New Panel Unit Root Tests under Cross Section Dependence for Practitioners *

Donggyu Sul
Department of Economics,
University of Auckland

June 30, 2005

Abstract

This paper studies the principle of common recursive mean adjustment and proposes a new detrending method in dynamic panel models. By utilizing recursive mean adjustment, this paper provides three unit root tests: a recursive mean adjusted (RMA) unit root test, a covariate RMA and a pooled RMA-feasible generalized least squares tests. The first two tests are designed for testing the cross sectional average of panel time series data to examine if the common factors in a panel are stationary or not. The third test is designed to test if the idiosyncratic errors are stationary or not. The proposed panel unit root test under cross section dependence is precise and powerful especially when T is larger than N.

Keywords: Recursive detrending, Dynamic factors, Panel unit root test. Covariate unit root test, Cross section dependence

JEL Classification Numbers: C33 Panel Data

*I thank Yoosoon Chang, Nelson C. Mark, Alan Rogers, Peter C.B. Phillips and Peter M. Robinson for many constructive comments and suggestions. My thanks also go to Ryan Greenaway-McGrevy for excellent research assistantship. Email: D.SUL@auckland.ac.nz.

1 Introduction

Since Quah (1994) opened the door to the panel unit root testing literature, several important theoretical developments have been made by several researchers. Levin, Lin and Chu (2002) generalize Quah's (1994) panel unit root test under the alternative of a homogenous panel. Im, Pesaran and Shin (2003), Choi (2001) and Maddala and Wu (1999) consider panel unit root tests under the alternative of a heterogeneous panel. Along with the theoretical development of panel unit root tests, their use in empirical research has grown exponentially. The most important reason for their popularity is that panel unit root tests reject the null hypothesis of a panel unit root more often than univariate unit root tests. This is a natural result because the goal of the panel unit root tests under the null hypothesis of a panel unit root is to amplify the power of tests through the pooling of information across units.

The high power of panel unit root tests, however, suffers from serious size distortions when a panel is contaminated by cross section dependence. Choi (2005), Moon and Perron (2004a, MP) and Phillips and Sul (2003, PS) propose panel unit root tests under cross section dependence which arises from unknown common factors. Their tests require the elimination of the unknown common factors. Bai and Ng (2004, BN) propose an effective defactoring method. More importantly, BN go beyond this to analyze the source of non-stationarity. To fix ideas, assume that $\{y_{it}\}$ consist of common factors F_{st} and idiosyncratic errors M_{it}.

$$y_{it} = \lambda_{is}' F_{st} + M_{it} \tag{1}$$

where t $(t = 1, .., T)$ indexes the time series observations, the index i $(i = 1, ..., N)$ stands for the ith cross sectional unit and $s(s = 1, ..., K)$ indexes the number of common factors. MP and PS examine if M_{it} is I(1) or I(0) while BN test the common factor and the idiosyncratic errors separately. To do so, the number of common factors must be estimated. In simulation studies BN show that their tests perform very well in large N sample.

There are two remaining issues in BN's tests. First, if $N < 20$, there is a somewhat serious problem in estimating the number of common factors precisely, and this results in the size distortion of the tests, as well as the poor power of the tests. Second, testing for nonstationarity in y_{it} requires testing for nonstationarity in F_{st}. If the common factors F_{st} are I(1) , then y_{it} becomes I(1) regardless of whether M_{it} is I(0) or I(1). Suppose that $K = 1$. Then testing for a panel unit root in y_{it} hinges on a univariate unit root test of F_t. If so, practitioners may wonder where the panel gain comes from.

This paper addresses these two issues and proposes powerful unit root tests by utilizing recursive mean adjustments. Several authors have used recursive mean adjustment in testing for unit roots. So and Shin (1999), Shin and So (2001), Chang, Park and Phillips (2001), Taylor (2002) propose various recursive mean adjustments for detrending and demeaning as well. This paper goes beyond this to explain how recursive mean adjustments reduce small sample bias and to provide a new recursive detrending method. Combining the proposed recursive mean adjustment (RMA) methods with covariate unit root tests produces a very powerful univariate unit root test. The covariate RMA unit root test can be used to detect if common factors are stationary or not. When $N \geq 20$, this test can be applied directly to defactored $\hat{F}_{st}$ by using BN's method. When $N < 20$, the cross sectional average of y_{it}, which may be a good proxy for F_{st}, can be used for this test. Moreover, this paper provides panel unit root tests which work well even for a small N but large T panels by utilizing the pooled recursive mean adjusted feasible generalized least squares (RMA-FGLS) estimator. The tests don't require the estimation the number of common factors, and also don't require a particular factor structure such as (1). The proposed panel unit root tests don't suffer from serious size distortion as N increases because RMA methods reduce small sample bias both under the null and the alternative.

