A PANIC Attack on Unit Roots and Cointegration

Jushan Bai* Serena Ng†

July 31, 2001

Preliminary and Incomplete

Abstract

This paper presents a toolkit for Panel Analysis of Non-stationarity in Idiosyncratic and Common components (PANIC). The point of departure is that the unobserved common factors can be consistently estimated, provided that the cross-section dimension is large. This allows us to decouple the issue of whether common factors exist from the issue of whether the factors are stationary. While unit root tests are imprecise when the common and idiosyncratic components have different orders of integration, direct testing on the two components are found to be more accurate in simulations. The tests are applied to a panel of real exchange rates.

Keywords: Common factors, common trends, principal components, unit root, cointegration, panel data

*Dept. of Economics, Boston College, Chestnut Hill, MA 02467 Email Jushan.Bai@bc.edu

†Dept. of Economics, Johns Hopkins University, Baltimore, MD 21218 Email: Serena.Ng@jhu.edu

1 Introduction

A hallmark development in time series econometrics during the last decade is the observation that a linear combination of N individually non-stationary series can be stationary. A system of variables with these properties is said to be cointegrated. The Granger Representation Theorem of Engle and Granger (1987) establishes that if a N-vector time series X_t has Wold representation $\Delta X_t = C(L)e_t$, then under cointegration, $C(1)$ has rank $k < N$ and ΔX_t has a vector error-correction representation (VECM). A large body of work has been devoted to the determination of k and efficient estimation of the parameters within the VECM framework under the assumption that N is small.

An implication of cointegration is that a system with N variables only has $k < N$ unit roots. It follows that under cointegration, non-stationarity in the data is driven by common stochastic trends. Idiosyncratic stochastic trends are possible only if there is no cointegration, in which case, $C(1)$ is full rank and ΔX_t is a VAR. The fact that a cointegrated system has common trends also implies that X_t can be decomposed into k non-stationary (T_t) and $r = N - k$ stationary (S_t) components. Stock and Watson (1988) suggest a T-S decomposition in which T_t is a random walk, while Gonzalo and Granger (1995) obtain a decomposition in which ΔT_t is serially correlated. Vahid and Engle (1993) suggest a different decomposition that allows common cycles to be identified along with the common trends. Nonetheless, the common stochastic trends are generally not regarded as objects of interest. In part, this is because the non-uniqueness of the T-S decomposition makes interpretation of the trends difficult. In part, this is also because under the small N assumption, consistent estimation of the cointegration parameters does not imply consistent estimation of the space spanned by the common trends. Hence even if the common trends can be given economic interpretation, the statistical properties of the estimated common trends are not well understood.

The starting point of our analysis is first, that common trends and cycles can be consistently estimated, and second, that cointegration restrictions are not necessary to identify these common variations. The main appeal of stepping outside the cointegration framework is that it makes it possible to decouple the issue of whether common factors exist from the issue of whether these factors are stationary. Once the common factors are estimated, a common-idiosyncratic (I-C) decomposition of the data follows. Our main contribution is to provide testing procedures to determine whether non-stationarity in a panel of data is induced by common sources or if it is of the idiosyncratic type. Because the common variations can be stationary or non-stationary, they are referred to as common factors rather than common trends. Such an analysis is made possible by use of large dimensional panels. Because N is large by assumption, pooling across units allows for consistent estimation of the common factors whether or not they are stationary. By virtue of the fact that N is large, more information can also be used for analysis. This not only improves statistical efficiency of the estimates, it also allows for identification of factors that are 'common' in a more general sense.

We consider a factor analytic framework

$$\begin{aligned} X_{it} &= c_i + \lambda_i' F_t + e_{it} \\ &= c_i + \lambda_i^{T\prime} F_t^T + \lambda_i^{S\prime} F_t^s + e_{it} \end{aligned}$$

where F_t is a $k \times 1$ vector of common factors, λ_i are the factor loadings, and e_{it} is an idiosyncratic error. The common factors can be non-stationary (F_t^T) or stationary (F_T^S). Thus, common factors with different orders of integration can co-exist. A factor model with N variables can have N idiosyncratic components and an unrestricted number of common factors that can be stationary and non-stationary In a I-C decomposition, non-stationarity in X_{it} can arise because of common stochastic trends, or because the idiosyncratic error is non-stationary. The factor model encompasses the multivariate local level model of Nyblom and Harvey (2000) as a special case. In that analysis, F_t^s is not permitted, N is fixed, and their object of interest is simply to determine k. Our interest is in understanding whether non-stationarity in X_{it} is due to F_t, e_{it}, or both, based on their consistent estimates. We refer to these procedures as Panel Analysis of Non-stationarity in the Idiosyncratic and Common components (PANIC).

Many macroeconomic issues of interest can be addressed using PANIC. Let X_{it} be output across countries. With integrated world markets, real output of country i may consist of a global growth component (F_t^T), a global cyclical component (F_t^S), and a idiosyncratic component which may itself be non-stationary. The framework then allows us to establish the source of non-stationarity and the relative importance of the country specific variations. In the same vein, let X_{it} be the inflation rate of good i, and core inflation be that component of inflation that is common to all goods. The proposed framework allows us to determine if an individual inflation series is non-stationary because core inflation is non-stationary, or because the series has idiosyncratic component has a unit root.

Testing stationarity of the common and idiosyncratic components separately can potentially a problem that has challenged unit root and stationarity tests. The problem arise when F_t is a weak random walk that is dominated by the variation of the stationary process e_{it}. It can also arise when e_{it} is $I(1)$ but F_t is stationary. In such cases, unit root tests on X_{it} tend to be oversized while stationarity tests tend to have no power. The issue is documented by Schwert (1989), and formally analyzed in Perron and Ng (1996) and Ng and Perron (1997). The problem arises because X_{it} has two components with different degrees of integration. However, if the non-stationary component is common to a large number of series, PANIC can be used to disentangle F_t from the idiosyncratic component. We can test whether these components are individually I(1) or I(0). Non-stationarity in any one component will then imply X_{it} is non-stationary.

