June, 1995

PANEL COINTEGRATION;

ASYMPTOTIC AND FINITE SAMPLE PROPERTIES OF POOLED TIME SERIES TESTS WITH AN APPLICATION TO THE PPP HYPOTHESIS

Peter Pedroni

Indiana University

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

email: PPEDRONI@INDIANA.EDU

*This is a substantially revised version of the draft that was presented at the North American and European Econometric Society Meetings in Quebec and Maastricht respectively during the summer of 1994, whose participants I thank for their many helpful comments. I also thank especially Richard Clarida, Heejoon Kang, Andrew Levin, Masao Ogaki, David Papell and Pierre Perron for helpful suggestions, and Rasmus Ruffer and Younghan Kim for valuable research assistance.

Abstract

This paper studies asymptotic and finite sample properties of statistics devised to test for the null of no cointegration in nonstationary pooled time series panels as both the cross section and time series dimensions grow large. The paper finds that for panels with homogenous long run parameters, the spurious regression coefficient estimates become consistent even under the null of no cointegration, and this generates a superconsistency result for panels whereby it becomes irrelevant for the asymptotic distributions whether the residuals are known or are estimated. For heterogenous panels, on the other hand, this asymptotic equivalency does not hold, and the use of stationarity tests that are convergent for raw data can even become divergent when applied to estimated residuals. Furthermore, the direct use of unit root tests for estimated residuals can generate data dependencies that are not present in unit root statistics that are applied to raw panel data. The paper therefore derives asymptotic distributions for panel cointegration statistics that circumvent these problems, and also reports on finite sample properties of these statistics based on Monte Carlo simulations for a varying number of time periods and cross sections. These statistics allow for complete heterogeneity of the dynamics and potential cointegrating relationships across members of the panel. Finally, an empirical application of these panel cointegration statistics is demonstrated for the purchasing power parity hypothesis, which has been difficult to evidence on the basis of conventional single series cointegration tests.

PANEL COINTEGRATION; ASYMPTOTIC AND FINITE SAMPLE PROPERTIES OF POOLED TIME SERIES TESTS WITH AN APPLICATION TO THE PPP HYPOTHESIS

I. Introduction

The use of cointegration techniques to test for the presence of long run relationships among integrated variables have enjoyed growing popularity in the empirical literature. Unfortunately, a common dilemma for practitioners has been the inherently low power of many of these tests when applied to time series available for the length of the post war period. Research by Shiller and Perron (1985), Perron (1989,1991) and recently Pierse and Snell (1995) have generally confirmed that it is the span of the data, rather than frequency that matters for the power of these tests. On the other hand, expanding the time horizon to include pre-war data can risk introducing unwanted changes in regime for the data relationships. In light of these data limitations, it is natural to question whether a practical alternative might not be to bring additional data to bear upon a particular cointegration hypothesis by drawing upon data from among similar cross sectional data in lieu of additional time periods.

For many important hypotheses to which cointegration methods have been applied, data is in fact commonly available on a time series basis for multiple countries, for example, and practitioners could stand to benefit significantly if there existed a straightforward manner in which to perform cointegration tests for pooled time series panels. Many areas of research come to mind, such as the growth and convergence literature, or the purchasing power parity literature, for which it is natural to think about long run time series properties of data that are expected to hold for groups of countries. Alternatively, examples are also readily available for issues that involve time series panels for yields across asset term structures or price movements across industries, to name only a few. For applications where the cross sectional dimension grows reasonably large, existing systems methods such as the Johansen (1989, 1991) procedure are likely to become infeasible, and panel methods may be more appropriate.

On the other hand, pooling time series has traditionally involved a substantial degree of sacrifice in terms of the permissible heterogeneity of the individual time series.[1] In order to ensure broad applicability of any panel cointegration test, it will be important to allow for as much heterogeneity as possible among the individual members of the panel. Therefore, one objective of this paper will be to construct panel cointegration test statistics that allow one to vary the degree of permissible heterogeneity among the members of the panel, and in the extreme case pool only the multivariate unit root information, leaving the form of the time series dynamics and the potential cointegrating vectors entirely heterogeneous across individual members.

Unfortunately, very little is known about the properties of cointegration tests for these types of applications in which both the time series and cross sectional dimensions grow large. Of course recent work by Levin and Lin (1994) and Quah (1994) has gone a long way toward furthering the understanding of asymptotics for nonstationary panels. Quah (1994) derives asymptotically normal distributions for standard unit root tests in panels for which the time series and cross sectional dimensions grow large at the same rate. Levin and Lin (1994) extend this work for the case in which both dimensions grow large independently and derive asymptotic distributions for panel unit root tests that allow for heterogeneous intercepts and trends across individual members. The authors also document large gains in power even for relatively small sized panels as compared to conventional single series tests.[2] Based on the relationship between cointegration tests and unit root tests in the conventional single series case, one might be tempted to think that the panel unit root statistics introduced in these studies might be directly applicable to tests of the null of no cointegration, with perhaps some changes in the critical values to reflect the use of estimated residuals.

Quite to the contrary, however, the relationship for nonstationary panels turns out to be considerably more involved. The asymptotic methods used for nonstationary panels dictate that in

[1] See for example Holz-Eakon, Newey and Rosen (1988) on the dynamic homogeneity restrictions required typically for the implementation of panel VAR techniques.

[2] For example, Levin and Lin show by Monte Carlo simulation that at a nominal size of 10%, with 100 time periods the power of pooled unit root tests jump to nearly 100% for 25 cross sections against stationarity alternates with heterogenous means and heterogenous means and trends as compared to 31% and 37% respectively for the conventional single series tests.

general the properties of unit root tests applied to estimated residuals can be considerably different than the case in which these tests are applied to raw data. In particular, the multivariate nature of cointegration introduces two key complicating features. The first involves the fact that regressors are typically not required to be exogenous for cointegrated systems, which introduces off diagonal terms in the residual asymptotic covariance matrix. As Phillips and Ouliaris (1990) demonstrate, for the single series case these terms drop out of the asymptotic distributions for unit root tests based on residuals. By contrast, for nonstationary panels, these effects are likely to be idiosyncratic across individual members and will not generally vanish from the asymptotic distribution for unit root tests. If these features of the data are ignored, the asymptotic method used to average over the cross sectional dimension of the data for raw unit root tests can introduce undesirable data dependencies into the asymptotic distributions when estimated residuals are used.

The second, more general complicating feature that estimated residuals bring to unit root tests for panels involves the dependency of the residuals on the distributional properties of the estimated coefficients of the spurious regression, which itself converges to a nonstandard random variable. For the single series case, this has the effect of altering the asymptotic distributions and the corresponding critical values required to reject the null of no cointegration as compared to the null of a raw unit root. For panels, because of the averaging that occurs in the cross sectional dimension, this can induce much more dramatic effects on the properties of the asymptotic distributions. Interestingly, the nature of the cross sectional asymptotics is such that the effect of this dependency on the distribution of the estimated coefficients will hinge critically on the precise nature of the alternate hypothesis that is being entertained for the cointegrating relationship. In fact, if the alternate hypothesis is such that any cointegrating relationship is necessarily similar among members of the panel, so that the coefficients of the spurious regression can be constrained to be homogenous, then the slope estimator becomes consistent at rate $\sqrt{N}$ even under the null of no cointegration as the cross sectional dimension, N, grows large. This has the interesting consequence of producing for this special case a type of superconsistency result for nonstationary panels such that the asymptotic distribution of the unit root tests will actually be invariant to whether the residuals are known or are estimated.

More generally however, when the form of the cointegrating relationship under the alternate hypothesis is not restricted to be homogeneous across individual members of the panel, then the effect of the dependency of the estimated residuals on the estimated coefficients of the spurious regression can be of a more sinister nature. In this case, the random variable nature of the estimated coefficients can have the effect of transforming a convergent panel unit root test statistic into a nonconvergent test statistic when it is applied to estimated residuals. The practical implications can be staggering. Consider a simple example. The critical value required to reject a unit root at the 10% level for a panel of 50 cross sections with zero mean and trend is -1.81 for the panel OLS autoregression rho-statistic and -1.28 for the corresponding t-statistic. By contrast, according to the asymptotic distributions presented in this paper, the appropriate critical value for the same case applied to estimated residuals would become -25.98 and -8.71 respectively, and a researcher who mistakenly reports significant values under the assumption of raw unit root tests would in actuality be reporting values that are in fact far to the right hand side of the true distributions. Thus, the error that is made in using critical values from raw unit root distributions for estimated residuals becomes much more severe for panels, and becomes worse as the cross sectional dimension grows large.

Notice, furthermore, that the more general heterogenous case for the alternate is by far the more important distribution to consider. This is because erring on the side of homogeneity will do more than just lead to a wildly inaccurate asymptotic distribution, since falsely imposing a homogeneous slope coefficient on the individual members of the panel will generate a component of the residual which is integrated even under the alternate of cointegration. Therefore, unless one has very strong convictions that any cointegrating vector under the alternate must necessarily be homogeneous, one should always use the unrestricted specification and its corresponding distribution. Thus, again, an important objective of this paper is to study the properties and derive asymptotic distributions for the case in which even the parameters of the long run cointegrating relationships are permitted to be heterogenous among cross sections.

In particular, the paper proposes three sets of statistics designed to test for the null of no cointegration for heterogenous panels and derives their asymptotic distributions as both the time series and cross sectional dimensions grow large. The statistics have their single series analogs in the autoregressive rho-statistic, the corresponding t-statistic, and a variance ratio statistic. Each are shown to be constructed in a manner such that their asymptotic distributions are free of nuisance parameters associated with any idiosyncratic temporal dependencies that may be present in the data. In particular, the distributions of all three panel cointegration statistics are shown to be asymptotically normal and to depend only on the moments of a vector Brownian motion functional $\Upsilon' \equiv (\int Q^2, \int QdQ, \tilde{\beta}^2)$, where $Q \equiv V - \tilde{\beta}W$ for independent standard Wiener processes W(r) and V(r) and $\tilde{\beta} \equiv \frac{\int VW}{\int W^2}$ is the distribution of the spurious regression coefficient for independent random walks. In this way, the distributions are specified in a form that depends only on the properties of standard Brownian motion despite the very heterogenous nature of the individual members of the panel. Section II presents these asymptotic distributions in addition to the other asymptotic results discussed in this introduction. Finite sample distributions are also investigated in Monte Carlo simulations for various combinations of time series and cross sectional dimensions as reported in the appendix. The proofs for each of the results in section II are collected in the mathematical appendix.

Finally, section III demonstrates a brief empirical application of these panel cointegration statistics to the hypothesis of exchange rate purchasing power parity. The conventional single series tests have been hard pressed to find evidence in support of the PPP hypothesis on the basis of country by country tests on data from the post Bretton-Woods floating exchange rate period from 1973 to the present. Since the null of no PPP is tested via the null of no cointegration, a widely held belief is that this result is due to the inherently low power of these tests for such a short time span. Point estimates for the long run relationships among nominal variables indicate considerable heterogeneity among a panel of 20 to 25 countries for the period. Consequently, it becomes essential to use panel cointegration tests that allow for heterogeneity of these relationships to avoid potential mispecification rather than unit root tests that necessarily impose homogeneity of the cointegrating relationship. On this bases, panel cointegration tests for both the annual and monthly data appear to clearly reject the null of no cointegration in favor of PPP, thereby overturning earlier conventional single series results.