The paper is organized as follows. Section 2 consists of two subsections. The first subsection explains how RMA can reduce small sample bias significantly for the case of unknown mean and suggests a new detrending method using the RMA principle. Based on the proposed RMA methods, univariate RMA unit root tests are developed in the second subsection. Moreover, to achieve more power, covariate RMA unit root tests are also developed. Asymptotic local power analyses with various unit root tests also are provided. Section 3 proposes new panel unit root tests and deals with how to construct a consistent covariance and variance matrix of regression errors both under the null and alternative. Section 4 reports results of Monte Carlo studies. Section 5 provides some practical guidelines and an empirical example of long run purchasing power parity. Section 6 concludes.

2 Common Recursive Mean Adjustment

This section justifies the use of the common recursive mean adjustment to reduce the small sample bias of autoregressive coefficients. The autoregression models considered in the paper fall into the following two categories:

$$\textbf{M1: (Unknown Constant)} \begin{cases} y_t = a\,(1-\rho) + \rho y_{t-1} + \varepsilon_t \\ y_t = a + x_t,\;\; x_t = \rho x_{t-1} + \varepsilon_t \end{cases}$$

$$\textbf{M2: (Linear Trends)} \begin{cases} y_t = a\,(1-\rho) + b\rho + b\,(1-\rho)\,t + \rho y_{t-1} + \varepsilon_t \\ y_t = a + bt + x_t,\;\; x_t = \rho x_{t-1} + \varepsilon_t \end{cases}$$

The regression error ε_t is covariance stationary. For the unit root case, the initialization of x_t is taken to be $x_{i0} = O_p(1)$ and is uncorrelated with $\{\varepsilon_t\}_{t\geq 1}$.

2.1 Principle of Common Recursive Demeaning and Detrending

Common recursive mean adjustment for fixed effects: Transform M1 as

$$y_t - c_{t-1} = a(1-\rho) + \rho(y_{t-1} - c_{t-1}) - (1-\rho)c_{t-1} + u_t,$$

and find c_{t-1} which satisfies the following conditions.

1. $\mathrm{E}c_{t-1}u_t = 0$
2. $\mathrm{E}c_{t-1} = a$
3. $\mathrm{E}\sum_{t=2}^{T}(y_{t-1} - c_{t-1})(c_{t-1} - \mathrm{E}c_{t-1}) < O(T)$.

If c_{t-1} satisfies all three conditions, the following transformed regression will reduce the small sample bias significantly

$$y_t - c_{t-1} = \rho(y_t - c_{t-1}) + e_t, \quad e_t = -(1-\rho)(c_{t-1} - \mathrm{E}c_{t-1}) + u_t \tag{2}$$

Under the null hypothesis of $\rho = 1$, it is easy to see $e_t = u_t$ so that the third condition is not required under the null. Several candidates for c_{t-1} satisfy the first two conditions. For example, the overall time series mean of y_{t-1} satisfies the first two conditions. That is, $\mathrm{E}\left(\frac{1}{T}\sum y_{t-1}\right)u_t = 0$ and $\mathrm{E}\left(\frac{1}{T}\sum y_{t-1}\right) = a$. However, the third condition is not satisfied at all.

A second candidate is found in BN. They suggest one uses $y_1 = c_{t-1}$. Note that $\mathrm{E}y_1\epsilon_t = 0$ and also $\mathrm{E}y_1 = a$. However under the alternative, the third condition is not satisfied by y_1. To see this

$$\mathrm{E}\sum_{t=2}^{T}(y_{t-1} - y_1)x_1 = \mathrm{E}\sum_{t=2}^{T}(x_{t-1} - x_1)x_1 = \sigma_x^2\left[\frac{1-\rho^T}{1-\rho} - T - 1\right] = O(T)$$

As $T \to \infty$, the probability limit of $\hat{\rho}$ in (2) is given by

$$\mathrm{plim}_{T\to\infty}\hat{\rho} = \frac{1+\rho}{2} < 1 \text{ for } \rho < 1 \tag{3}$$

The most successful candidate known by the author is the common recursive mean. So and Shin (1999) originally introduced the recursive mean adjustment in univariate autoregressions to reduce the small sample bias of least square estimators. Later Shin and So (2001) extended their recursive mean adjustment to a unit root test for the case of an unknown mean. Choi, Mark and Sul (2005) extend the univariate recursive mean adjustment method to the panel context. Define the recursive mean as

$$c_{t-1} = \frac{1}{t-1}\sum_{s=1}^{t-1} y_s.$$

It is easy to see the first two conditions are satisfied with $\frac{1}{t-1}\sum_{s=1}^{t-1} y_s$. The third condition is satified given that

$$\mathrm{E}\sum_{t=2}^{T}\left(x_{t-1}-\frac{1}{t-1}\sum_{s=1}^{t-1}x_s\right)\frac{1}{t-1}\sum_{s=1}^{t-1}x_s = O\left(\ln T\right) < O\left(T\right).$$

This condition provides the consistency of $\hat{\rho}^{rc}$ which is the point estimate in (2) with $c_{t-1} = (t-1)^{-1}\sum_{s=1}^{t-1} y_s$.

This principle, however, cannot directly apply to the case of a linear trend. Chang, Park and Phillip (2001) propose a recursive detrending method to make the regression error become a martingale difference sequence. Their detrending estimator suffers from a serious upward bias when $\rho < 1$.