PANIC makes use of information in a panel and yet it is not designed specifically for panel unit root tests, though they can be performed. The testing is applied to the estimated common and idiosyncratic components, one at a time. Thus, it can determine the proportion of the series in the panel that has unit roots. While pooling of the the tests based on observed data is invalid when the cross-sections are correlated, the I-C decomposition also allows us to pool on the basis of tests on the idiosyncratic components. Thus, panel unit root tests can also be obtained to test the joint stationarity of these idiosyncratic components.

The rest of the paper is organized as follows. The PANIC procedures are presented in Section 2. Asymptotic properties of the proposed tests are given in Section 3. Simulations are presented in Section 4, along with an application to real exchange rate data. Section 5 concludes. Proofs are given in the Appendix.

2 The PANIC Procedures

We consider the model

$$X_{it} = c_i + \lambda_i' F_t + e_{it} \tag{1}$$

$$F_{mt} = \alpha_m F_{mt-1} + u_{mt} \quad m = 1, \dots k \tag{2}$$

$$e_{it} = \rho_i e_{it-1} + \epsilon_{it} \quad i = 1, \dots N. \tag{3}$$

Equation (1) represents a series X_{it} as the sum of a common component $\lambda_i' F_t$ and an idiosyncratic component. The common factors F_t are k dimensional. Factor m is non-stationary if $\alpha_m = 1$. Our analysis permits some, none, or all of the factors to be non-stationary. The idiosyncratic component is stationary if $\rho_i < 1$ and has a unit root if $\rho_i = 1$. Furthermore, ϵ_{it} is allowed to be weakly correlated in the cross section dimension, so that (1) is an 'approximate factor model' in the sense of Chamberlain and Rothschild (1983).

We observe X_{it}, $i = 1, \dots N$, $t = 1, \dots T$. Non-stationarity in X_{it} can be due to the presence of common stochastic trends, a non-stationary idiosyncratic component, or both. These components are not, in general, identifiable from the observed data. Accordingly, the key to PANIC is an I-C decomposition of the data into common and idiosyncratic components. We use the method of principal components (PC) to estimate F_t. The PC has been suggested by Harris (1997) as an estimator of cointegrating vectors, but estimation of F_t is not considered. In fact, consistent estimation of F_t is not possible when N is small. However, this is no longer the case when N is large. This is because the idiosyncratic variations must be dominated by those of the common factors when the data are averaged across N. The principal components estimator effectively provides the weights for averaging. Asymptotic properties of the estimator when N and T are both large are analyzed in Forni, Hallin, Lippi and Reichlin (2000), Stock and Watson (1998), and Bai (2001a,b).

PANIC has two modules. The PANIC-UR allows us to test the null hypothesis that e_{it} has a unit root, while the PANIC-S takes stationarity as the null hypothesis. Specific implementation of the principal component estimator depends on the null hypothesis to be tested.

2.1 The PANIC-S

The PANIC-S procedure is set up to test

$$\begin{aligned} H_0 &: \rho_i < 0, \quad \text{or } e_{it} \quad \text{is } I(0) \quad \text{for all } i. \\ H_1 &: \rho_i = 1, \quad \text{or } e_{it} \quad \text{is } I(1) \quad \text{for some } i. \end{aligned}$$

Since e_{it} is stationary for every i, estimation of F_t is based on the data in level form.

S1: Estimate F_t by the method of principal components. Denote this by $\widehat{F}_t$.

S2: Test the null hypothesis that demeaned $\widehat{F}_{mt}, m = 1, \ldots k$ is stationary. Denote the statistic by $S_F(m)$.

S3: Given $\widehat{F}_t$, obtain $\widehat{e}_{it}$ as the residual from the regression

$$X_{it} = c_i + \lambda_i' \widehat{F}_t + e_{it}.$$

S4: Test the null hypothesis that $\widehat{e}_{it}$ is stationary, for each $i = 1, \ldots N$. Denote the test by $S_e(i)$.

We use the KPSS statistic of Kwiatkowski, Phillips, Schmidt and Shin (1992) to test stationarity.

2.2 The PANIC-UR

The PANIC-UR procedure is set up to test

$$\begin{aligned} H_0 &: \rho_i = 1 \quad \text{or } e_{it} \quad \text{is } I(1) \quad \text{for all } i. \\ H_1 &: \rho_i < 1 \quad \text{or } e_{it} \quad \text{is } I(0) \quad \text{for some } i. \end{aligned}$$

Since e_{it} is non-stationary for every i under the null hypothesis, we apply the method of principal components to the first differenced data,

$$\Delta X_{it} = \lambda_i' \Delta F_t + \Delta e_{it}. \tag{4}$$

UR1: Estimate $\widehat{\Delta F_t}$ and $\widehat{\lambda}_i$ by the method of principal components.

UR2: Given $\widehat{\Delta F_t}$, define for each $m = 1, \dots k$,

$$\widehat{F}_{mt} = \sum_{s=2}^{t} \widehat{\Delta F_{ms}}.$$

UR3: Test for the null hypothesis that demeaned $\widehat{F}_{mt}$ has a unit root for each $m = 1, \dots k$. Denote this test by $UR_F(m)$.