II. Asymptotic Properties of Panel Regressions for Integrated Regressors

2.1 The Panel Models and Basic Methodology

We will in general refer to two sets of panel regression models. The first we will refer to as homogenous, reflecting the notion that the individual cross sections of the panels are assumed to share common long run features under the alternate hypothesis of cointegration, including common cointegrating vector coefficients, means, or deterministic time trends when applicable. Thus, we will write these type of regressions as

Model 1 *(homogeneous panel):*

$$y_{it} = X_{it}\beta + e_{it} \tag{1}$$

$$y_{it} = \alpha + X_{it}\beta + e_{it} \tag{1'}$$

$$y_{it} = \alpha + \delta t + X_{it}\beta + e_{it} \tag{1''}$$

for which the variables y and X are subscripted by the index i to indicate that they exist for each of $i = 1, \dots , N$ different cross sections. The parameters α, β, δ on the other hand, are not subscripted by the index *i* to indicate that they will be assumed to be homogenous across individual cross sections of the panel under the alternate hypothesis that they reflect the coefficients of a cointegrating relationship. By comparison, we will refer to the second set of models as heterogenous, reflecting the fact that the individual cross sections for these panels are not necessarily assumed to be homogenous under the alternate hypothesis. Thus, we will write these type of regressions as

Model 2 *(heterogenous panel):*

$$y_{it} = X_{it}\beta_i + e_{it} \tag{2}$$

$$y_{it} = \alpha_i + X_{it}\beta_i + e_{it} \tag{2'}$$

$$y_{it} = \alpha_i + \delta_i t + X_{it}\beta_i + e_{it} \tag{2''}$$

for which the coefficients α,β,δ are now subscripted with the cross sectional index *i*. The distinction between these two types of models is critical in terms of the type of alternate hypotheses that they will permit under cointegration testing. For the first set, the alternate to no cointegration must be that if the individual are cointegrated, then they will exhibit the *same* long run cointegrating relationships. For the second, the alternate to no cointegration may be that the individual cross sections contain cointegrating relationships that are free to take on different values for different members of the panel.

Of course models need not be restricted to exclusively homogenous or heterogeneous long run parameters. In some cases one may also wish to consider hybrid models which retain some features of both model types (1) and (2). For example, the model given by

Model 3 *(hybrid, with common time dummies):*

$$y_{it} = \delta_t + \alpha_i + X_{it}\beta_i + e_{it} \tag{3}$$

can be useful if one wishes to permit common aggregate shocks that are shared across individuals while still allowing for both dynamic and long run idiosyncratic responses.

In each model, unless otherwise indicated, all other assumptions regarding the properties of the stochastic processes will be the same. Specifically, let Z_i be defined by the vector $Z_i \equiv (y_i, X_i)'$. Then we will assume that each individual Z_i is generated by the process

$Z_{it} = Z_{it-1} + \xi_{it}$,[3] for which the following standard assumption holds for each individual cross section:

Assumption 1.1 (invariance principle): *The process* $\xi_{it} \equiv (\xi_{it}^{y}, \xi_{it}^{x})^{\prime}$ *satisfies* $\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} \xi_{it} \rightarrow B_i(\Omega_i)$, *as* $T \rightarrow \infty$, *where -> signifies weak convergence, and* $B_i(\Omega_i)$ *is vector Brownian motion defined over the real interval* $r \in [0,1]$, *with asymptotic covariance* Ω_i.

Thus, assumption 1.1 is simply a statement that the standard functional central limit theorem is assumed to hold individually for each cross sectional series as T grows large. The conditions on the error process required for this convergence are relatively weak and includes the entire class of stationary ARMA processes.[4] The asymptotic covariance matrix Ω_i is given by

$$\Omega_i \equiv \lim_{T \rightarrow \infty} E\left(T^{-1}(\sum_{t=1}^{T} \xi_{it})(\sum_{t=1}^{T} \xi_{it}^{\prime}) \right) \qquad \textbf{(4)}$$

which can be decomposed as $\Omega_i \equiv \Omega_i^{o} + \Gamma_i + \Gamma_i^{\prime}$ where Ω_i^{o} is the contemporaneous covariance among the components of ξ_{it} for a given cross section *i*, and Γ_i captures the dynamic covariances among the components of ξ_{it} for a given cross section *i*. The off diagonal terms of the asymptotic covariance matrix Ω_i, denoted ω_{21i}, $\omega_{21i}^{\prime}$, thus capture the idiosyncratic endogenous feedback among the I(1) variables, and in keeping with the cointegration literature we do not require that the regressors X_{it} be exogenous. The fact that Ω_i is permitted to vary across individual sections of the panel reflects the fact that in general we will permit all dynamics that are absorbed in the asymptotic covariance matrix to be heterogeneous for any of the models (1) through (3), regardless of whether or not the long run parameters α, δ, β are treated as varying across individual sections.

[3] We will also assume for convenience that initial conditions are given by Z_{io} constant to avoid complications from possible covariation of the initial condition with subsequent errors, which is considered in Quah (1994).

[4] See standard references, eg. Phillips (1986,1987), Phillips and Durlauf (1986), for further discussion of the conditions under which assumption 1.1 holds more generally.

A number of consistent estimators are available for the individual Ω_i, typically based on kernel estimators for the Γ_i component. Typical estimators take the form

$$\hat{\Omega}_i \equiv \hat{\Omega}_i^o + \hat{\Gamma}_i + \hat{\Gamma}_i^{\prime}$$

$$where \quad \hat{\Omega}_i^o \equiv T^{-1} \sum_{t=1}^{T} \hat{\xi}_{it} \hat{\xi}_{it}^{\prime} \,, \quad \hat{\Gamma}_i \equiv T^{-1} \sum_{s=1}^{k_i} w_{sk_i} \sum_{t=1}^{T} \hat{\xi}_{it} \hat{\xi}_{it-s}^{\prime} \qquad \textbf{(5)}$$

for some lag window w_{sk}, where $\hat{\xi}_{it}$ is obtained from an autoregression $Z_{it} = \rho_i Z_{it-1} + \xi_{it}$ individually for each i. A commonly used lag window is based on the Newey-West estimator with $w_{sk_i} = 1 - \frac{s}{k_i + 1}$. More recently, Andrews and Monahan (1992), Lee and Phillips (1994) and Levin and den Haan (1995) have proposed robust asymptotic covariance estimators based on prewhitening of the residuals $\hat{\xi}_{it}$ to improve finite sample estimates of Ω_i. Based on the preliminary investigations of Park and Ogaki (1991) for these type of estimators in the context of standard cointegrating regressions, the finite sample improvements from these prewhitening procedure may be particular attractive in the present context of panel cointegration tests with relatively small time series dimensions.

In addition to the conditions for the invariance principle with regard to the time series dimension, we will also assume the following condition in keeping with a panel data approach

Assumption 1.2 (cross sectional independence): *The individual processes are assumed to be independent cross sectionally, so that* $E[\xi_{it}, \xi_{jt}] = 0$ *forall* $i \neq j$. *More generally, the asymptotic covariance matrix for a panel of size NxT is given as* $I_N \otimes \Omega_i > 0$, *which is block diagonal positive semidefinite with the ith diagonal block given by the asymptotic covariances for cross section i.*

Formally, this condition will be required to apply standard central limit theorems in the cross sectional dimension in the presence of heterogenous errors. In practice, the assumption is not as restrictive as it may first appear given the possibility of incorporating aspects of the hybrid models of type (3) to capture any disturbances that are common across members of the panel as is frequently done for panels. On the other hand, the condition that the covariance matrix for the panel as a whole be positive semidefinite rules out any singularities that would reflect potential intra-panel stochastic cointegrating relationships.[5]

Together, conditions 1.1 and 1.2 regarding the time series and cross sectional properties of the error processes for the time series panel will provide us with the basic methodology for investigating the asymptotic properties of various statistics as both N and T grow large. Thus, the first assumption will allow us to make use of standard convergency results regarding asymptotics in the time series dimension for each of the individual cross sections. In particular, we will make use of the fact that the following convergencies, developed in Phillips and Durlauf (1986) and Park and Phillips (1988), must also hold for each of the individual cross sections *i* as T grows large, so that

$$T^{-2} \sum_{t=1}^{T} Z_{it-1} Z_{it-1}' \quad \rightarrow \quad \Omega_i^{1/2} \int_0^1 Z_i(r) Z_i(r)' dr \, \Omega_i^{1/2} \tag{6a}$$

$$T^{-1} \sum_{t=1}^{T} Z_{it-1} \xi_{it}' \quad \rightarrow \quad \Omega_i^{1/2} \int_0^1 Z_i(r) \, dZ_i(r)' dr \, \Omega_i^{1/2} \; + \; \Gamma_i \tag{6b}$$

where $Z_i(r) \equiv (V_i(r), W_i(r))'$ is a vector Brownian motion such that V(r) and W(r) are independent standard Wiener processes, and Ω_i and its components are defined as in (4) above. The decomposition that is implicit in these convergencies, in terms of the transformation of discrete statistics that are heterogenous across *i* to continuous statistics that are expressed as a product of standard Brownian motion and the idiosyncratic asymptotic covariance terms, is key in permitting the use of panel data methods effectively to variables with such general and heterogenous error processes. In particular, the fact that the elements of $Z_i(r)$ are standard implies that the distributions of functionals of $Z_i(r)$ will be identical across individual *i* so that more standard central limit theorems can be applied to sums of these standardized Brownian

[5] In contrast to the common time dummies, the presence of stochastic cointegrating relationship among cross sections will generally impact the limiting distributions if not properly accommodated.

motion functionals as N grows large. In this case the resulting distributions will be asymptotically normal and free of nuisance parameters with moments determined solely by the properties of the Brownian motion functionals, as we will see.

2.2 Asymptotic Properties of Spurious Regressions in Panels

In order to facilitate analysis of the asymptotic properties of residual stationarity based tests, we first investigate a number of interesting properties of panel regressions based on models (1) and (2) for the case in which the residual process e_{it} is integrated of order one so that the regressions can be characterized as "spurious" in the long run for each individual cross section. Define $\hat{\beta}_{NT}$ as the pooled panel OLS estimator of the spurious regression slope coefficient β for models of type (1), and let $\hat{\beta}_{iT}$ refer to individual equation OLS estimator of the spurious regression slope coefficients β_i in models of type (2), or a hybrid model such as (3). Clearly, the asymptotic properties of the $\hat{\beta}_{iT}$ estimator will not differ from the standard results for single series spurious regressions. In contrast, the asymptotic properties of the estimator $\hat{\beta}_{NT}$ for models of type (1) can differ markedly. In particular, for the estimator $\hat{\beta}_{NT}$ we have the following proposition.