Recursive detrending adjustment: The common mean adjustment principle does not work for M2. Intuitively the two nuisance parameters a and b cannot be eliminated by the use of a common mean adjustment. Several detrending methods have been suggested (for example, see Taylor (2002), So and Shin (1999), Chang, Park and Phillips (2001)), but none of them reduce the small sample bias significantly when $\rho < 1$.[1]. Here we give an example of BN's detrending method.

$$y_t^* = \rho y_{t-1}^* + e_t \tag{4}$$

where

$$y_t^* = y_t - y_1 - \frac{(y_T - y_1)}{T-1}(t-1)$$

and

$$e_t = -(1-\rho)x_1 - (1-\rho)\frac{x_T - x_1}{T-1}(t-1) - \rho\frac{(x_T - x_1)}{T-1} + u_t.$$

Hence when $\rho < 1$, the probability limit of $\hat{\rho}$ in (4) is given by

$$\operatorname{plim}_{T\to\infty}(\hat{\rho}-\rho) = \frac{4}{6}(1-\rho) > 0 \text{ for } \rho < 1. \tag{5}$$

Here we provide a new detrending method which reduces the small sample bias significantly. First, define the common recursive mean adjustment as

$$d_{t-1} = \frac{2}{t-1}\sum_{s=1}^{t-1} y_s = 2a + b(t-1) + 2\bar{x}_{t-1},$$

where $\bar{x}_{t-1} = (t-1)^{-1}\sum_{s=1}^{t-1} x_s$. Observe this

$$y_t - d_{t-1} = -a(1-\rho) + b\rho + \rho(y_{t-1} - d_{t-1}) - 2(1-\rho)\bar{x}_{t-1} + u_t,$$

[1]Sul, Phillips and Choi (2005) prove that So and Shin (1999) detrending method cannot eliminate a trend coefficient.

since

$$y_{t-1} - d_{t-1} = -a + (x_{t-1} - 2\bar{x}_{t-1}).$$

The trend is eliminated but the constant is still present. Taking an overall mean adjustment yields

$$y_t - \bar{y} - 2(\bar{y}_{t-1} - \mu) = \rho[y_{t-1} - \bar{y}_{-1} - 2(\bar{y}_{t-1} - \mu)] + (e_t - \bar{e})$$

where $\mu = T^{-1}\sum \bar{y}_{t-1}$, $\bar{y}_{-1} = T^{-1}\sum y_{t-1}$ and $e_t = -2(1-\rho)\bar{x}_{t-1} + u_t$. This procedure reduces the small sample bias significantly.

Table 1 shows the dramatic bias reduction using the recursive mean adjustment. It is worthwhile noting that the variance of the recursive mean adjusted (RMA) estimator is far less than that of the OLS estimator, especially for the case of the linear trend. we also investigated whether or not the recursive mean adjustment works with a general AR(p) specification by means of Monte Carlo simulation and found that the proposed new estimator works very well.[2] In the next section, we provide an explicit bias formulae for RMA estimators.

2.2 Recursive Mean Adjusted Unit Root Tests

As Table 1 revealed, the relative variance of the RMA estimator compared to that of the OLS estimator decreases as T increases. This useful fact can be used for testing unit roots. Shin and So (2001) already proposed a univariate unit root test based on recursive mean adjustment under M1. Here we complete their task by adding the case of a linear trend. Consider a modified AR(p) model given by

$$\begin{array}{rr} y_t - c_{t-1} = \rho(y_{t-1} - c_{t-1}) + \sum_{j=1}^{p}\phi_j \Delta y_{t-j} + u_t & \text{for constant} \\ y_t - 2c_{t-1} = \beta + \rho(y_{t-1} - 2c_{t-1}) + \sum_{j=1}^{p}\phi_j \Delta y_{t-j} + u_t & \text{for linear trend} \end{array} \tag{6}$$

and denote

$$t^{rc} = \frac{\hat{\rho}^{rc}}{\sqrt{V(\hat{\rho}^{rc})}} \text{ and } t^{r\tau} = \frac{\hat{\rho}^{r\tau}}{\sqrt{V(\hat{\rho}^{r\tau})}} \tag{7}$$

where $\hat{\rho}^{rc}$ and $\hat{\rho}^{r\tau}$ are point estimates in (6). Let $\rho = 1 + c/T$.