UR4: For each $i = 1, \dots N$, denote by $UR_e(i)$ the test for the null hypothesis that there is a unit root in $\widehat{e}_{it}$, where $\widehat{e}_{it}$ is defined as follows:

- If F_{mt} is I(0) for every $m = 1, \dots k$, then for $t = 2, \dots T$, let $\widetilde{e}_{it} = X_{it} - \widehat{\lambda}_i' \widehat{F}_t$ and define

$$\widehat{e}_{it} = \widetilde{e}_{it} - \widetilde{e}_{i2}.$$

- If we cannot reject $\widehat{F}_{mt}$ is I(1) for some m, then obtain $\widehat{e}_{it}$ as the residual from the regression

$$X_{it} = c_i + \lambda_i' \widehat{F}_t + e_{it}.$$

We consider the ADF test of Said and Dickey (1984) based on $\widehat{e}_{it}$.

3 Properties of PANIC

The validity of PANIC-S and PANIC-UR hinges on the fact that consistent estimates of F_t can be obtained by the method of principal components when N is large. The factors and their loadings are not, in general, separately identifiable. Assuming that $F'F/T = I_k$, we construct the principal component estimator for F_t as $\sqrt{T}$ times the eigenvector corresponding to the largest eigenvalue of the $T \times T$ matrix XX'. Denote this by $\widehat{F}_t$. Then $N \times k$ matrix λ is estimated as $\widehat{\Lambda} = \widehat{F}'X/T$. Consistency results are obtained in Bai (2001a) for stationary panels and in Bai (2001b) when the factors are non-stationary. The convergence rate of a rescaled version of $\widehat{F}_t$ is $\min[\sqrt{n}, T]$ for the stationary case and $\min[\sqrt{n}, T^{3/2}]$ in the non-stationary case. Since the space spanned by $\widehat{F}_t$ is preserved after scaling, these convergence results imply that PANIC can treat F_t as though it was known. Once consistent estimation of the factors is granted, the limiting distributions of the statistics are then obtained by application of the functional central limit theorem. That is, for a time series z_t satisfying mixing conditions stated in Phillips and Perron (1988)

$$\frac{1}{\sqrt{T}} \sum_{s=1}^{t} z_s \Rightarrow \sigma_z W_z(r),$$

where $\sigma_z = \lim_{T\to\infty} T^{-1}E(\sum_{j=1}^{T} z_j)^2$ and $W_z(r)$ is a standard Brownian motion. Following convention, limits for the demeaned and detrended processes are denoted $\sigma_z W_z^c(r)$ and $\sigma_z W_z^\tau(r)$ respectively. Furthermore, let $\bar{z} = \frac{1}{T}\sum_{t=1}^{T} z_t$. Then

$$\frac{1}{\sqrt{T}}\sum_{s=1}^{t}(z_s - \bar{z}) \Rightarrow \sigma_z\Big(W_z(r) - rW_z(1)\Big) \equiv \sigma_z V_z(r)$$

is a Brownian bridge.

Theorem 1 *PANIC-S: Suppose the KPSS statistic developed in Kwiatkowski et al. (1992) is used to test stationarity. Let V_{um} $(m = 1, \dots k)$ be Brownian bridges independent of V_{ei} $(i = 1, \dots N)$, which are N mutually independent Brownian bridges.*

1. *If F_{mt} is stationary, then*
$$S_F(m) \Rightarrow \int_0^1 V_{um}(r)^2 dr.$$
2. *If F_{mt} is $I(0)$ for every m, then for each $i = 1, \dots N$,*
$$S_e(i) \Rightarrow \int_0^1 V_{ei}(r)^2 dr.$$
3. *Suppose $\bar{k}$ of the factors are I(1). Then $S_e(i)$ has the same limiting distribution as the statistic developed in Shin (1994) for testing the null hypothesis of cointegration in a equation with $\bar{k}$ integrated regressors.*

Since $\widehat{F}_t$ can be consistently estimated, the test can treat the estimated factors as though they were known. As stated in part (1) of the Theorem, $S_F(m)$ has the same distribution as derived in Kwiatkowski et al. (1992). At the 5% level, this is 0.463. The limiting distribution for the $S_e(i)$ test depends on whether F_t is I(1) or I(0). Part (2) of the theorem states that if F_{mt} is I(0) for every $m = 1, \dots k$, $S_e(i)$ has the same limit as the KPSS stationary test. At the 5% level, this is also 0.463. If $\bar{k}$ factors are I(1), stationarity of e_{it} implies cointegration between X_i and a subset of F of dimension $\bar{k}$. Then $S_e(i)$ has the same limiting distribution as the Shin test for the null hypothesis of cointegration, as indicated in part (3) of the Theorem. The critical values are those in Shin (1994) for $\bar{k}$ regressors and a mean included in the cointegrating regression. At the 5% level, this is 0.314 for $\bar{k} = 1$ and 0.221 when $\bar{k} = 2$. In each case, the null hypothesis is rejected when the test statistic exceeds the critical value.

Theorem 2 *PANIC-UR: Suppose the ADF statistic of Said and Dickey (1984) is used. Let W_{um} $(m = 1, \dots k)$ and W_{ei} $(i = 1, \dots N)$ be standard Brownian motions.*

1. *If F_{mt} is I(1), then*

$$UR_F(m) \Rightarrow \frac{\int_0^1 W_{um}^c(r)dW_{um}(r)}{\int_0^r W_{um}^c(r)^2 dr}.$$

2. *if F_{mt} is $I(0)$ for every m, then for every $i = 1, \dots N$,*

$$UR_e(i) \Rightarrow \frac{\int_0^1 W_{ei}(r)dW_{ei}(r)}{\int_0^1 W_{ei}(r)^2 dr}.$$

3. *If F_{mt} is $I(1)$ for some $\bar{k}$ factors, $UR_e(i)$ has the same limiting distribution as the residuals based cointegration test of Phillips and Ouliaris (1990) with $\bar{k}$ integrated regressors.*