Proposition 2.1 (Spurious Regressions for Homogeneous Panels): *If $e_{it} \sim I(1)$ for all i, then for all models (1) with homogenous coefficients and $\Omega_i = \Omega$,*

(a). $\hat{\beta}_{NT} \sim O_p(N^{-1/2})$ *so that $\hat{\beta}_{NT}$ is a consistent estimator in the sense that $plim_{N,T\to\infty} \hat{\beta}_{NT} = 0$.*

(b). $t_{\hat{\beta}_{NT}} \sim O_p(T^{1/2})$ *so that $t_{\hat{\beta}_{NT}}$ diverges as $N,T\to\infty$.*

The consistency result for $\hat{\beta}_{NT}$ is in contrast to the classical spurious regression case of the single equation OLS estimators $\hat{\beta}_{iT}$, which are $O_p(1)$ nonconvergent as in Phillips (1986). Not surprisingly, the distinction is of fundamental importance for the properties of residual stationarity tests for cointegration that are based on estimates of the long run parameters α,δ,β associated with models (1) and (2) as we will see in the next subsection. First, however, it is worthwhile to note several more immediate implications of proposition 3.1 for spurious regressions in panels. Notice for example that cross sectional asymptotics is not in itself sufficient to remedy the spurious regression problem in the sense that standard t-statistic will continue to diverge as N and

T grow large. Thus, a researcher who chooses to ignore the unit root properties of the variables of a panel will still face the problem of falsely inferring a significant relationship among integrated variables with certainty as the time series dimension grows large in the absence of cointegration even as the cross sectional dimension grows large. Furthermore, the standardized asymptotic distributions of both $\hat{\beta}_{NT}$ and $t_{\hat{\beta}_{NT}}$ will in general be contaminated by data dependencies associated with dynamics reflected in the nuisance parameters of the long run covariance matrix Ω_i. However, the following general result does provide some insights regarding the conditions under which this data dependency does not pertain for certain special case distributions.

Lemma 2.1 *Let L_i be the lower triangular Cholesky decomposition of Ω_i. Then under the conditions of proposition 2.1, the convergence*

$$\sqrt{N}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} L_{22i}^{-2}\, x_{it}^2\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left((L_{11i}L_{22i})^{-1} x_{it} y_{it} - \frac{L_{21i}}{L_{11i}} x_{it}^2\right) \rightarrow N(0,2/3)$$

holds as $N,T\rightarrow\infty$.

Indirectly, this lemma illustrates the precise nature of the nuisance parameter dependencies for estimators such as $\hat{\beta}_{NT}$ and $t_{\hat{\beta}_{NT}}$. In particular, the left hand side of the expression more closely resembles the panel OLS estimator $\hat{\beta}_{NT}$ in the event that $L_{21i} = 0$, which corresponds to the case of strict exogeneity for the regressors. When the dynamics are also assumed to be homogeneous across individual sections so that the elements of L are no longer indexed by *i*, then the following special case results are obtained.

Proposition 2.2 (Special Case Distributions of $\hat{\beta}_{NT}$): *If the dynamics associated with the serial correlation patterns for all models (1) are also homogenous and the regressors are strictly exogenous, so that $\Omega_i = \Omega$ for all i with $\omega_{21} = 0$, then as $N,T\rightarrow\infty$*

(a). $\sqrt{\tfrac{3}{2}N}\,\hat{\beta}_{NT} \rightarrow N(0,\frac{\omega_{11}}{\omega_{22}})$

(b). $T^{-1/2} t_{\hat{\beta}_{NT}} \rightarrow N(0,2/3)$

In the event that the disturbances are i.i.d. so that $\Omega_i = L_i L_i^{/} = \Omega^o$, then the asymptotic variances ω_{11}, ω_{22} of part (a) are replaced by the more standard sample variance and part (b) continues to hold as is. In fact, these special cases correspond to the more restrictive assumptions that have commonly been used in more traditional panel data regressions. Table B.I of the appendix reports the results of Monte Carlo simulations for the finite sample distributions of $\sqrt{N}\hat{\beta}_{NT}$ and $t_{\hat{\beta}_{NT}}$ respectively for various combinations of N and T for 10,000 draws of independent random walks under the null spurious with no estimated constants and also for the case in which homogenous constants are estimated. The tables demonstrate that in this case the asymptotic distributions are very well approximated even for relatively small samples, and that furthermore the invariance to the estimation of homogenous constants holds well in relatively small samples.

In general however, it should be clear that these assumptions are likely to be overly restrictive for most of the purposes to which one would like to use cointegration tests. Of course in principle one might imagine constructing a feasible statistic free of nuisance parameters even in the most general case on the basis of the relationships in lemma 2.1. But lemma 2.1 clearly indicates that the feasible statistic would in general not be a simple transformation of either $\hat{\beta}_{NT}$ or $t_{\hat{\beta}_{NT}}$. Instead, as we shall see in the next subsections, transformations based on residual stationarity tests will provide more straightforward feasible statistics. Furthermore, in contrast to statistics based on lemma 2.1, statistics based on estimated residuals can be made to accommodate models such as (2) that also include heterogenous long run parameters.

2.3 Asymptotic Properties of Cointegration Tests in Panels

In this subsection we consider the asymptotic properties and distributions of test statistics for cointegration based on estimated residuals of the spurious regressions of models (1) through (3). We will consider both residual based stationarity tests as well as residual based variance ratio tests. In particular, for the latter, define the panel autoregressive coefficient estimator $\hat{\rho}_{NT}$ as

$$\hat{\rho}_{NT} - 1 \equiv \left(\sum_{i=1}^{N} \sum_{t=1}^{T} e_{it-1}^{2} \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(e_{it-1} \Delta e_{it} - \hat{\lambda}_i \right) \qquad \textbf{(7)}$$

under the null of nonstationarity for the residuals e_{it}, where $\hat{\lambda}_i$ is based on the estimated residuals of the autoregression $e_{it} = \rho e_{it-1} + \mu_{it}$ such that $\hat{\lambda}_i = T^{-1}\sum_{s=1}^{k_i} w_{sk_i}\sum_{t=1}^{T}\hat{\mu}_{it}\hat{\mu}_{it-s}$ for some lag window w_{sk_i}, so that $\hat{\lambda}_i$ is a scalar analogous to $\hat{\Gamma}_i$ defined in (5) above, in that it is based on autoregressions of the scalar time series e_{it} rather than the vector time series Z_{it}. Thus, $\hat{\rho}_{NT}$ can be thought of as a pooled version of the unstandardized Phillips and Perron statistic. The nonparametric estimation of $\hat{\lambda}_i$ is not critical for the first two result, though, as they will hold equally well for parametric unit root tests constructed along the lines of the Augmented Dickey-Fuller statistic. The first result is a direct consequence of proposition 2.1 for spurious regressions.

Proposition 3.1 (Asymptotic Equivalency Results): *If $e_{it} \sim I(1)$ for all i, then*

(a). *The distribution of residual stationarity tests in all models (1) with homogeneous coefficients are invariant to whether the residuals e_{it} are known or are estimated.*

(b). *The same is not true for models of type (2) and (3) with heterogeneous coefficients.*

Thus part (a) provides a type of superconsistency result that operates even in the absence of cointegration for certain panels as N grows large. As a consequence, tests of cointegration for homogenous panels of model type (1) will have the same critical values for large N as raw unit root tests applied to similarly dimensioned panels. The following result derives then as a direct corollary to proposition 3.1.

Corollary 3.1 (Special Case Asymptotic Distributions for Homogenous Panels): *Panel unit root tests applied to the estimated residuals of models (1), (1'), or (1'') with homogenous coefficients will be distributed as follows under the null of no cointegration.*

(a). $T\sqrt{N}(\hat{\rho}_{NT}-1) \rightarrow N(0,2)$ *as* $N,T\rightarrow\infty$ *for general* $\Omega_i = \Omega > 0$ *disturbances.*

(b). $\sqrt{NT(T-1)}\,(\hat{\rho}_{NT}-1) \rightarrow N(0,2)$ *as* $N\rightarrow\infty$ *for i.i.d. disturbances, regardless of* T.

(c). $t_{\hat{\rho}_{NT}} \rightarrow N(0,1)$ *as* $N\rightarrow\infty$ *for i.i.d. disturbances, regardless of* T, *and as* $N,T\rightarrow\infty$ *for general* $\Omega_i = \Omega > 0$ *disturbances.*

Part (a) of the corollary is in fact the distribution that is obtained by Levin and Lin (1994) for simple unit root tests in panels. Result (b) is a special case asymptotic distribution that holds regardless of the time series dimension T, so that under certain conditions cointegration tests can be performed even for panels with extremely small time series dimensions such as T=5. Clearly, this result can also be expected to hold for the case in which there exists some degree of serial correlation as long as it is homogenous across the individual cross sections *i,* and of low order relative to the T, since in this case the dynamics could also be estimated parametrically as $N \rightarrow \infty$. What is perhaps more interesting, however, is that appropriate standardization for the distribution of the corresponding t-statistic in part (c) is the same for the two cases. In fact this result helps to explain why Monte Carlo studies for panel unit root tests such as reported in Levin and Lin indicate very close approximations to the asymptotic distributions even for very small T. The answer appears to be that for the special case i.i.d. random walks that are used in the simulations, the time series dimension is in fact not relevant as N grows sufficiently large.

Table B.II reports the results of Monte Carlo simulations for the finite sample distributions of the three statistics $T\sqrt{N}\,(\hat{\rho}-1)$, $\sqrt{NT(T-1)}\,(\hat{\rho}-1)$ and $t_{\hat{\rho}-1}$ of corollary 3.1 for various combinations of N and T for 10,000 draws of independent random walks under the null spurious with no estimated constants, as well as the case in which homogenous constants are estimated. The tables reflect the fact that although the distribution of the residual stationarity test applied to estimated residuals is asymptotically equivalent to the case in which the residuals are known, for finite samples the critical values will not be identical.[6] Specifically, in the case of cointegration testing, the asymptotic distribution additionally depends on $\hat{\beta}_{NT}$ converging to zero as N grows large, so that the residual stationarity tests do not reach their asymptotic distributions as quickly for finite N. Still, the asymptotic distributions are fairly well approximated, particularly in the case of the t-statistic, as N grows large even for small T. For example, in a Monte Carlo simulation for N=200, T=10, the mean and variance were calculated to be -0.02 and 1.00 respectively for the case of $t_{\hat{\rho}-1}$ with no constant.