Proposition 1 *(without covariate) The limiting distribution of the test statistics are given by*

$$t^{rc} \xrightarrow{d} A + cB, \quad t^{r\tau} \xrightarrow{d} C + cD$$

[2] To save space, the simulation results are not reported here but will be available upon request of the author.

where

$$
\begin{aligned}
A &= \left[\int_0^1 J^c dW\right]\left[\int_0^1 (J^c)^2\, dr\right]^{-\frac{1}{2}},\ C = \left[\int_0^1 J^\tau dW\right]\left[\int_0^1 (J^\tau)^2\, dr\right]^{-\frac{1}{2}} \\
B &= \Phi\left\{\left[\int_0^1 (J^c)^2\, dr\right]^{\frac{1}{2}} + \left[\int_0^1 \tilde{J}^c dr\right]\left[\int_0^1 (J^c)^2\, dr\right]^{-\frac{1}{2}}\right\} \\
D &= \Phi\left\{\left[\int_0^1 (J^\tau)^2\, dr\right]^{\frac{1}{2}} + \left[\int_0^1 \tilde{J}^\tau dr\right]\left[\int_0^1 (J^\tau)^2\, dr\right]^{-\frac{1}{2}}\right\}
\end{aligned}
$$

and $\Phi = \left(1 - \sum_{i=1}^{p-1}\phi_i\right)^{-1}$, $J = J(r) = \int_0^r e^{c(r-s)} dW(s)$, $\bar{J} = \bar{J}(r) = r^{-1}\int_0^r J(s)\, ds$, $\dot{J} = 2\bar{J} - \int_0^1 J(s)\, ds + 2\int_0^1 \bar{J} dr$, $J^c = J - \bar{J}$, $J^\tau = J - \dot{J}$, $\tilde{J}^c = J^c \bar{J}$, $\tilde{J}^\tau = J^\tau\left(\bar{J} - \int_0^1 \bar{J} dr\right)$ *and* $J(r)$ *is the Ornstein-Uhlenbeck process.*

The proof of proposition 1 is straightforward, hence it is omitted. The result for the linear trend case is new. The critical value for RMA unit root test can be obtained by letting $c = 0$ and replacing J by the standard Brownian motion W. The critical values are reported in Table 2. In contrast to the DF critical values, those for the RMA tests are less time variant. Practitioners can use -1.88 for the case of constant and -1.86 for the linear trend case, respectively, regardless of the size of T. For small T, the use of such critical values yields a slight size distortion.

When a covariate is available, the principle of RMA can be directly applied to obtain more power. Following Hansen (1995), consider the following covariate augmented DF, CADF(p, q_1, q_2) regressions for the unknown constant:

$$
y_t = \alpha + \rho y_{t-1} + \sum_{j=1}^{p}\phi_j \Delta y_{t-j} + \sum_{j=0}^{q_1}\phi_j^- \Delta x_{t-j} + \sum_{j=1}^{q_2}\phi_j^+ \Delta x_{t+j} + u_t
$$

Define

$$
v_t = \sum_{j=0}^{q_1}\phi_j^- \Delta x_{t-j} + \sum_{j=1}^{q_2}\phi_j^+ \Delta x_{t+j} + u_t,
$$

$$
\Omega = \sum_{l=-\infty}^{\infty} E\left[\begin{pmatrix} v_t \\ u_t \end{pmatrix}\begin{pmatrix} v_{t-k} & u_{t-k} \end{pmatrix}\right] = \begin{pmatrix} \sigma_v^2 & \sigma_{vu} \\ \sigma_{vu} & \sigma_u^2 \end{pmatrix}
$$

and $\lambda^2 = \sigma_{vu}^2\left[\sigma_u^2\sigma_v^2\right]^{-1}$ and $R^2 = \sigma_u^2/\sigma_v^2$.

The modified covariate augmented DF regression by using RMA is given by

$$
\begin{aligned}
y_t - c_{t-1} &= \rho(y_{t-1} - c_{t-1}) + \textstyle\sum_{j=1}^{p}\phi_j \Delta y_{t-j} + \sum_{j=-q_2}^{q_1}\psi_j \Delta x_{t-j} + u_t && \text{for M1} \\
y_t - 2c_{t-1} &= \beta + \rho(y_{t-1} - c_{t-1}) + \textstyle\sum_{j=1}^{p}\phi_j \Delta y_{t-j} + \sum_{j=-q_2}^{q_1}\psi_j \Delta x_{t-j} + u_t && \text{for M2}
\end{aligned}
\tag{8}
$$

The covariate RMA (CRMA) test statistics are defined as

$$
t_{\text{crma}}^{rc} = \left(\hat{\rho}_{\text{crma}}^{rc} - 1\right)/\sqrt{V\left(\hat{\rho}_{\text{crma}}^{rc}\right)},\ t_{\text{crma}}^{r\tau} = \left(\hat{\rho}_{\text{crma}}^{r\tau} - 1\right)/\sqrt{V\left(\hat{\rho}_{\text{crma}}^{r\tau}\right)}.
$$

where ρ^{rc}_{crma} and $\rho^{r\tau}_{\text{crma}}$ are point estimates in (8) for M1 and M2, respectively. Let $\rho = 1 + c/T$.

Proposition 2 *(CRMA test) The limiting distribution of the test statistics are given by*

$$t^{rc}_{crma} \xrightarrow{d} \lambda A + \frac{c}{R}B + \left(1-\lambda^2\right)^{1/2} N\left(0,1\right),$$
$$t^{r\tau}_{crma} \xrightarrow{d} \lambda C + \frac{c}{R}D + \left(1-\lambda^2\right)^{1/2} N\left(0,1\right)$$

The proof of Proposition 2 is obvious and hence it is omitted. When $\rho = 1$, the limiting distribution is dependent on the nuisance parameter λ. In contrast to the original CADF test proposed by Hansen, the 5% critical values for the RMA unit root test is not far from the 5% critical value for normal distribution. For example, the 5% critical values of the RMA unit root tests for unknown constant and linear trend cases with $T = 150$ are given by -1.88 and -1.86, respectively, which are equivalent to the 5% critical values for $\lambda = 1$ but to the 3% critical values for $\lambda = 0$.