PANIC-UR is based on the the first differenced model which satisfies the assumptions of Bai (2001a). Since $\widehat{F}_t$ can be consistently estimated, $UR_F(m)$ has the same limiting distribution as derived in Fuller (1976) for the constant only case. The 5% asymptotic critical value is -2.86. The critical values of $UR_e(i)$ depend on whether F_t has a unit root. If F_{mt} is I(0) for every m, the critical values are those of a unit root test and are tabulated in Fuller (1976) for the case of no constant. At the 5% significance level, this is -1.95. Essentially, when F_t is $I(0)$, $\widehat{c}_i$ is $\sqrt{T}$ consistent and its effects on $\widehat{e}_{it}$ are asymptotically negligible. However, when $\bar{k}$ of F_t series are I(1), testing the null hypothesis that e_{it} is I(1) is the same as testing the null hypothesis that X_{it} does not cointegrate with $\bar{k}$ integrated factors. $UR_e(i)$ then becomes the residuals based test for no cointegration with $\bar{k}$ regressors. The critical values are tabulated in Phillips and Ouliaris (1990). When a mean is in the regression, the critical value at the 5% level is -3.37 for $\bar{k} = 1$ and -3.77 for $\bar{k} = 2$.

As defined, $\widehat{e}_{i1} = \widehat{e}_{i2} = 0$ for every i, which provides an initial condition assumption. One simpler alternative computationally is to take $\widehat{e}_{it}$ is the cumulative sum of $\widehat{\Delta e_{it}}$. Another method is to obtain $\widehat{e}_{it}$ as the residuals from a regression of X_{it} on a constant and $\widehat{F}_t$, which effectively sets the mean of the process to zero. The latter two methods appear to have similar properties in finite samples, though our proof is based explicitly on the assumption that $\widehat{e}_{i2} = 0$.

The number of factors k is unknown. Bai and Ng (2000) showed that $P(\widehat{k} = k) \to 1$ as $N, T \to \infty$ when F_{mt} is stationary for every m if $\widehat{k}$ is the minimizer of

$$\text{PC(m)} = \log \frac{1}{N} \sum_{i=1}^{N} \widehat{\sigma}_i^2(m) + kg(N,T) \tag{5}$$

where $\widehat{\sigma}_i^2(m) = \frac{1}{T}\sum_{t=1}^{T} \widehat{e}_{it}^2(m)$ is the sample variance from estimation of a model with m factors, and $g(N,T) \to 0$ as $N, T \to$ and $\min[N,T]g(N,T) \to \infty$. Bai (2001b) showed if we use the first differenced model and let $\widehat{\sigma}_i^2 = T^{-1}\sum_{j=1}^{T} \widehat{\Delta e_{it}^2}$, the PC criterion remains valid. This is also true even if ΔF_t or Δe_{it} is over-differenced. Furthermore, the PC consistently estimates the total number of factors whether or not the factors are non-stationary. Bai and Ng (2000) showed that

$$g(N,T) = \frac{N+T}{NT} \log \left[\frac{NT}{N+T} \right]$$

satisfies the conditions required for consistent estimation of k.

3.1 Panel Tests

Thus far, we have introduced PANIC as univariate tests on the common and idiosyncratic components. A common criticism of univariate unit root tests is low power, especially when T is small. This has generated substantial interest to improve power. A popular method is to pool information across units, leading to panel unit root tests. A survey of panel unit root tests is offered by Maddala and Wu (1999). The early test developed in Levin and Lin (1993) imposes substantial homogeneity but independence across units. Subsequent developments have led to tests that allow for heterogeneous intercepts and slopes, while maintaining the assumption of independence across units. This assumption is restrictive, and if violated, can lead to over-rejections of the null hypothesis. O´ Connell (1998) provides a solution to this problem assuming that all cross-section correlation is of the stationary type.

If cross-section correlation can be adequately represented by common factors, the idioyscractic components will more likely satisfy the assumption of independence across units. Thus, while panel unit root tests applied to X_{it} may be inappropriate, panel unit root tests can be applied to e_{it}. PANIC allows consistent estimates of e_{it} to be obtained. To allow as much heterogeneity across units as possible, we consider panel unit root tests based on meta analysis.

Theorem 3 *Let $p_{Se}(i)$ and $p_{URe}(i)$ be the p-value associated with the stationarity and unit root tests on $\widehat{e}_{it}$, respectively. Consider pooled tests defined by*

$$\begin{aligned} P_{Se} &= -2 \sum_{i=1}^{n} \log p_{Se}(i) \\ P_{URe} &= -2 \sum_{i=1}^{n} \log p_{URe}(i). \end{aligned}$$

If F_{mt} is I(0) for every $m = 1, \ldots k$,

$$\frac{P_{Se} - 2N}{\sqrt{4N}} \overset{d}{\Rightarrow} N(0,1) \quad \textit{and} \quad \frac{P_{URe} - 2N}{\sqrt{4N}} \overset{d}{\Rightarrow} N(0,1).$$

Theorems 1 and 2 show tests involving $\widehat{e}_{it}$ are functionals of Brownian motions that are independent across i. The p-values are thus independent U[0,1] random variables. Taking logarithms, summing over i, and applying central limit theorem gives the result stated. The pooled test $-2\sum_{i=1}^{N} \ln p_{Sx}(i)$ was first proposed in Maddala and Wu (1999) for fixed N. Choi (2001) extended the analysis to allow $N \to \infty$. But pooling is invalid when the units are not independent. Theorem 3 shows if the factors are stationary, pooling over tests based on $\widehat{e}_{it}$ is yields a test statistic that is standard normal. However, when a subset of F_t is I(1), the limiting distributions of the pooled tests will be mixture of χ^2 random variables

4 Monte Carlo Simulations

We simulate data using Equations (1)-(3). We let λ be a $N \times r$ matrix, ϵ a $T \times N$ matrix of $N(0,1)$ variables, and $u_t \sim N(0,1)\sigma_F$. The following variations are considered:

- σ_F=20, 10, 5, and 1.
- α=0, .5, .8, .9, .95, 1;
- ρ=0, .5, .8, .9, .95, 1;
- $k = 0, 1, 2$;
- N=10, 20, 30, 100;
- T=100, 200.