In any case, it is important to keep in mind that although they can be very convenient

[6] For example if we compare these tables to similar tables reported for panel unit root simulations in Levin and Lin (1994), or Pedroni (1993) for $\sqrt{NT(T-1)}\,(\hat{\rho}-1)$ for various N and T.

under the appropriate circumstances, the results of corollary 3.1 should be used only in the special case of model (1), for which the alternate hypothesis of cointegration requires that the cointegrating relationships are also identical for each cross section. While some applications may fit this description, more generally one should expect the alternate hypothesis to be such that models of type (2) or (3) must be employed, since the alternate to no cointegration is likely to be that any one of a number of different and unknown cointegrating relationships hold for the individual members of the panel. In this case the asymptotic distributions for cointegration tests will differ substantially from the raw unit root tests. Furthermore, it becomes important in the general case for heterogenous panels to take account of the idiosyncratic feedback effect among variables. In particularly, we obtain the following very general result, which includes the case of heterogenous panels.

Proposition 3.2 (Asymptotic Distributions of Residual Tests of Cointegration in Heterogenous Panels): *Based on the estimated residuals of Models (2), define the following statistics.*

$$Z_{\hat{v}_{NT}} \equiv \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{L}_{11i}^{-2} \hat{e}_{it-1}^{2} \right)^{-1}$$

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

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

where $\tilde{\sigma}_{NT} \equiv \frac{1}{N} \sum_{i=1}^{N} (\hat{L}_{11i} \hat{\sigma}_i)^2$, $\hat{\lambda}_i = \frac{1}{2}(\hat{\sigma}_i^2 - \hat{s}_i^2)$, $\hat{s}_i^2 \equiv T^{-1} \sum_{t=1}^{T} \hat{\mu}_{it}^2$ *and* $\hat{L}_i$ *is based on consistent estimation of* $\hat{\Omega}_i$. *Let* Θ, ψ *signify the mean and variance for the vector Brownian motion functional* $\Upsilon' \equiv (\int Q^2, \int QdQ, \tilde{\beta}^2)$, *where* $\tilde{\beta} \equiv \frac{\int VW}{\int W^2}$, $Q \equiv V - \tilde{\beta} W$ *and* $\psi_{(i)}$ *, i = 1,2,3 refers to the i x i*

upper sub-matrix of ψ. *Then the asymptotic distributions of these statistics are free of any* *nuisance parameters associated with* Ω_i *, and for* $N,T \to \infty$ *under the null of no cointegration, are given by:*

$$T N^{3/2} Z_{\hat{v}_{NT}} - \Theta_1^{-1}\sqrt{N} \;\to\; N(0, \phi_{(1)}^{\prime}\psi_{(1)}\phi_{(1)})$$

$$T\sqrt{N}\,(Z_{\hat{\rho}_{NT}} - 1) - \Theta_2\Theta_1^{-1}\sqrt{N} \;\to\; N(0, \phi_{(2)}^{\prime}\psi_{(2)}\phi_{(2)})$$

$$Z_{t_{NT}} - \Theta_2(\Theta_1(1+\Theta_3))^{-1/2}\sqrt{N} \;\to\; N(0, \phi_{(3)}^{\prime}\psi_{(3)}\phi_{(3)})$$

with $\phi_{(1)}^{\prime} = -\Theta_1^{-2}$ *;* $\phi_{(2)}^{\prime} = -(\Theta_1^{-1}, \Theta_2\Theta_1^{-2})$ *;*
$\phi_{(3)}^{\prime} = (\Theta_1^{-1/2}(1+\Theta_3)^{-1/2}, -\frac{1}{2}\Theta_2\Theta_1^{-3/2}(1+\Theta_3)^{-1/2}, -\frac{1}{2}\Theta_2\Theta_1^{-1/2}(1+\Theta_3)^{-3/2})$

Notice that since by construction $\hat{L}_i$ is the lower triangular Cholesky decomposition of $\hat{\Omega}_i$ such that $L_{22i} = \sqrt{\omega_{22i}}$, this means that $\hat{L}_{11i}^2 = \hat{\omega}_{11i} - \frac{\hat{\omega}_{21i}^2}{\hat{\omega}_{22i}}$ is interpreted as the asymptotic conditional variance for the spurious regression. For independent random walks drawn from the standard normal distribution, L_{11i}^2 will be unity and the distributions in proposition 3.1 will obtain from a straightforward pooling of the more conventional single series cointegration tests. More generally, however, this asymptotic conditional covariance will not be unitary, and will tend to differ across members of the panel, so that this term will not be eliminated asymptotically. Thus, in contrast to the single series case, simple panel cointegration tests will in general contain nuisance parameters in the asymptotic distributions. However, as proposition 3.1 indicates, with an appropriate estimator for L_{11i}^2, panel cointegration statistics can be constructed that are free of nuisance parameters and will converge to the distributions as given, whose moments depend only on the properties of standard Brownian motion.

The following table gives the corresponding moments for the models of interest.

Table I. *Let* Θ, ψ *signify the mean and variance for the vector Brownian motion* *functional* $\Upsilon' \equiv (\int Q^2, \int QdQ, \tilde{\beta}^2)$, *where* $\tilde{\beta} \equiv \frac{\int VW}{\int W^2}$, $Q \equiv V - \tilde{\beta} W$. *Likewise, let* Θ', ψ' *and* Θ'', ψ'' *signify the mean and variance of the same functionals constructed from the demeaned and detrended Wiener processes* V', W' *and* V'', W'' *respectively. Then the approximations*

$$\Theta = \begin{bmatrix} 0.250 \\ -0.693 \\ 0.889 \end{bmatrix} \quad \psi = \begin{bmatrix} 0.110 & & \\ -0.011 & 0.788 & \\ 0.243 & -1.326 & 3.174 \end{bmatrix};$$

$$\Theta' = \begin{bmatrix} 0.116 \\ -0.698 \\ 0.397 \end{bmatrix} \quad \psi' = \begin{bmatrix} 0.011 & & \\ -0.013 & 0.179 & \\ 0.026 & -0.238 & 0.480 \end{bmatrix}; \quad \Theta'' = \begin{bmatrix} 0.056 \\ -0.590 \\ 0.182 \end{bmatrix}$$

$$\psi'' = \begin{bmatrix} 0.001 & & \\ -0.001 & 0.034 & \\ 0.003 & -0.042 & 0.085 \end{bmatrix}$$

are obtained on the basis of Monte Carlo simulations for 100,000 draws from pairs of independent random walks with T=1,000, N=1.

This enables us to derives asymptotic distributions for the panel cointegration statistics as N grows large on the basis of these moments for large T. The results are summarized in the following three corollaries to proposition 3.2. It should also be noted that the inclusion of any effects that are common across the members of the panels, such as common intercepts, trends or time dummies will not affect the asymptotic distributions for any of these distributions since these parameters can be estimated arbitrarily well as N grows large. Consequently, the following corollaries also apply equally well to any hybrid models such as (3) that also include common effects such as time dummies.

Corollary 3.2 (Variance Ratio Tests for Cointegration in Heterogenous Panels): *Let* $Z_{\hat{v}_{NT}}, Z'_{\hat{v}_{NT}}, Z''_{\hat{v}_{NT}}$ *refer to the panel variance ratio statistics computed for models (2), (2') and (2'') respectively. Based on the empirical moments given above for large T, the following approximations obtain as* $N \to \infty$ *under the null of no cointegration:*

(a). $TN^{3/2} Z_{\hat{v}_{NT}} - 4.00\sqrt{N} \;\to\; N(0, 27.81)$

(b). $TN^{3/2} Z'_{\hat{v}_{NT}} - 8.62\sqrt{N} \;\to\; N(0, 60.75)$

(c). $TN^{3/2} Z^{//}_{\hat{\nu}_{NT}} - 17.86\sqrt{N} \rightarrow N(0, 101.68)$

The usage for these panel variance ratio statistics is similar to the single series case, in that large positive values indicate rejections. Table B.III of the appendix reports finite sample distributions for various combinations of N and T based on Monte Carlo simulations for 20,000 draws of independent random walks. The next two corollaries relate to the more common residual based tests that use the first order autoregressive parameter estimates and the corresponding t-statistic.

Corollary 3.3 (Residual Autoregression Tests for Cointegration in Heterogenous Panels): *Let* $Z_{\hat{\rho}_{NT}}, Z^{/}_{\hat{\rho}_{NT}}, Z^{//}_{\hat{\rho}_{NT}}$ *refer to the panel variance ratio statistics computed for models (2), (2') and (2'') respectively. Based on the empirical moments given above for large T, the following approximations obtain as* $N \rightarrow \infty$ *under the null of no cointegration:*

(a). $T\sqrt{N}(Z_{\hat{\rho}_{NT}} - 1) + 2.77\sqrt{N} \rightarrow N(0, 24.91)$

(b). $T\sqrt{N}(Z^{/}_{\hat{\rho}_{NT}} - 1) + 6.02\sqrt{N} \rightarrow N(0, 31.27)$

(c). $T\sqrt{N}(Z^{//}_{\hat{\rho}_{NT}} - 1) + 10.54\sqrt{N} \rightarrow N(0, 39.52)$

Corollary 3.4 (T-Statistic Tests for Cointegration in Heterogenous Panels): *Let* $Z_{\hat{t}_{NT}}, Z^{/}_{\hat{t}_{NT}}, Z^{//}_{\hat{t}_{NT}}$ *refer to the panel variance ratio statistics computed for models (2), (2') and (2'') respectively. Based on the empirical moments given above for large T, the following approximations obtain as* $N \rightarrow \infty$ *under the null of no cointegration:*

(a). $Z_{\hat{t}_{NT}} + 1.01\sqrt{N} \rightarrow N(0, 1.50)$

(b). $Z^{/}_{\hat{t}_{NT}} + 1.73\sqrt{N} \rightarrow N(0, 0.93)$

(c). $Z^{//}_{\hat{t}_{NT}} + 2.29\sqrt{N} \rightarrow N(0, 0.66)$

For the panel rho- and t-statistics, large negative values indicate rejection of the null. In comparing these distributions to the ones applicable for raw panel unit root tests reported in Levin and Lin (1994), we see that the consequence of using estimated residuals is to affect not only the asymptotic variance, but also the rate at which the mean diverges asymptotically. In fact, for both statistics we see that for the standard case (a), while the mean does not diverge when applied to raw data, it does become divergent when applied to estimated residuals for heterogenous panels. Ignoring the consequences of the estimated regressors problem for the asymptotic bias in panels appears to lead the raw unit root statistic to become divergent when applied to residuals in these cases.