Figure 1 shows the asymptotic local power of five unit root tests for M1: DF, Elliott, Rothenberg and Stock (1996)'s DFGLS, Hansen (1995)'s CADF, RMA and CRMA tests. The initial value u_1 is assumed to be distributed as $N\left(0,1\right)$. For the covariate tests, we set $R^2 = \lambda^2 = 0.8$ which is the case of a weak covariate. As $R^2 \to 0$, the asymptotic local power of the CADF and CRMA tests approaches 1. Note that DFGLS test becomes the point optimal test when $u_1 = 0$. However, Müeller and Elliott (2003) point out that the DFGLS test is not optimal when $u_1 - N\left(0,1\right)$. When $c << 0$, the asymptotic local power of the DFGLS test is lower than that of the DF test. Figure 2 shows the asymptotic local power for M2. The RMA test is worse than the DFGLS test when $u_1 = 0$. However, again, the CRMA test provides the best power even with a weak covariate.

3 Panel unit root tests under cross section dependence

Recently several panel unit root tests under cross section dependence have been proposed. Most of them are assumed that the cross section dependence arises from unknown common factors. Forni, Hallin, Lippi and Reichlin (2000) and Bai and Ng (2002,4) assume that panel data $\{y_{it}\}$ follow (1) while MP and PS assume that $\{y_{it}\}$ follow a panel AR(p) process but that their innovations follow (1).

There are at least two methods to handle or eliminate cross section dependence. The first method is rather direct. BN suggest to estimate F_{st} and defactor F_{st} from y_{it}. MP and PS suggest to estimate λ_{is} and orthogonalize λ_{is} by transforming y_{it}. This method is useful especially when $N > T$, and is efficient when the cross section dependence can be modelled by (1). The problem of this method is, however, that it requires knowledge of the number of factors, K. Bai and Ng (2002) and MP suggest various criteria to select the optimal number of K, but they confess that if $N < 20$, the number of common factors is difficult to estimate. Naturally if the number of common factors is poorly estimated, the panel unit root tests perform poorly.

However, for small N but relatively large T, the cross section dependence can be asymptotically handled well by employing panel feasible generalized least squares (PFGLS) estimation. The foremost merit of the PFGLS estimator is that it does not require the cross section dependence to follow (1). Another merit is that the limiting distribution of PFGLS is free from cross section dependence as it is shown by PS. To see this, assume that $F_{st} = \rho F_{st-1} + f_{st}$ and $M_{it} = \rho M_{it-1} + m_{it}$. Then eq. (1) can be rewritten as

$$y_{it} = \alpha_i + \rho y_{it-1} + u_{it}, \quad u_{it} = \sum_{s=1}^{K} \lambda_{is} f_{st} + m_{it} \tag{9}$$

and let $\hat{\Sigma}_{u,\text{pfgls}}$ be a consistent estimator for the variance-covariance matrix of u_{it}, and let $\hat{\upsilon}_{ij}$ be the ith and jth element of $\hat{\Sigma}^{-1}_{u,\text{pfgls}}$. Then the PFGLS estimator is defined as

$$\hat{\rho}_{\text{PFGLS}} = \frac{\sum_{i=1}^{N} \sum_{j=i}^{N} \hat{\upsilon}_{ij} \sum_{t=1}^{T} \tilde{y}_{it-1} \tilde{y}_{jt}}{\sum_{i=1}^{N} \sum_{j=i}^{N} \hat{\upsilon}_{ij} \sum_{t=1}^{T} \tilde{y}^2_{it-1}}$$

where '$\sim$' implies the demeaned time series. The limiting distribution of $\hat{\rho}_{\text{PFGLS}}$ under the null of a panel unit root is given by

$$T\left(\hat{\rho}_{\text{PFGLS}} - 1\right) \xrightarrow{d} \left(\sum_{i=1}^{N} \int_0^1 \tilde{W}_i dW_i\right) \left(\sum_{i=1}^{N} \int_0^1 \tilde{W}_i^2\right)^{-1} \quad \text{as } T \to \infty \tag{10}$$

where $\tilde{W}_i = W_i - \int W_i$ and W is a standard Brownian motion.