Simulations are performed assuming k (the total number of common factors) and $\bar{k}$ (the number of I(1) common factors) are known. Naturally, $\bar{k} = k$ if $\alpha_m = 1$ for every $m = 1, ldotsk$. We report results for $k = 1$, T=200 and N=20. Table 1 reports rejection rates for unit-root tests applied to $\{X_{1t}\}$, $\{\widehat{F}_{1t}\}$, and $\{\widehat{e}_{1t}\}$. The number of lags for the ADF test is set to $4(T/100)^{1/4}$ throughout. UR_{x1} should have a rejection rate of .05 when $\alpha = 1$ or $\rho = 1$. This is true only when $\alpha = 1$ but is not the case when $\rho = 1$, $\alpha < 1$. The problem is once again that the random walk component is small relative to the stationary variations of the idiosyncratic component, so that ΔX_{it} has a moving average root that is close to unity. The consequence of the near common factor is severe over-rejections of the unit root hypothesis. Such size distortions evidently arise because X_{it} has two components with different degrees of integration. Thus, if one of these components is common, PANIC will be able to consistently estimate it and separate unit root tests then become possible. The last six rows of Table 1 most clearly illustrate what PANIC can deliver. The column UR_{F1} is the rejection rate of the ADF test applied to $\widehat{F}_1$. It is around .05 when α is 1 and exhibit little size distortion. At other values of α, it reflects power. The column UR_{e1} is the rejection rate when the test is applied to $\widehat{e}_1$. It too should be .05 when $\rho = 1$, and the results show that the finite sample rejection rates are indeed close to the nominal size. In contrast, the UR_{x1} column shows that the

ADF test applied to the observed data does not give the correct inference. Table 2 reports results with different values for the variance σ_F. The finding is similar to that of Table 1.

More tables will be added.

5 Application to PPP

The PANIC can be used to shed new light on a much research issue:- the validity of the purchasing power parity (PPP) hypothesis. Under PPP, real exchange rates should be mean reverting and thus stationarity. The low power of unit root tests over short span has often been used to rationalize non-rejections of the unit root null. This has prompted the development of panel unit root tests which exploit information in the cross section dimension to increase power. But a major shortcoming of panel unit root tests is the assumption of independence across units. Indeed, because real exchange rates are often defined using the same base country, cross-section correlation arises almost by construction. Such strong cross-section correlation amounts to a common factor that cannot be aggregated away. Panel unit root tests based on pooling in the cross-section dimension will depend on this nuisance cross-section correlation parameter. As O´ Connell (1998) found, the tests are biased towards the alternative hypothesis. That is, if the null hypothesis is I(1), we tend to accept stationarity. If the null is stationary, we tend to accept nonstationarity. O'Connell suggest removing the cross-section correlation by a GLS transformation of the data. This presupposes that the common variation is stationary, which need not be the case. A PANIC allows us to test if real exchange rates are non-stationary because of a common non-stationary component, or if country-specific variations are the source of non-stationarity.

Appendix

Lemma 1 *Consistency of $\widehat{F}_t$.*

- *Suppose F_t is I(0) and Assumptions A to G of Bai (2001a) hold.. Then $\widehat{F}_t$ is $\sqrt{N}$ consistent if $\sqrt{N}/T \to 0$. If $\sqrt{n}/T \to \tau > 0$, $\widehat{F}_t$ is T consistent.*
- *Suppose F_t is I(1) and Assumptions A-F of Bai (2001b) hold. Then $\widehat{F}_t$ is $\sqrt{N}$ consistent if $N/T^3 \to 0$. If $N/T^3 \to \tau > 0$, $\widehat{F}_t$ is consistent at rate $T^{3/2}$.*

Stationarity and unit root applied to $\widehat{F}_t$ have the same distributions as though $\widehat{F}_t$ is known. Theorems 1.1 and 2.1 are stated in terms of the KPSS and the ADF test, but other tests are equally valid. The remaining proofs assume that F_t is known.

Proof of the Theorem 1, part 2

For a given i, consider the regression

$$X_{it} = c_i + \gamma_i' F_t + e_{it}.$$

Because F_t is stationary and an intercept is allowed in the regression, without loss of generality, we assume $E(F_t) = 0$. Let $\widehat{e}_{it}$ be the estimated residuals. By direct calculations,

$$\widehat{e}_{it} \quad = \quad e_{it} - (\widehat{c}_i - c_i) - (\widehat{\lambda} - \lambda)' F_t.$$

Note that $\widehat{c}_i$ and $\widehat{\lambda}_i$ are $\sqrt{T}$ consistent.

$$\begin{aligned}
\frac{1}{\sqrt{T}} \sum_{j=1}^{t} \widehat{e}_{ij} &= \frac{1}{\sqrt{T}} \sum_{j=1}^{t} e_{ij} - \frac{t}{\sqrt{T}} (\widehat{c}_i - c_i) - \frac{(\widehat{\lambda}_i - \lambda_i)'}{\sqrt{T}} \sum_{j=1}^{t} F_j \\
&= \frac{1}{\sqrt{T}} \sum_{j=1}^{t} e_{ij} - \frac{t}{T} \sqrt{T} (\widehat{c}_i - c_i) - \sqrt{T} (\widehat{\lambda}_i - \lambda_i)' \frac{1}{T} \sum_{j=1}^{t} F_j \\
&= \frac{1}{\sqrt{T}} \sum_{j=1}^{t} e_{ij} - \frac{t}{T} \sqrt{T} (\widehat{c}_i - c_i) - O_p(\frac{1}{\sqrt{T}})
\end{aligned}$$