Finite sample distributions for each of these statistics are also reported in table B.III for various combinations of N and T for 20,000 draws of independent random walks under the null of no cointegration, and again, the finite sample distributions appear to be remarkably close to the asymptotic distributions even for panels with relatively small cross section and time series dimensions. For example, with N=20, T=25, the Monte Carlo experiment produces a mean and variance of -4.87 and 1.58 respectively for the standard case t-statistic, as compared to the values of -4.52 and 1.50 respectively that are predicted from the asymptotic distribution in (a) of corollary 3.4.[7]

III. An Empirical Application to the Purchasing Power Parity Hypothesis

Next, we consider the application of these panel cointegration statistics to the hypothesis of long run purchasing power parity, which can be interpreted in terms of a cointegrating relationship between nominal exchange rates and aggregate price differentials between countries. Since nominal exchange rates and price levels have been well documented to be integrated processes, a test for the null hypothesis of no purchasing power parity (PPP) becomes a test for the presence of a unit root in the residuals of an equation such as

$$s_{it} = \alpha_i + \beta_i p_{it} + e_{it} \qquad \textbf{(8)}$$

---

[7] Of course for finite samples, the degree to which these empirical sample distributions are relevant also depends on the extent to which higher order serial correlation is present. However, preliminary experiments regarding size distortions in the presence of heterogenous serial correlation within the panel appear to indicate relatively good approximations even in these cases.

where s_{it} is the log nominal bilateral U.S. dollar exchange rate at time *t* for country *i* , and p_{it} is the log price level differential between country *i* and the U.S. at time *t*. In the event that these variables are not strictly independent across countries and we wish to take account of the fact they are likely to share common disturbances, we can estimate common time dummies δ_t in the style of the hybrid model (3) as

$$s_{it} = \delta_t + \alpha_i + \beta_i p_{it} + e_{it} \qquad \textbf{(8)'}$$

One specific form of the PPP hypothesis states that the long run equilibrium relationship between s_{it} and p_{it} is directly proportional so that β_i is necessarily unity. Under this fairly narrow interpretation of PPP, the residuals e_{it} are interpreted as the log real exchange rate, and are computed directly by $e_{it} = s_{it} - p_{it}$ in order to test whether the log real exchange rate is best characterized as a unit root process. More generally, theory tells us that the relationship need not necessarily be exactly proportional. For example, the presence of international transportation costs, measurement errors (Taylor, 1988), differences in price indices (Patel, 1990), and differential productivity shocks (Fisher and Park, 1991) have been used to explain why in theory β_i should not be fixed at unity even if there do exist mechanisms that induce long run tendency toward an equilibrium relationship between aggregate prices and nominal exchange rates. On the other hand, it is unlikely that these factors should indicate any one particular value for β_i that is different from unity, and consequently empirical tests of the PPP hypothesis are typically performed on the basis of estimated residuals $\hat{e}_{it}$, since β_i itself must be estimated.

Furthermore, since the relative importance of various explanations for the nonunity of the relationship may differ among countries, there is no a priori reason to expect that these parameters should necessarily be the same for different countries. The point estimates of the intercept, α_i and slope β_i of regression (8) performed on a country by country basis for a collection of 25 different countries appear to confirm the suspicion that homogeneity of the these parameters would be an unwarranted assumption. As the first two columns of Table II indicate, the point estimates for the intercept vary considerably across countries, as do the also the slope estimates. The first number in each column indicates the estimate based on the annual data, while the second, in parentheses, indicate estimates based on the monthly data. The point estimates do not differ much between the annual and monthly data in most cases, but they do vary considerably among countries, with point estimates for the slope varying between 0.15 (0.48) for Belgium and 2.26 (2.12) for India, and 1.81 (1.85) for Japan. The intercepts vary even more, from -7.21 (-7.22) for Italy, and 1.31 (1.45) for Mexico, and 0.53 (0.53) for the U.K. For these reasons, the null of no PPP is best tested on the basis of the absence of a cointegrating relationship rather than on the basis of a raw unit root test. Furthermore, without strong evidence to the contrary, the corresponding parameters of the spurious regression must be allowed to vary across countries if

Table II. *Purchasing power parity estimates of individual country slopes and intercepts and individual cointegration statistics compared to heterogeneous panel cointegration statistics for the null of no PPP. Annual, T=20, and monthly, T=246, IFS data, June 1973 to Dec. 1994.*

Country	$\hat{\alpha}_i$ *(intercept)*		$\hat{\beta}_i$ *(slope)*		$TN^{3/2} Z_{\hat{v}_{NT}}$		$T\sqrt{N}(Z_{\hat{\rho}_{NT}} - 1)$		$Z_{\hat{t}_{NT}}$ *(PP)*		$Z_{\hat{t}_{NT}}$ *(ADF)*		K_i *(lags)*	
Belgium	-3.63	(-3.66)	0.15	(0.48)	6.86	(9.29)	-6.32	(-6.20)	-1.78	(-1.77)	-1.93	(-1.49)	1	(16)
Denmark	-1.97	(-2.01)	1.21	(1.43)	10.57	(12.12)	-6.48	(-6.61)	-1.82	(-1.82)	-1.79	(-1.68)	1	(13)
France	-1.90	(-1.92)	1.72	(1.71)	5.41	(14.71)	-7.40	(-7.75)	-1.96	(-1.97)	-2.02	(-2.16)	1	(6)
Germany	-0.67	(-0.70)	0.74	(0.70)	12.31	(15.09)	-7.07	(-7.91)	-1.91	(-2.00)	-1.88	(-1.77)	1	(17)
Ireland	0.35	(---)	0.74	(---)	6.88	(---)	-7.98	(---)	-2.05	(---)	-2.18	(---)	1	(-)
Italy	-7.21	(-7.22)	0.80	(0.88)	6.64	(13.72)	-8.05	(-7.53)	-2.03	(-1.94)	-1.96	(-1.62)	1	(11)
Netherland	-0.79	(-0.82)	0.68	(0.69)	12.48	(15.38)	-7.28	(-7.74)	-1.94	(-1.96)	-1.94	(-1.88)	1	(20)
Sweden	-1.81	(-1.80)	1.15	(1.24)	11.18	(12.86)	-6.32	(-6.56)	-1.77	(-1.80)	-1.83	(-1.57)	1	(20)
Switzerland	-0.51	(-0.54)	0.98	(1.16)	14.21	(17.42)	-8.43	(-9.21)	-2.12	(-2.14)	-2.13	(-2.09)	1	(17)
U.K.	0.53	(0.53)	0.57	(0.68)	16.18	(20.23)	-9.75	(-10.17)	-2.31	(-2.26)	* -3.36	(-2.38)	3	(21)
Canada	-0.20	(-0.20)	1.10	(1.42)	9.71	(9.05)	-7.30	(-6.11)	-2.03	(-1.72)	-2.51	(-1.62)	1	(16)
Japan	-5.14	(-5.19)	1.81	(1.85)	11.59	(13.75)	-9.68	(-8.93)	-2.33	(-2.12)	-2.22	(-1.81)	1	(18)
Greece	-4.60	(-4.57)	1.01	(1.03)	4.10	(12.21)	-6.24	(-6.59)	-1.76	(-1.81)	-2.87	(-1.88)	3	(14)
Iceland	-3.50	(---)	0.99	(---)	6.87	(---)	-7.75	(---)	-2.02	(---)	-2.37	(---)	3	(-)
Portugal	-4.79	(-4.77)	0.96	(1.02)	5.04	(9.49)	-5.39	(-4.84)	-1.69	(-1.53)	-2.73	(-1.60)	5	(16)
Spain	-4.74	(-4.74)	0.80	(0.86)	11.80	(12.27)	-6.66	(-6.17)	-1.83	(-1.75)	-2.04	(-1.50)	1	(16)
Turkey	-6.07	(-5.93)	1.11	(1.09)	7.55	(9.64)	-4.73	(-4.12)	-1.45	(-1.29)	-1.79	(-1.87)	3	(15)
Australia	-0.10	(---)	1.43	(---)	12.26	(---)	-8.12	(---)	-2.03	(---)	-1.98	(---)	1	(-)
N. Zealand	-0.42	(-0.38)	0.84	(1.19)	10.27	(21.26)	-8.56	(-11.18)	-2.14	(-2.36)	-2.14	(-2.75)	1	(18)
S. Africa	-0.42	(---)	1.14	(---)	5.47	(---)	-7.87	(---)	-2.05	(---)	-3.08	(---)	3	(-)
Chile	-4.94	(-4.84)	1.11	(1.18)	7.86	* (37.73)	-9.41	* (-50.19)	-2.68	* (-5.51)	-1.62	(-2.26)	1	(21)
Mexico	1.31	(1.45)	1.03	(1.04)	5.32	(18.87)	-9.73	(-9.57)	-2.33	(-2.06)	-2.10	(-2.61)	1	(10)
India	-2.42	(-2.37)	2.26	(2.12)	5.72	(5.04)	-12.93	(-8.62)	* -4.06	(-2.24)	-2.41	(-1.90)	1	(6)
Korea	-6.57	(-6.56)	1.08	(0.98)	10.61	(8.54)	-7.88	(-5.12)	-2.06	(-1.68)	-2.77	(-1.45)	5	(6)
Pakistan	-2.62	(---)	3.03	(---)	2.50	(---)	-4.88	(---)	-1.61	(---)	-2.14	(---)	4	(-)
Pooled (20)					35.93	* (54.37)	* -34.61	* (-34.90)	* -9.41	* (-8.95)	* -9.36	(-8.09)		
Pooled (25)					35.14		* -37.87		* -10.27		* -10.74			

Notes: *Estimated equation is (8):* $s_{it} = \alpha_i + \beta_i p_{it} + e_{it}$ *, where* s_{it} *is the log nominal bilateral U.S. dollar exchange rate at time t for country i , and* p_{it} *is the log price level differential between country i and the U.S. at time t. Estimates for monthly data are in parentheses. Pooled (20) excludes those countries for which monthly data is unavailable or incomplete over the chosen time span. An asterisk indicates rejections at the 10% level or better.*

they are to be pooled for the purposes of a panel cointegration test to avoid misspecification from a false homogeneity restriction.

Typically, cointegration tests of the PPP hypothesis have been conducted on a single country by country basis. For the most part, however, tests conducted on a single country basis have been largely unable to reject the null of no PPP for the post Bretton Woods period of floating exchange rates. Studies of this type include for example Taylor (1988), Corbai and Ouliaris (1988), Baillie and Selover (1987) and Patel (1990), to name only a few. On the other hand, studies that employ data spanning much longer time frames, based on Lee's (1976) annual exchange rate and CPI data from 1900 to 1973 combined with post Bretton Woods IFS data, generally do reject the null of no PPP. This leaves open the question of whether the failure to reject the null in the post 1973 data alone is due to an inherent regime change that no longer favors PPP or whether it is merely a reflection of the notoriously low power of cointegration tests when applied to short time spans. Adding further suspicion that low power is to blame for this result, a recent study by Fisher and Park (1991) found that for many cases, when the null hypothesis was reversed, the post Bretton Woods data was equally unable to reject the opposite null of PPP.

