.003	0.219	0.779	-0.014	0.027	-0.391	-16.888	-2.827	67.123	0.585
m=4	0.053	-0.908	1.815	0.001	0.209	0.784	-0.011	0.022	-0.390	-20.841	-3.157	81.835	0.560
m=5	0.044	-0.925	1.850	0.001	0.182	0.687	-0.008	0.016	-0.344	-24.775	-3.452	98.278	0.553
m=6	0.038	-0.938	1.876	0.001	0.166	0.633	-0.006	0.013	-0.317	-28.720	-3.726	113.131	0.542
m=7	0.033	-0.945	1.890	0.0003	0.141	0.540	-0.005	0.009	-0.271	-32.538	-3.976	126.059	0.525
Demeaned and Detrended													
m=2	0.047	-0.662	1.323	0.001	0.048	0.154	-0.002	0.004	-0.077	-17.359	-2.872	66.387	0.555
m=3	0.041	-0.717	1.433	0.001	0.057	0.197	-0.002	0.005	-0.099	-21.116	-3.179	81.832	0.548
m=4	0.036	-0.759	1.518	0.0004	0.063	0.228	-0.002	0.005	-0.114	-24.930	-3.464	97.362	0.543
m=5	0.032	-0.792	1.583	0.0003	0.065	0.241	-0.002	0.005	-0.121	-28.850	-3.737	113.145	0.538
m=6	0.028	-0.816	1.632	0.0002	0.065	0.247	-0.002	0.004	-0.124	-32.716	-3.986	127.989	0.530
m=7	0.025	-0.835	1.670	0.0001	0.065	0.250	-0.002	0.004	-0.125	-36.494	-4.217	140.756	0.518

Notes: *Simulations are based on 100,000 draws of m independent random walks of length T=1000. See text for definitions of each of the moments.*