But $\widehat{c}_i = \bar{X}_i - \widehat{\lambda}_i' \bar{F}$. We have $\sqrt{T}(\widehat{c}_i - c_i) = \frac{1}{\sqrt{T}} \sum_{j=1}^{T} e_{it} + o_p(1)$. Thus,

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{t} \widehat{e}_{ij} = \frac{1}{\sqrt{T}} \sum_{j=1}^{t} e_{ij} - \frac{t}{T} \frac{1}{\sqrt{T}} \sum_{j=1}^{T} e_{ij} + o_p(1).$$

Let $t = [Tr]$ and $t/T \to r$. By the functional central limit theorem, $\frac{1}{\sqrt{T}} \sum_{j=1}^{t} e_{ij} \Rightarrow W_{ei}(r)\sigma_{ei}$. Thus,

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{t} \widehat{e}_{ij} \Rightarrow \Big[W_{ei}(r) - r W_{ei}(1) \Big] \sigma_{ei}.$$

Let s_{ei}^2 be a consistent estimate of σ_{ei}^2, the long run variance of e_i. Then

$$S_e(i) = \frac{\frac{1}{T}\sum_{j=1}^{T}(\frac{1}{\sqrt{T}}\sum_{j=1}^{t}\widehat{e}_{ij})^2}{s_{ei}^2} \Rightarrow \int_0^1 (W_{ei}(r) - rW_{ei}(1))^2 dr \equiv \int_0^1 V_{ei}(r)^2 dr.$$

Proof of Theorem 2, part 2

Again, we assume F_t is known. The model is $X_{it} = c_i + \lambda_i' F_t + e_{it} = \widehat{c}_i + \widehat{\lambda}_i' F_t + \widehat{e}_{it}$. Set $\widehat{e}_{i1} = \widehat{e}_{i2} = 0$. Then $X_{i2} = \widehat{c}_i + \widehat{\lambda}' F_2$ which implies $\widehat{c}_i = X_{i2} - \widehat{\lambda}_i' F_2$. It follows that $\widehat{e}_{it} = X_{it} - \widehat{c}_i - \widehat{\lambda}_i' F_t = \widetilde{e}_{it} - \widetilde{e}_{i2}$, where $\widetilde{e}_{it} = X_{it} - \widehat{\lambda}_i' F_t$ by definition. We have

$$\begin{aligned}
\widehat{e}_{it} &= e_{it} - (\widehat{\lambda}_i - \lambda_i)' F_t - (\widehat{c}_i - c_i) \\
&= [1 \quad -\sqrt{T}(\widehat{\lambda}_i - \lambda_i)' \quad (\widehat{c}_i - c_i)] \begin{bmatrix} e_{it} \\ \frac{1}{\sqrt{T}} F_t \\ 1 \end{bmatrix} \equiv \widehat{b}' \zeta_t. \\
\Delta \widehat{e}_{it} &= \Delta e_{it} - (\widehat{\lambda}_i - \lambda_i)' \Delta F_t \\
&= [1 \quad -\sqrt{T}(\widehat{\lambda}_i - \lambda_i)'] \begin{bmatrix} \Delta e_{it} \\ \frac{1}{\sqrt{T}} \Delta F_t \end{bmatrix} \equiv \widehat{h} \eta_t.
\end{aligned}$$

Note that $\widehat{b} \Rightarrow b \equiv (1, b_2, b_3)'$ and $\widehat{h} \Rightarrow h \equiv (1, h_2)$. We want the limiting distribution of $\widehat{\rho}_i - 1$, where $T(\widehat{\rho}_i)$ is obtained a first order autoregression in $\widehat{e}_{it}$.

$$T(\widehat{\rho}_i - 1) = \frac{T^{-1}\sum_{t=2}^{T} \widehat{e}_{it-1} \Delta \widehat{e}_{it}}{T^{-2}\sum_{t=2}^{T} \widehat{e}_{it-1}^2}.$$

The numerator is

$$\begin{aligned}
T^{-1}\sum_{t=2}^{T} \widehat{e}_{it-1} \Delta e_{it} &= \widehat{b}' T^{-1} \sum_{t=2}^{T} \zeta_{t-1} \eta_t \widehat{h} \\
&= \widehat{b}' \begin{bmatrix} T^{-1}\sum_{t=2}^{T} e_{it-1}\Delta e_{it} & T^{-3/2}\sum_{t=2}^{T} e_{it-1}\Delta F_t \\ T^{-3/2}\sum_{t=2}^{T} F_{t-1}\Delta e_{it} & T^{-5/2}\sum_{t=2}^{T} F_{t-1}\Delta F_t \\ T^{-1}\sum_{t=2}^{T} \Delta e_{it} & T^{-3/2}\sum_{t=2}^{T} \Delta F_t \end{bmatrix} \widehat{h} \\
&\Rightarrow b' \begin{bmatrix} \sigma_{\epsilon i}^2 \int_0^1 W_{\epsilon i}(r) dW_{\epsilon i}(r) & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} h = \sigma_{\epsilon i}^2 \int_0^1 W_{\epsilon i}(r) dW_{\epsilon i}(r).
\end{aligned}$$

Analogous calculations give

$$T^{-2}\sum_{t=2}^{T} \widehat{e}_{it-1}^2 = \widehat{b}' [T^{-2}\sum_{t=2}^{T} \zeta_t \zeta_t'] \widehat{b} \Rightarrow \sigma_{\epsilon i}^2 \int_0^1 W_{\epsilon i}(r)^2 dr.$$

Combining the results,

$$T(\widehat{\rho}_i - 1) \Rightarrow \frac{\int_0^1 W_{\epsilon i}(r) dW_{\epsilon i}(r)}{\int_0^1 W_{\epsilon i}(r)^2 dr}.$$

The t statistic follows.