The panel cointegration test statistics developed in the previous section provide an obvious opportunity to make progress toward sorting out this issue. In lieu of increasing the time span of the data, the panel cointegration methodology permits one to bring additional data to bear upon the hypothesis by increasing the number of countries over the same time span, while at the same time allowing the structural dynamics and long run parameters to be completely heterogenous across the countries. Table II compares the results of applying conventional cointegration tests on a country by country basis to the PPP hypothesis versus applying the pooled panel cointegration tests developed in section II of this study. Columns 3 and 4 present the estimates for the variance ratio test statistic and the rho statistic respectively. Columns 5 and 6 report estimates for t-statistics, in the first case for the Phillips and Perron statistic, in the second for the Augmented Dickey Fuller statistic.[8] Reading across the rows of the table are the

[8] Raw unit root tests were also performed on s_{it} and p_{it} to confirm their nonstationarity.

individual country by country results for each of 25 countries, whose nominal exchange rates were permitted to float over at least some interval of the post Bretton Woods period. Both annual and monthly IFS data were used for these estimates wherever possible. Most studies based on the post Bretton Woods data typically use the relatively high frequency monthly IFS data given the very few number of annual observations that are available. Since pooling across countries allows one a viable alternative to increasing the time span or frequency, it is interesting to consider the consequences of using annual data. Thus, both cases are reported for ease of comparison, with the monthly estimates appearing in parentheses for each column. Finally, the last column reports the highest order lag that was required for each of the countries. More precisely the highest order lag was selected on the basis of the procedure recommended in Campbell and Perron (1991) for the ADF statistic, and the same value was used as a conservatively high kernel estimator truncation value for the other nonparametric statistics to facilitate more direct comparison, especially between the PP and ADF versions of the t-statistics.

Not surprisingly, on a country by country basis, the data is hard pressed to reject the null of no cointegration, for either the monthly or annual data. Of course this is largely consistent with previous studies. The 10% critical values in given in Phillips and Ouliaris are 27.85, -17.04 and -3.07 for the variance ratio statistic, the rho statistic, and the t-statistics respectively. On a single country basis, of 20 countries, only Chile produces rejections on the basis of the monthly data on three of the four statistics, but is unable to reject the null of no PPP for any of the statistics on the basis of the annual data. Using the annual data, the U.K. produces a rejection for one of the four statistics, as does India, and in these cases the monthly data does not reject. Considering the number of countries (25 annual) and number of different statistics, it can be safely said that on a country by country basis, both the annual and monthly data are clearly unable to reject the null. In most cases, the estimates are well to the left of the mean of the distribution, but to the right of the 10% critical value. The last two rows of the table report the results of the analogous panel cointegration statistics for the pooled sample where the statistics were computed as indicated in proposition 3.2, allowing for heterogenous slopes and intercepts and idiosyncratic serial correlation patterns and feedback effects and with the idiosyncratic lag truncations given by the corresponding country by country entries in column 7. The row labelled

"Pooled (20)" reports the values of these statistics applied to the panels of 20 countries for which both annual and monthly data are available for easy comparison, and thus excludes five countries for which the monthly data was either unavailable or incomplete over the chosen time period. The last row, labelled "Pooled (25)" reports the similar values for the statistics applied to the annual data of all 25 countries.

In contrast to the individual country by country results, by pooling together across countries, the heterogenous panel cointegration statistics obtain critical values that are clearly able to reject the null of no cointegration. Based on the asymptotic distributions reported in corollaries 3.3 through 3.5 the 10% critical values can be computed for the case of N=20 as 48.54, -34.09, and -8.97 for the panel variance ratio statistic, rho-statistic and t-statistics respectively. Thus, in comparing the monthly and annual panel cointegration test statistics in the row labelled Pooled (20), we see that for the monthly data the panel variance ratio statistic, the panel rho-statistic, and the nonparametric t-statistic all reject the null of no cointegration at approximately the 10% level or better. Only the parametric t-statistic is unable to reject at the 10% level for the monthly data. On the other hand, the parametric t-statistic does reject for the annual data, with only the variance ratio statistic unable to reject in this case. The same continues to hold for the case in the annual data from all 25 countries are used, since the corresponding 10% critical values become 53.09, -37.27 and -9.89 for N=25 according to corollaries 3.2b through 3.4b.

Considering that movements in nominal exchange rate and price variables for these countries are likely to be responding to similar common disturbances, it also makes sense to investigate the consequences of including common time dummies to capture these effects in the pooled regressions. Table IIb reports these results on the basis of regression equation (8)', for easy comparison.

Table IIb. *Purchasing power parity estimates of heterogeneous panel cointegration statistics for the null of no PPP with common time dummies.*

Country	$TN^{3/2}Z_{\hat{v}_{NT}}$	$T\sqrt{N}(Z_{\hat{\rho}_{NT}}-1)$	$Z_{\hat{t}_{NT}}$ *(PP)*	$Z_{\hat{t}_{NT}}$ *(ADF)*

Pooled (20)	36.20 * (62.07)	* -37.03 * (-39.35)	* -9.44 * (-9.53)	* -11.47 (-7.84)
Pooled (25)	38.33	* -40.05	* -10.50	* -13.30

Notes: *Estimated equation is (8)' :* $s_{it} = \alpha_i + \delta_t + \beta_i p_{it} + e_{it}$*. Otherwise, all notes are same as for Table II.*

Indeed, we see that for all but one of the 12 statistics reported, the results are strengthened by the inclusion of the time dummies.

Finally, Table B.III of the appendix was constructed in order to indicate the closeness of these asymptotic distributions to the empirical finite sample distributions for even relatively small sized panels based on independent random walks, and this provides an alternative possibility for comparing the estimated statistics. In this case, the approximate 10% critical values are somewhat larger in absolute value, producing 54.97 (51.01), -33.05 (-34.99), and -9.77 (-9.13) for the case of N=20, T=25 (T=250) for the heterogenous, demeaned panel variance ratio statistic, panel rho-statistic, and panel t-statistics respectively according to Table B.III. On this basis, the monthly data continues to reject for the panel variance ratio statistic and the panel rho statistic, but the annual data produces statistics that fall slightly below the 10% empirical critical values in this case in Table II, without time dummies. On the other hand, for N=25, T=25 the empirical critical values based on the same Monte Carlo simulations produce 10% critical values of 59.91, -36.03, and -10.77 respectively (not reported in the table). Thus, when the full sample of 25 countries is used, even on the basis of small sample empirical distributions the annual data continue to reject the null at the approximately the 10% level or better for the heterogenous panel rho-statistic and parametric panel t-statistic. Evidently, this slight improvement likely comes from the slightly larger size of the panel, which is just enough to boost the statistics to the corresponding 10% critical value, rather than idiosyncratic features of the additional countries, since individually they do not appear to disproportionately favor a rejection as compared to the other twenty. Alternatively, the inclusion of common time dummies as in Table IIb appears to preserve the rejections even for the most difficult case with only N=20 using annual data. Thus, overall, on the basis of both the asymptotic critical values from section II and the empirical Monte Carlo distributions of the appendix, the pooled data do appear to give fairly strong support for the PPP hypothesis.

IV. Concluding Remarks

The intention of the purchasing power parity application has been to demonstrate the manner in which the panel cointegration test statistics developed in this study can be successfully implemented even for relatively small panel data sets. By specifying a cointegration hypothesis that is considered to hold or not hold uniformly across members of a panel, the additional power obtained from pooling the multivariate unit root information alone across the panel can be sufficient to overturn results based on a member by member basis. More generally, one can argue that, for the very same reasons that one would expect the long run structural parameters α_i, β_i to differ across countries, one might expect them to change over long periods of time for any one country. In this case, one can also view pooling across members of a panel as a direct alternative to increasing the time span of a single member over a period in which these parameters may be changing. Either way, the broad applicability of such a procedure to numerous other empirical issues should be apparent.

Needless to say, the empirical application to the PPP hypothesis also raises a number of interesting issues, both empirical and methodological. For example, while empirical research has not generally uncovered strong evidence for cross country cointegrating relationships among exchange rates, the possibility certainly exists in many applications, including PPP. In this case, it will be interesting to know how inclusion of possible intra-panel cointegrating relationships under the null might affect the type of asymptotic distributions studied in this paper regarding the existence of intra-member cointegrating relationships. Furthermore, the finite sample distributions indicated by the Monte Carlo simulations in the appendix are performed for independent random walks, and the extent to which they reliably indicate the properties of more general finite sample regressions depends on the degree of temporal dependence present in the data. It is well known that the for small time series dimensions, the size distortions from estimating the asymptotic covariance matrix can be substantial in the case of high degrees of serial correlation. Preliminary Monte Carlo evidence indicates, however, that under certain conditions, the relative size distortions become much less severe as the number of cross sections are increased. It will be interesting to consider such issues in future research, which is also currently being undertaken.

REFERENCES

Abuaf and Jorion (1990) "Purchasing Power Parity in the Long Run," *The Journal of Finance*, 45, 157-74.

Andrews, D. and C. Monahan (1992) "An Improved Heteroscadasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 60, 953-66.

Baillie and Selover (1987) "Cointegration and Model of Exchange Rate Deterimination," International Journal of Forcasting, 3, 43-51.

Corbae, D. and S. Ouliaris (1988) "Cointegration and Tests of Purchasing Power Parity," *Review of Economics and Statistics*, 70, 508-11.

Campbell, J. and P. Perron (1991) "Pitfalls and Opportunities; What Macro-economists Should Know About Unit Roots," *NBER Macroeconomics Annual*, 6.

Cheung, Y. and K Lai (1993) "Long-Run Purchasing Power Parity During the Recent Float," *Journal of International Economics*, 34, 181-92.

Den Haan, W. and A. Levin (1995) "Inferences from parametric and non-parametric covariance matrix estimation procedures, International Finance Discussion Papers, No. 504.

Fisher E. and J. Park (1991) "Testing Purchasing Power Parity Under the Null Hypothesis of Cointegration," *The Economic Journal*, 101, 1476-84.

Holz-Eakon, D., W. Newey, and H. Rosen (1988) "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56.

Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1991) "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," Econometrica 59, 1551-80.

Lee, M. (1978) PURCHASING POWER PARITY, Dekker, New York.

Levin, A. and C. Lin (1994) "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," University of California, San Diego Discussion Paper.

MacDonald, R. (1993) "Long-Run Purchasing Power Parity: Is it for Real?" *Review of Economics and Statistics*, 690-95.

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

Park, J. and M Ogaki (1991) "Seemingly Unrelated Canonical Cointegrating Regressions," The Rochester Center for Economic Research, Working Paper No. 280.

Park, J. and M. Ogaki (1991) "Inference in Cointegrated Models Using VAR Prewhitening to Estimate Shortrun Dynamics," The Rochester Center for Economic Research, No. 281.

Park, J.Y. and P.C.B Phillips (1989) "Statistical Inference in Regressions with Integrated Processes: Part II," *Econometric Theory*, 5, 95-131.

Park, J.Y. and P.C.B Phillips (1988) "Statistical Inference in Regressions with Integrated Processes: Part I," *Econometric Theory*, 4, 468-497.

Patel, J. (1990) "Purchasing Power Parity as a Long Run Relation," *Journal of Applied Econometrics*, 5, 367-79.

Pedroni P. (1993) "Endogenous Growth, Human Capital, and Cointegration for Multi-Country Panels," Mimeo, Columbia University.

Perron, P. (1989) "Testing for a Random Walk: A Simulation Experiment of Power When the Sampling Interval Is Varied," in B. Jaj, ed., Advances in Econometrics and modelling (Kluwer Academic Publishers, Dordrecht) 47 -68.