However, for finite T, if $\hat{\Sigma}_{u,\text{pfgls}}$ is biased then $\hat{\rho}_{\text{PFGLS}}$ also becomes biased. For example, the sample covariance matrix, $\hat{\Sigma}_u$, constructed by least squares dummy variable (LSDV) residuals is biased. The LSDV residuals are given by

$$\hat{u}_{it} = u_{it} + \left(1 - \hat{\rho}_{\text{LSDV}}\right) \tilde{y}_{it}$$

Hence the estimator for the ith and jth element σ_{ij} for the covariance matrix Σ_ϵ is given by

$$\hat{\sigma}_{ij,\text{lsdv}} = \frac{1}{T} \sum \hat{\epsilon}_{it} \hat{\epsilon}_{jt} = \sigma_{ij} + O_p\left(T^{-1}\right)$$

since $\left(\hat{\rho}_{\text{LSDV}} - 1\right)$ is $O_p\left(T^{-1}\right)$. Hence for finite T, the panel FGLS test suffers from a size distortion. More importantly, the size distortion increases as N increases. we will revisit this issue later in detail.

This small sample bias problem also causes difficulty in estimating the factor loading coefficients of λ_{is} even when N is large. MP and PS panel unit root tests require an orthogonalization procedure which eliminates the common factors. Their estimates for λ_{is}, however, depend on $\hat{\Sigma}_{u,\text{pfgls}}$. When the estimated $\hat{\Sigma}_{u,\text{pfgls}}$ is biased, the estimates of λ_{is} suffer from bias also. To avoid this problem, PS impose the null hypothesis of $\rho_i = 1$ to calculate $\hat{\Sigma}_{u,\text{pfgls}}$. This restriction reduces the size distortion but makes the test inconsistent under the alternative, which results in poor power of the test.

In the next subsection we show how the recursive mean adjustment method reduces the small sample bias. To do so, we assume Σ_u is known and we develop N asymptotic theory for the pooled recursive mean adjusted generalized least squares estimator. Next, we shall discuss how this N asymptotic theory can be used in practice.

3.1 Pooled recursive mean adjusted generalized least squares (RMA-GLS) estimator

Consider the simple panel AR(1) models given by

$$\begin{aligned} y_{it} - c_{it-1} &= \rho\left(y_{it-1} - c_{it-1}\right) + u_{it} && \text{for constant} \\ y_{it} - 2c_{it-1} &= \beta_i + \rho\left(y_{it-1} - 2c_{it-1}\right) + u_{it} && \text{for linear trend} \end{aligned} \tag{11}$$

where $u_{it} - N\left(0, \Sigma_u\right)$ and the off-diagonal terms of Σ_u are not equal to zero. In the end of this subsection, we discuss the general AR(p) case. Here we assume that the covariance and variance matrix Σ_u is known. In the next section, the feasible generalized least squares estimator will be discussed. Let $\mathbf{y}_t = \left(y_{1t}, ..., y_{Nt}\right)$, $\mathbf{c}_{t-1} = \left(c_{1t}, ..., c_{Nt}\right)$, $\mathbf{u}_t = \left(u_{1t}, ..., u_{Nt}\right)$, and $\Sigma_u^{-1} = \Lambda'\Lambda$. Now denote the transformed vector $\mathbf{y}_t^+ = \mathbf{y}_t\Lambda'$, $\mathbf{c}_{t-1}^+ = \mathbf{c}_{t-1}\Lambda'$. Furthermore, let y_{it}^+, c_{it-1}^+ and u_{it}^+ denote the ith elements of $\mathbf{y}_t^+$, $\mathbf{c}_{t-1}^+$ and $\mathbf{u}_t^+$, respectively.

$$\begin{aligned} y_{it}^+ - c_{it-1}^+ &= \rho\left(y_{it-1}^+ - c_{it-1}^+\right) + u_{it}^+ && \text{for constant} \\ y_{it}^+ - 2c_{it-1}^+ &= \beta_i + \rho\left(y_{it-1}^+ - 2c_{it-1}^+\right) + u_{it}^+ && \text{for linear trend} \end{aligned} \tag{12}$$

Note that u_{it}^+ is not cross sectionally dependent. The estimator in (12) is called the pooled RMA-GLS estimator. Following Harris and Tzavalis (1999), as $N \to \infty$ for fixed T, we have

Proposition 3 *The probability limit of the pooled recursive mean adjusted estimator under the null hypothesis of a panel unit root is given by*

$$plim_{N\to\infty}\left(\hat{\rho}_{pgls}^{rk} - 1\right) = 0,$$

and

$$\frac{\hat{\rho}_{pgls}^{rk} - 1}{\sqrt{Var\left(\hat{\rho}_{pgls}^{rk}\right)}} \xrightarrow{d} N\left(0, 1\right)$$

where $k = c$ or τ for constant and linear trend, respectively.

Appendix A provides the proof of Proposition 3 but it is instructive to sketch its outline here. For fixed effects, it is easy to see why the pooled RMA estimator is consistent. Under the null, we have

$$\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it-1}^{+}u_{it}^{+}=\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\left(\frac{1}{t-1}\sum_{s=1}^{t-1}y_{is}^{+}\right)u_{it}^{+}\right]=0$$

since $Ey_{it-j}^{+}u_{it}^{+}=0$ for $j>0$ as long as u_{it}^{+} is not serially correlated. For the linear trend case, note that

$$\begin{aligned}&\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=1}^{T}y_{it-1}^{+}\right)\left(\sum_{t=1}^{T}u_{it}^{+}\right)\\ =\ &2\times\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=1}^{T}\left(\frac{1}{t-1}\sum_{s=1}^{t-1}y_{is}^{+}\right)\right)\left(\sum_{t=1}^{T}u_{it}^{+}\right)\end{aligned}$$

as long as u_{it}^{+} is not serially correlated.