Table 1: Rejection rates for H_0: e_{it} is I(1)

				$\sigma^F = 20$				$\sigma^F = 5$			
T	N	ρ	α	UR_{x1}	UR_{F1}	UR_{e1}	UR^*_{e1}	UR_{x1}	UR_{F1}	UR_{e1}	UR^*_{e1}
200	20	0.00	0.00	1.00	1.00	0.71	0.70	1.00	1.00	0.71	0.72
200	20	0.00	0.50	1.00	1.00	0.70	0.69	1.00	1.00	0.68	0.68
200	20	0.00	0.80	0.96	0.96	0.67	0.67	0.98	0.97	0.70	0.71
200	20	0.00	0.90	0.68	0.67	0.68	0.69	0.71	0.68	0.72	0.72
200	20	0.00	0.95	0.26	0.25	0.69	0.68	0.32	0.27	0.72	0.70
200	20	0.00	1.00	0.07	0.07	0.74	0.74	0.08	0.07	0.75	0.74
200	20	0.50	0.00	1.00	1.00	0.80	0.79	1.00	1.00	0.81	0.81
200	20	0.50	0.50	1.00	1.00	0.81	0.81	1.00	1.00	0.81	0.80
200	20	0.50	0.80	0.98	0.98	0.81	0.81	0.98	0.97	0.81	0.79
200	20	0.50	0.90	0.65	0.64	0.81	0.80	0.73	0.67	0.80	0.79
200	20	0.50	0.95	0.30	0.27	0.82	0.82	0.32	0.24	0.81	0.81
200	20	0.50	1.00	0.07	0.06	0.90	0.89	0.08	0.05	0.88	0.89
200	20	0.80	0.00	1.00	1.00	0.81	0.81	1.00	1.00	0.81	0.82
200	20	0.80	0.50	1.00	1.00	0.81	0.81	0.99	1.00	0.81	0.82
200	20	0.80	0.80	0.97	0.97	0.82	0.84	0.97	0.97	0.82	0.82
200	20	0.80	0.90	0.68	0.66	0.83	0.83	0.72	0.67	0.83	0.83
200	20	0.80	0.95	0.28	0.26	0.81	0.82	0.33	0.25	0.81	0.81
200	20	0.80	1.00	0.07	0.06	0.76	0.76	0.11	0.07	0.78	0.77
200	20	0.90	0.00	0.97	1.00	0.78	0.76	0.91	1.00	0.77	0.77
200	20	0.90	0.50	0.98	1.00	0.74	0.75	0.93	1.00	0.75	0.74
200	20	0.90	0.80	0.95	0.97	0.76	0.76	0.91	0.97	0.76	0.76
200	20	0.90	0.90	0.67	0.68	0.76	0.77	0.69	0.69	0.76	0.75
200	20	0.90	0.95	0.29	0.28	0.75	0.75	0.34	0.27	0.74	0.74
200	20	0.90	1.00	0.06	0.05	0.31	0.33	0.11	0.06	0.33	0.32
200	20	0.95	0.00	0.91	1.00	0.53	0.54	0.65	1.00	0.52	0.52
200	20	0.95	0.50	0.94	1.00	0.52	0.53	0.75	1.00	0.52	0.51
200	20	0.95	0.80	0.91	0.97	0.52	0.53	0.74	0.97	0.51	0.50
200	20	0.95	0.90	0.61	0.65	0.53	0.52	0.54	0.66	0.49	0.51
200	20	0.95	0.95	0.26	0.24	0.54	0.53	0.26	0.26	0.53	0.53
200	20	0.95	1.00	0.06	0.06	0.10	0.10	0.07	0.06	0.09	0.09
200	20	1.00	0.00	0.73	1.00	0.06	0.06	0.27	0.98	0.06	0.06
200	20	1.00	0.50	0.79	1.00	0.05	0.06	0.41	1.00	0.04	0.05
200	20	1.00	0.80	0.80	0.96	0.05	0.05	0.47	0.95	0.05	0.05
200	20	1.00	0.90	0.56	0.68	0.06	0.06	0.32	0.62	0.06	0.06
200	20	1.00	0.95	0.23	0.27	0.05	0.06	0.17	0.28	0.04	0.04
200	20	1.00	1.00	0.05	0.05	0.02	0.02	0.07	0.07	0.02	0.01

Table 2: Rejection rates for H_0: e_{it} is I(1)