Perron, P. (1991) "Test Consistency with Varying Sampling Frequency," *Econometric Theory*, 7, 341-68.

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

Phillips, P.C.B. (1987) "Time Series Regression with a Unit Root," *Econometrica*, 55, 227-301.

Phillips and Durlauf, (1986) "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53, 473-95.

Phillips, P. and 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.

Pierce and Snell (1995) "Temporal Aggregation and the Power of Tests for a Unit Root," *Journal of Econometrics*, 65, 333-45.

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

Shiller and Perron (1985) "Testing the Random Walk Hypothesis: Power Versus Frequency of Observation, *Economic Letters*, 18, 381-86.

Taylor, M. (1988) "An Empirical Examination of Long-Run Purchasing Power Parity Using Cointegration Techniques," *Applied Economics*, 20, 1369-82.

MATHEMATICAL APPENDIX

Proposition 2.1:

(a). Write

$$\sqrt{N}\hat{\beta}_{NT} = \frac{\frac{1}{\sqrt{N}}\sum_i^N \left(T^{-2}\sum_t^T x_{it}y_{it} \right)}{\frac{1}{N}\sum_i^N \left(T^{-2}\sum_t^T x_{it}^2 \right)'} \rightarrow \frac{\frac{1}{\sqrt{N}}\sum_i^N \left(\Omega_i^{1/2}\int_0^1 Z_i(r)Z_i(r)'dr\Omega_i^{1/2} \right)_{21}}{\frac{1}{N}\sum_i^N \left(\Omega_i^{1/2}\int_0^1 Z_i(r)Z_i(r)'dr\Omega_i^{1/2} \right)_{22}} \sim O_p(1) \qquad \textbf{(A1)}$$

by virtue of (6a). If $\Omega_i = \Omega$ for all i, then the numerator converges to an $O_p(1)$ random variable by a standard central limit theorem, and the denominator converges to a nonzero constant by the law of large numbers.

(b). Write

$$T^{-1/2}t_{\hat{\beta}_{NT}} = \sqrt{N}\hat{\beta}_{NT}\left(\frac{T^{-1}\hat{\sigma}_{NT}^2}{\frac{1}{N}\sum_i^N (T^{-2}\sum_t^T x_{it}^2)} \right)^{-1/2} \sim O_p(1) \qquad \textbf{(A2)}$$

since $T^{-1}\sigma_{NT}^2 = \frac{1}{NT^2}\sum_i^N\sum_t^T \hat{e}_{it}$ *converges to a nonzero constant for* $e_{it} \sim I(1)$.

Lemma 2.1:

Expanding (A1) in terms of the elements of $Z_i(r)$ gives

$$\frac{\frac{1}{\sqrt{N}}\sum_i^N \left(T^{-2}\sum_t^T x_{it}y_{it} \right)}{\frac{1}{N}\sum_i^N \left(T^{-2}\sum_t^T x_{it}^2 \right)} \rightarrow \frac{\frac{1}{\sqrt{N}}\sum_i^N \left(L_{11i}L_{22i}\int_0^1 V_iW_i + L_{21i}L_{22i}\int_0^1 W_i^2 \right)}{\frac{1}{N}\sum_i^N \left(L_{22i}^2\int_0^1 W_i^2 \right)} \qquad \textbf{(A3)}$$

where the index r has been dropped for notational convenience. Rearranging (A3) gives

$$\sqrt{N}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} L_{22i}^{-2}x_{it}^{2}\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left((L_{11i}L_{22i})^{-1}x_{it}y_{it} - \frac{L_{21i}}{L_{11i}}x_{it}^{2}\right) \rightarrow \frac{\frac{1}{\sqrt{N}}\sum_{i}^{N}\left(\int_0^1 V_i W_i\right)}{\frac{1}{N}\sum_{i}^{N}\left(\int_0^1 W_i^2\right)} \qquad \textbf{(A4)}$$

Since $V_i(r), W_i(r)$ are standard independent Wiener processes, we can calculate the moments of the functionals in (A4) as

$$E_t\int_0^1 W_i(r)V_i(r)dr = 0 \;\; ; \quad E_t\int_0^1 W_i^2(r)dr = \int_0^1 r\,dr = \frac{1}{2} \qquad \textbf{(A5)}$$

$$\begin{aligned} Var_t\int_0^1 W_i(r)V_i(r)dr &= E_t\int_0^1\int_0^1 W_i(r)W_i(s)V_i(r)V_i(s)ds\,dr \\ &= 2\int_0^1\int_0^q E_tW_i(q)W_i(s)V_i(q)V_i(s)ds\,dq = 2\int_0^1\int_0^q s^2 ds\,dq = \frac{1}{6} \end{aligned} \qquad \textbf{(A6)}$$

Thus, by applying a standard central limit theorem to the numerator and law of large numbers to the denominator in (A4) we obtain convergence to a random variable with distribution N(0,2/3).

Proposition 2.2:

(a). Let $\omega_{21} = 0$. Then for $\Omega_i = \Omega$, $L_{11} = \left(\omega_{11} - \frac{\omega_{21}^2}{\omega_{22}}\right)^{1/2} = \sqrt{\omega_{11}}$. Substituting these into Lemma 2.1 gives the result.

(b). Under these restrictions, the numerator of the ratio in (A2) converges to $\frac{1}{2}L_{22}^2 = \frac{1}{2}\omega_{22}$ by virtue of (A5), and since $plim_{N,T\rightarrow\infty}\hat{\beta}_{NT} = 0$, then

$$T^{-1}\sigma_{NT}^2 \rightarrow \frac{1}{N}\sum_{i}^{N} L_{11i}^2\int V_i^2 + L_{11i}L_{21i}\int V_iW_i + L_{21i}^2\int W_i^2 - \hat{\beta}_{NT}\frac{1}{N}\int W_i^2 \rightarrow \frac{1}{2}L_{11}^2 = \frac{1}{2}\omega_{11} \qquad \textbf{(A7)}$$

which gives the result when substituted into (A5).

Proposition 3.1: *Using the relationship $\hat{e}_{it} = e_{it} - (\hat{\beta} - \beta)x_{it}$ for $\beta = 0$, under the null spurious,* *write*

$$T(\hat{\rho}-1) = \frac{\frac{1}{T}\sum_i^N \sum_t^T \hat{e}_{it-1}\Delta\hat{e}_{it}}{\frac{1}{T^2}\sum_i^N \sum_t^T \hat{e}_{it-1}^2} = \frac{\frac{1}{T\sqrt{N}}\sum_i^N \sum_t^T e_{it-1}\Delta e_{it} - \frac{1}{\sqrt{N}}\left[\sqrt{N}\hat{\beta}\right]\left[\frac{1}{T\sqrt{N}}\sum_i^N \sum_t^T x_{it-1}\Delta e_{it}\right]}{\frac{1}{T^2\sqrt{N}}\sum_i^N \sum_t^T e_{it-1}^2 - \frac{1}{\sqrt{N}}\left[N\hat{\beta}^2\right]\left[\frac{1}{T^2\sqrt{N}}\sum_i^N \sum_t^T x_{it-1}^2\right]} \qquad \textbf{(A8)}$$

so that the terms that are subtracted from the numerator and denominator represent the influence on the distribution of estimating β.

(a). *For* $\hat{\beta} = \hat{\beta}_{NT}$, $\sqrt{N}\hat{\beta} \sim O_p(1)$, *so that the bracketed expressions are each* $O_p(1)$ *and the entire subtracted terms each become* $O_p(N^{-1/2})$, *compared to the first terms which are* $O_p(1)$. *Thus, the contribution of the subtracted terms vanishes asymptotically, and*

$$\begin{aligned} plim_{N,T\to\infty}\left(\frac{1}{T\sqrt{N}}\sum_i^N \sum_t^T \hat{e}_{it-1}\Delta\hat{e}_{it}\right) &= plim_{N,T\to\infty}\left(\frac{1}{T\sqrt{N}}\sum_i^N \sum_t^T e_{it-1}\Delta e_{it}\right) \\ plim_{N,T\to\infty}\left(\frac{1}{\sqrt{N}T^2}\sum_i^N \sum_t^T \hat{e}_{it-1}^2\right) &= plim_{N,T\to\infty}\left(\frac{1}{\sqrt{N}T^2}\sum_i^N \sum_t^T e_{it-1}^2\right) \end{aligned} \qquad \textbf{(A9)}$$

(b). *For* $\hat{\beta} = \hat{\beta}_{iT}$, $\sqrt{N}\hat{\beta} \sim O_p(N^{1/2})$ *and the entire subtracted terms become* $O_p(1)$ *so that they are of the same order as the other terms. Thus, the influence of the subtracted term does not vanish asymptotically and condition (A9) does not hold.*

Corollary 3.1:

(a). *Since under the null* $plim_{N,T}\hat{\beta}_{NT} \to 0$ *from proposition 2.2(a), use (A9) to write*

$$T\sqrt{N}(\hat{\rho}_{NT}-1) = \frac{\frac{1}{T\sqrt{N}}\sum_i^N \sum_t^T \hat{e}_{it-1}\Delta\hat{e}_{it}}{\frac{1}{T^2 N}\sum_i^N \sum_t^T \hat{e}_{it-1}^2} \to \frac{\frac{1}{T\sqrt{N}}\sum_i^N L_{11}^2\int_0^1 V_i dV_i}{\frac{1}{T^2 N}\sum_i^N L_{11}^2\int_0^1 V_i^2} \to N(0,2) \qquad \textbf{(A10)}$$

since the Brownian motion functional of the numerator has moments 0, 1/2 and the functional of the denominator has mean 1/2.