It is worthwhile noting that the pooled RMA estimator is more efficient than the pooled mean unbiased estimator proposed by Harris and Tzavalis (1999) and Phillips and Sul (2004). Let $\hat{\rho}^{c}$ be the pooled mean unbiased estimator. That is

$$\hat{\rho}^{c}=\hat{\rho}+bias\,(1,T)$$

where $\hat{\rho}$ is the LSDV estimator and $bias\,(1,T)$ is the mean bias function provided by Harris and Tzavalis (1999) and Phillips and Sul (2004). For the case of an unknown constant, $bias\,(1,T)=3/T$ while for the case of a linear trend, it becomes $7.5/T$ with moderately large T. The variance of the mean unbiased estimator $\hat{\rho}^{c}$ is larger than the variance of $\hat{\rho}^{pk}$ asymptotically. Harris and Tzavalis (1999) provide the asymptotic variance of $\hat{\rho}^{c}$ under the assumption of normality in the error ε_{it},given by

$$\text{plim}_{N\to\infty}V\left(\hat{\rho}^{c}\right)=\begin{cases}\frac{3\left(17T^{2}+14T+14\right)}{5T(T+2)^{3}} & \text{for M1}\\ \frac{15\left(193T^{2}-728T+1147\right)}{112(T+2)^{3}(T-2)} & \text{for M2}\end{cases}$$

As $T\to\infty$, the variance ratio is given by

$$\lim_{T\to\infty}\left[\text{plim}_{N\to\infty}\frac{V\left(\hat{\rho}^{c}\right)}{V\left(\hat{\rho}^{rk}\right)}\right]=\begin{cases}1.17 & \text{for M1}\\ 2.87 & \text{for M2}\end{cases}$$

For small T, the exact asymptotic variance ratios for the panel AR(1) model are plotted in Figure 3. Even when T is small (say $T>15$), the pooled RMA estimator is more efficient than the pooled mean unbiased estimator for both the fixed effects and the incidental linear trend cases.

For panel AR(p) regressions, proposition 3 still holds for M1 but not for M2. The pooled RMA estimator for the incidental linear trend becomes inconsistent for the panel AR(p) case when p> 1. However, as T increases, the inconsistency dissapates very rapidly. Table 3 reports the simulation results for the AR(2) case with an incidental linear trend. The data generating process is given by $\Delta y_{it} = \phi \Delta y_{it-1} + \Delta u_{it}$ where $u_{it} - iidN(0, I_N)$. We set $N = 1000$ and consider two values of $\phi = 0.4$ and -0.4. Note that the pooled RMA estimator suffers from an upward bias while the LSDV estimator suffers from a downward bias with small T. The upward bias of the pooled RMA estimator dissappears rapidly as T increases. Note that column D shows the MSE ratio of the pooled RMA to the LSDV estimator. Hence multiplying column D by column C yields the actual MSE for the pooled RMA estimator.

3.2 Panel RMA-FGLS Tests

In practice, Σ_u is unknown and should be estimated. Moreover proposition 3 does not hold any longer with unknown Σ_u. Under cross section dependence, the pooled recursive mean adjustment estimator is slightly biased. Assume the cross section dependence follows (9). The probability limit of the bias of the pooled RMA estimator for fixed effects is then given by

$$\text{plim}_{N\to\infty} (\hat{\rho}^{rc} - 1) = \vartheta\eta + \eta(g_{FT} - Eg_{FT}) + o_p(T^{-1})$$

where $g_{FT} = \left[\sum_{t=1}^T (F_{t-1} - \bar{F}_{t-1})(F_t - \bar{F}_{t-1})\right]\left[\sum_{t=1}^T (F_{t-1} - \bar{F}_{t-1})^2\right]^{-1}$, $F_t = \sum_{j=0}^\infty \rho^j f_{t-j}$, $\eta = \sigma_f^2\mu_\lambda^2 \left(\sigma^2 + \sigma_f^2\mu_\lambda^2\right)^2$, $\mu_\lambda^2 = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \lambda_i^2$, $\sigma_f^2 = V(f_t)$, and ϑ is the bias of the univariate RMA estimator reported in Table 1. Note that η represents the degree of cross section dependence. Since ϑ is inconsequential, the N asymptotic bias of the pooled RMA estimator also becomes tiny but random. Hence as T increases, the bias of the pooled RMA estimator dissapates very rapidly.