				$\sigma^F = 10$			$\sigma^F = 1$		
T	N	ρ	α	UR_{x1}	UR_{F1}	UR_{e1}	UR_{x1}	UR_{F1}	UR_{e1}
200	20	0.00	0.00	1.00	1.00	0.71	1.00	1.00	0.70
200	20	0.00	0.50	1.00	1.00	0.70	1.00	1.00	0.68
200	20	0.00	0.80	0.97	0.96	0.67	0.99	0.97	0.70
200	20	0.00	0.90	0.68	0.67	0.68	0.82	0.68	0.72
200	20	0.00	0.95	0.27	0.25	0.69	0.47	0.27	0.71
200	20	0.00	1.00	0.08	0.07	0.74	0.19	0.07	0.71
200	20	0.50	0.00	1.00	1.00	0.80	1.00	1.00	0.81
200	20	0.50	0.50	1.00	1.00	0.82	1.00	1.00	0.81
200	20	0.50	0.80	0.98	0.98	0.81	0.99	0.97	0.80
200	20	0.50	0.90	0.67	0.64	0.81	0.86	0.68	0.79
200	20	0.50	0.95	0.32	0.28	0.82	0.55	0.25	0.81
200	20	0.50	1.00	0.09	0.06	0.90	0.22	0.05	0.88
200	20	0.80	0.00	1.00	1.00	0.81	0.98	1.00	0.81
200	20	0.80	0.50	0.99	1.00	0.81	0.98	1.00	0.81
200	20	0.80	0.80	0.97	0.97	0.82	0.97	0.97	0.83
200	20	0.80	0.90	0.70	0.66	0.83	0.86	0.68	0.83
200	20	0.80	0.95	0.31	0.26	0.81	0.62	0.28	0.81
200	20	0.80	1.00	0.08	0.06	0.76	0.31	0.07	0.77
200	20	0.90	0.00	0.95	1.00	0.78	0.74	0.99	0.77
200	20	0.90	0.50	0.96	1.00	0.74	0.77	1.00	0.75
200	20	0.90	0.80	0.93	0.97	0.76	0.78	0.96	0.75
200	20	0.90	0.90	0.67	0.68	0.76	0.68	0.68	0.76
200	20	0.90	0.95	0.31	0.28	0.75	0.46	0.27	0.74
200	20	0.90	1.00	0.08	0.05	0.31	0.24	0.06	0.33
200	20	0.95	0.00	0.82	1.00	0.53	0.32	0.87	0.52
200	20	0.95	0.50	0.88	1.00	0.52	0.35	0.96	0.52
200	20	0.95	0.80	0.86	0.97	0.52	0.42	0.92	0.51
200	20	0.95	0.90	0.57	0.65	0.53	0.36	0.62	0.50
200	20	0.95	0.95	0.26	0.24	0.54	0.26	0.27	0.53
200	20	0.95	1.00	0.07	0.06	0.10	0.12	0.06	0.09
200	20	1.00	0.00	0.52	1.00	0.06	0.06	0.28	0.06
200	20	1.00	0.50	0.64	1.00	0.05	0.08	0.50	0.04
200	20	1.00	0.80	0.67	0.96	0.05	0.11	0.59	0.05
200	20	1.00	0.90	0.46	0.69	0.06	0.09	0.42	0.06
200	20	1.00	0.95	0.21	0.26	0.05	0.06	0.21	0.04
200	20	1.00	1.00	0.06	0.05	0.02	0.07	0.06	0.02

References

Bai, J. and Ng, S. (2000), Determining the Number of Factors in Approximate Factor Models, *forthcoming in Econometrica*.

Bai, J. S. (2001a), Inference on Factor Models of Large Dimensions, mimeo, Boston College. Revision requested from *Econometrica*.

Bai, J. S. (2001b), Estimating Cross-Section Common Stochastic Trends in Non-Stationary Panel Data, mimeo, Boston College. Submitted.

Chamberlain, G. and Rothschild, M. (1983), Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets, *Econometrica* **51**, 1305–1324.

Choi, I. (2001), Unit Root Tests for Panel Data, *Journal of International Money and Finance* **20**, 249–272.

Engle, R. F. and Granger, C. (1987), Cointegration and Error-Correction: Representation, Estimation, and Testing, *Econometrica* **55**, 251–276.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), The Generalized Dynamic Factor Model: Identification and Estimation, *Review pf Economics and Statistics* **82:4**, 540–554.

Fuller, W. A. (1976), *Introduction to Statistical Time Series*, John Wiley, New York.

Gonzalo, J. and Granger, C. (1995), Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Busness and Economic Statistics* **13**, 27–35.

Harris, D. (1997), Principal Components Analysis of Cointegrated Time Series, *Econometric Theory* **13:4**, 529–557.

Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. (1992), (Testing the Null Hypothesis of Stationarity Against the Alternative Of a Unit Root, *Journal of Econometrics* **54**, 159–178.

Levin, A. and Lin, C. F. (1993), Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. UCSD Discussion Paper 93-56.

Maddala, G. S. and Wu, S. (1999), A Comparative Study of Unit Root Tests with Panel data and a New Simple Test, *Oxford Bulletin of Economics and Statistics* **Special Issue**, 631–652.

Ng, S. and Perron, P. (1997), Estimation and Inference in Nearly Unbalanced and Nearly Cointegrated Systems, *Journal of Econometrics* **79**, 54–81.

Nyblom, J. and Harvey, A. (2000), Tests of Common Stochastic Trends, *Econometric Theory* **16:2**, 176–199.

O´ Connell, P. (1998), The Overvaluation of Purchasing Power Parity, *Journal of International Economics* **44**, 1–19.

Perron, P. and Ng, S. (1996), Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties, *Review of Economic Studies* **63**, 435–465.

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

Phillips, P. C. B. and Perron, P. (1988), Testing for a Unit Root in Time Series Regression, *Biometrika* **75**, 335–346.

Said, S. E. and Dickey, D. A. (1984), Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika* **71**, 599–607.

Schwert, G. W. (1989), Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics* **7**, 147–160.

Shin, Y. (1994), A Residuals Based Test for the Null of Cointegration Against the Alternative of No Cointegration, *Econometric Theory* **10:1**, 91–115.

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

Stock, J. H. and Watson, M. W. (1998), Diffusion Indexes, NBER Working Paper 6702.

Vahid, F. and Engle, R. F. (1993), Common Trends and Common Cycles, *Journal of Applied Econometrics* **8**, 341–360.