(b). *Again, using the results of proposition 3.1, write the rho-statistic in terms of y as*

$$\left(\frac{1}{T^2\sqrt{N}}\sum_i^N\sum_t^T y_{it-1}^2\right)^{-1}\frac{1}{T\sqrt{N}}\sum_i^N\sum_t^T y_{it-1}\Delta y_{it} = \left(\frac{1}{\sqrt{N}}\sum_i^N T^{-2}\sum_t^T y_{it-1}^2\right)^{-1}\frac{1}{\sqrt{N}}\sum_i^N \frac{1}{2}\left((T^{-1/2}Y_T)^2 - T^{-1}\sum_t^T \Delta y\right.$$

where for $\Delta y_{it} \sim iid\ N(0,1)$, we have $E[(T^{-1/2}Y_{iT})^2] = 1$, $Var[(T^{-1/2}Y_{iT})^2] = 2$, $E[T^{-1}\sum_t^T \Delta y_{it}] = T^{-1}$ $Var[T^{-1}\sum_t^T \Delta y_{it}] = 2T^{-2}$ *and* $E[T^{-2}\sum_t^T y_{it-1}^2] = T^{-2}\sum_t^T E[y_{it-1}^2] = T^{-2}\sum_t^T (t-1) = \frac{1}{2}T(T-1)$. *Thus, as $N \to \infty$, the numerator converges to $N(0, T(T-1)/T^2)$ and the denominator converges to $\frac{1}{2}T(T-1)/T^2$, so that taken together these give the stated result.*

(c). *For the t-statistic, with iid disturbances, $\hat{\sigma} \to \sigma = 1$ as $N \to \infty$, and the remainder of the denominator converges to the square root of the term to which the denominator of the rho-statistic converges, giving the stated result.*

Proposition 3.2:

Using the relationships $\hat{e}_{it} = y_{it} - \hat{\beta}_{iT}\, x_{it}$, $\Delta\hat{e}_{it} = \Delta y_{it} - \hat{\beta}_{iT}\,\Delta x_{it}$, *write*

$$\equiv \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}\left(T^{-1}\sum_{t=1}^{T} y_{it-1}\Delta y_{it} - \hat{\beta}_{iT}\, T^{-1}\sum_{t=1}^{T} y_{it-1}\Delta x_{it} - \hat{\beta}_{iT}\, T^{-1}\sum_{t=1}^{T} x_{it-1}\Delta y_{it} + \hat{\beta}_{iT}^{2} T^{-1}\sum_{t=1}^{T} x \right.}{\frac{1}{N}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}\left(T^{-2}\sum_{t=1}^{T} y_{it-1}^{2} - 2\hat{\beta}_{iT}\, T^{-2}\sum_{t=1}^{T} y_{it-1}x_{it-1} + \hat{\beta}_{iT}^{2} T^{-2}\sum_{t=1}^{T} x_{it-1}^{2}\right)} \quad \textbf{(A11)}$$

Based on relationships (6a) and (6b), the terms of (A11) converge to the following

$$T^{-1}\sum_{t=1}^{T} y_{it-1}\Delta y_{it} \;\rightarrow\; L_{11i}^{2}\int V_i dV_i + L_{11i}L_{21i}\left(\int V_i dW_i + \int W_i dV_i\right) + L_{21i}^{2}\int W_i dW_i + \Gamma_{11i} \quad \textbf{(A12)}$$

$$T^{-1}\sum_{t=1}^{T} y_{it-1}\Delta x_{it} \;\rightarrow\; L_{11i}L_{22i}\int V_i dW_i + L_{21i}L_{22i}\int W_i dW_i + \Gamma_{21i} \quad \textbf{(A13)}$$

$$T^{-1}\sum_{t=1}^{T} x_{it-1}\Delta y_{it} \;\rightarrow\; L_{11i}L_{22i}\int W_i dV_i + L_{21i}L_{22i}\int W_i dW_i + \Gamma_{21i} \quad \textbf{(A14)}$$

$$T^{-1}\sum_{t=1}^{T} x_{it-1}\Delta x_{it} \;\rightarrow\; L_{22i}^{2}\int W_i dW_i + \Gamma_{22i} \quad \textbf{(A15)}$$

$$T^{-2}\sum_{t=1}^{T} y_{it-1}^{2} \;\rightarrow\; L_{11i}^{2}\int V_i^{2} + L_{11i}L_{21i}\int V_i W_i + L_{21i}^{2}\int W_i^{2} \quad \textbf{(A16)}$$

$$T^{-2}\sum_{t=1}^{T} y_{it-1}x_{it-1} \;\rightarrow\; L_{11i}L_{22i}\int V_i W_i + L_{21i}L_{22i}\int W_i^{2} \quad \textbf{(A17)}$$

$$T^{-2}\sum_{t=1}^{T} x_{it-1}^{2} \;\rightarrow\; L_{22i}^{2}\int W_i^2 \qquad \textbf{(A18)}$$

Furthermore, setting N=1 in (A3) gives

$$\hat{\beta}_{iT} \;\rightarrow\; \frac{L_{11i}L_{22i}\int V_iW_i + L_{21i}L_{22i}\int W_i^2}{L_{22i}^2\int W_i^2} \qquad \textbf{(A19)}$$

Thus, substituting expressions (A12) through (A19) into (A11) gives **(A20)**

$$1) \rightarrow \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}L_{11i}^{2}\left(\int V_i dV_i - \frac{\int V_iW_i}{\int W_i^2}\left(\int V_i dW_i + \int W_i dV_i\right) + \left(\frac{\int V_iW_i}{\int W_i^2}\right)^2\int W_i dW_i\right)}{\frac{1}{N}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}L_{11i}^{2}\left(\int V_i^2 - \frac{\left(\int V_iW_i\right)^2}{\int W_i^2}\right)} \;\rightarrow\; \frac{\frac{1}{\sqrt{N}}}{\frac{1}{N}}$$

where the last convergency is based on the definition of Q_i from Lemma 3.1 and consistent estimation of Ω_i, based for example on (5), so that $\hat{L}_{11i} \rightarrow L_{11i}$ as $T\rightarrow\infty$.

For the t-statistics, considider that $\hat{\mu}_{it} = \Delta\hat{e}_{it} - (\hat{\rho}-1)\hat{e}_{it-1} \rightarrow \Delta\hat{e}_{it}$ since $\hat{\rho}\rightarrow 1$ as $T\rightarrow\infty$. Thus, given that $T^{-1}\sum_t \Delta Z_{it}\Delta Z_{it}' = T^{-1}\sum_t \xi_{it}\xi_{it}' \rightarrow L_i'L_i = \Omega_i$, the term $\tilde{\sigma}_{NT}^2$ can be evaluated as

(A21)

$$\equiv \frac{1}{N}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}\hat{\sigma}_{iT}^2 \rightarrow \frac{1}{N}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}\left(L_{11i}^2 + L_{21i}^2 - 2\hat{\beta}_{iT}L_{21i}L_{22i} + \hat{\beta}_{iT}^2L_{22i}^2\right) \rightarrow \frac{1}{N}\sum_{i=1}^{N}\hat{L}_{11i}^{-2}\left(L_{11i}^2 + L_{11i}^2\left(\frac{\int V_iW_i}{\int W_i}\right.\right.$$

where the last convergency is obtained by substituting in the square of expression (A19). Hence

$$Z_{t_{\hat{\rho}_{NT}}} \;\rightarrow\; \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int Q_i dQ_i}{\sqrt{\left(\frac{1}{N}\sum_{i=1}^{N}\int Q_i^2\right)\left(1 + \frac{1}{N}\sum_{i=1}^{N}\tilde{\beta}_i^2\right)}} \qquad \textbf{(A22)}$$

The statistic $TN^{3/2}Z_{\hat{v}_{NT}}$ is given simply by setting the numerator of (A20) to unity. Thus, the *distributions of each of these statistics is free of nuisance parameters and will be governed only by the first two moments of the vector Brownian motion functional* $\Upsilon \equiv (\int Q^2, \int Q dQ, \tilde{\beta}^2)'$.

To obtain the precise relationship in terms of these moments, rewrite the statistics as

$$T N^{3/2} Z_{\hat{v}_{NT}} - \Theta_y^{-1}\sqrt{N} \;\rightarrow\; \sqrt{N}\left[\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{y\,i}\right)^{-1} - (\Theta_y)^{-1}\right] \qquad \textbf{(A23)}$$

$$_{T} - 1) - \Theta_x\Theta_y^{-1}\sqrt{N} \;\rightarrow\; \sqrt{N}\left[N^{-1}\sum_{i=1}^{N}\Upsilon_{x\,i} - \Theta_x\right]\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{y\,i}\right)^{-1} + \Theta_x\sqrt{N}\left[\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{y\,i}\right)^{-1}\right. \qquad \textbf{(A24)}$$

$$\hat{t}_{NT} - \Theta_x\Theta_y^{-1/2}(1+\Theta_z)^{-1/2}\sqrt{N} \;\rightarrow\; \sqrt{N}\left[N^{-1}\sum_{i=1}^{N}\Upsilon_{x\,i} - \Theta_x\right]\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{y\,i}\right)^{-1/2}\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{z\,i}\right)^{-1/}$$

$$\textbf{(A25)}$$

$$\sqrt{N}\left[\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{y\,i}\right)^{-1/2} - (\Theta_y)^{-1/2}\right]\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{z\,i}\right)^{-1/2} + \Theta_x\Theta_y^{-1/2}\sqrt{N}\left[\left(N^{-1}\sum_{i=1}^{N}\Upsilon_{z\,i}\right)^{-1/2} - (1+\right.$$

where indices y,x,z have been used in place of 1,2,3 respectively. As $N \rightarrow \infty$*, the summations in curved brackets converge to the means of the respective random variables by virtue of a law of large numbers. This leaves the expressions involving each of the square bracketed terms as a continuously differentiable transformation of a sum of i.i.d. random variables. In general, for a continuously differential transformation* Z_N *of an i.i.d. vector sequence* X_i *with vector mean u and covariance* Σ*,*

$$Z_N = \sqrt{N}\left(g\left(N^{-1}\sum_{i=1}^{N}X_i\right) - g(\mu)\right) \;\rightarrow\; N(0, \alpha'\Sigma\alpha) \qquad \textbf{(A26)}$$

as $N \rightarrow \infty$*, where the ith element of the vector* α *is given by the partial derivative* $\alpha_i = \frac{dg}{dg_i}(\mu_i)$ *(cf. Dhrymes, 1974). Thus, setting* $\mu_i = \Theta_i$, $\Sigma = \psi_{(i)}$, $\alpha = \phi_{(i)}$ *for each of the statistics provides the desired results.*

Corollaries 3.2-3.4:

Expanding the terms for the variances in proposition 3.2 gives

$$T\sqrt{N}\,Z_{\hat{v}_{NT}} - \Theta_y^{-1}\sqrt{N} \rightarrow N(0, \Theta_y^{-4}\psi_y) \qquad \textbf{(A27)}$$

$$T\sqrt{N}\,(Z_{\hat{\rho}_{NT}} - 1) - \Theta_x\Theta_y^{-1}\sqrt{N} \rightarrow N(0,\ \Theta_y^{-2}\psi_x - \Theta_x^{-2}\Theta_y^{-4}\psi_y - 2\Theta_x\Theta_y^{-3}\psi_{xy}) \qquad \textbf{(A28)}$$

$$T\sqrt{N}\,Z_{t_{NT}} - \Theta_x(\Theta_y(1+\Theta_z))^{-1/2}\sqrt{N} \rightarrow N(0,\zeta) \qquad \textbf{(A29)}$$

where

$$+\Theta_z)^{-1}\psi_x + \frac{1}{4}\Theta_x^2\Theta_y^{-3}(1+\Theta_z)^{-1}\psi_y + \frac{1}{4}\Theta_x^2\Theta_y^{-1}(1+\Theta_z)^{-3}\psi_z - \Theta_x\Theta_y^{-2}(1+\Theta_z)^{-1}\psi_{xy} - \Theta_x\Theta_y^{-1}(1+\Theta_z)^{-2}\psi_{xz} + \frac{1}{2}\Theta_x^2\Theta_y^{-2}($$

.

Substituting the empirical moments for large T *, N=1 into these expressions gives the reported approximations for the asymptotic distributions as* $N \rightarrow \infty$.