Define $\hat{\Sigma}_{u,\text{prma}}$ as the pooled RMA estimate of Σ_u. After replacing Σ_u by $\hat{\Sigma}_{u,\text{prma}}$ in (12). Let $t^{rk} = \frac{\hat{\rho}^{rk}_{\text{rma-fgls}} - 1}{\sqrt{V(\hat{\rho}^{rk}_{\text{rma-fgls}})}}$ where $k = c$ for fixed effects and τ for incidental linear trends. The resulting estimator becomes the pooled RMA-FGLS estimator. For fixed N, as $T \to \infty$, the limiting distribution of the RMA-FGLS tests are given by

$$t^{rk}_{\text{rma-fgls}} \underset{d}{\to} \left(\sum_{i=1}^N \int_0^1 W_i^k(r)^2 dr\right)^{-1/2} \sum_{i=1}^N \int_0^1 W_i^k(r)\, dW(r), \text{ for } k = c, \tau \tag{13}$$

where $W_i^c(r) = W_i(r) - \frac{1}{r}\int_0^r W_i(s)\,ds$, $W_i^\tau(r) = W_i(r) - \frac{2}{r}\int_0^r W_i(s)\,ds - \int_0^1 W_i(s)\,ds + 2\int_0^1 \frac{1}{r}\left(\int_0^r W_i(s)\,ds\right) dr$.

The asymptotic local power for the pooled RMA-FGLS tests for M1 and M2 are shown in Figure 4 and 5, respectively. Even with $N = 2$, the local power increases dramatically. As N increases further, the local power of the pooled RMA tests increase but at a decreasing rate.

When $\rho_i \neq \rho$, the pooled RMA-FGLS estimator becomes inconsistent even when $T \to \infty$. There are two sources of this inconsistency. To see this, let $\rho_i = \rho + \zeta_i$ where $\zeta_i - iid\left(0, \sigma_\zeta^2\right)$. The first source arises from the inconsistency of the pooled RMA estimator under heterogeneity of ρ_i. As $T \to \infty$ with fixed N, the inconsistency is given by

$$\text{plim}_{T\to\infty}\left(\hat{\rho}_{\text{prma}} - \rho\right) = \frac{N^{-1}\sum_{i=1}^{N}\left[\zeta_i/\left(1-\rho_i^2\right)\right]}{N^{-1}\sum_{i=1}^{N}\left[1/\left(1-\rho_i^2\right)\right]} > 0. \tag{14}$$

Hence the pooled RMA residuals carry this inconsistency so that the estimate of $\hat{\Sigma}_{u,\text{prma}}$ is also inconsistent. The second source of the inconsistency under the alternative arises from the use of the inconsistent estimate of $\hat{\Sigma}_{u,\text{prma}}$ to obtain the pooled RMA-FGLS estimator. The second source of the inconsistency amplifies the first source of inconsistency. More importantly the power of the pooled RMA-FGLS test is seriously hampered by this inconsistency since the direction of the inconsistency is upward rather than downward.

Under such circumstances, the covariance matrix should be estimated without imposing the homogeneity restriction $\rho_i = \rho$. Since the univariate RMA estimators don't suffer from small sample biases for both the unknown constant and linear trend cases, the variance-covariance estimate $\hat{\Sigma}_u$ can be constructed from the univariate RMA residuals. Define $\hat{\Sigma}_{u,\text{rma}} = (T-1)^{-1}\sum_{t=2}^{T}\hat{\mathbf{u}}_t\hat{\mathbf{u}}_t'$ where $\hat{\mathbf{u}}_t = (\hat{u}_{1t}, ..., \hat{u}_{Nt})$ and $\hat{u}_{1t}$ is the univariate RMA residual from (11). The transformed series $\mathbf{y}_t^\dagger = \mathbf{y}_t\hat{\Lambda}_{\text{rma}}$ where $\hat{\Sigma}_{u,\text{rma}}^{-1} = \hat{\Lambda}_{\text{rma}}\hat{\Lambda}_{\text{rma}}'$ is cross sectionally independent as $T \to \infty$. Denote P_i as the individual p-value for the t-test for $y_{it}^\dagger$, where $y_{it}^\dagger$ is the ith element of the vector $\mathbf{y}_t^\dagger$. Then following Maddala and Wu (1999) and Choi (2001), the following statistic can be used to test the null hypothesis of $H_1 : \rho_i = 1$ for all i as $T \to \infty$.

$$P = -2\sum_{i=1}^{N}\ln(P_i) \to_d \chi_{2N}^2 \quad \text{for fixed } N \tag{15}$$

Note that the alternative hypothesis of the pooled RMA-FGLS tests is $\rho < 1$ while the alternative of the Meta test is $\rho_i < 1$ for some i. However the Meta test has one disadvantage: The p-values must be calculated. This is a big burden to practitioners.

Here we suggest a rather cunning way to increase the power of the test by using the pooled RMA-FGLS when $\rho_i \neq \rho$. As we discussed above, the pooled RMA-FGLS estimator becomes inconsistent when $\rho_i \neq \rho$. This inconsistency can be attenuated by using $\hat{\Sigma}_{u,\text{rma}}$ rather than $\hat{\Sigma}_{u,\text{prma}}$. Note that under the null hypothesis, $\hat{\Sigma}_{u,\text{rma}}^{-1} - \hat{\Sigma}_{u,\text{prma}}^{-1} = o_p(1)$. Hence the pooled RMA-FGLS tests based on $\hat{\Sigma}_{u,\text{rma}}$ have the same limiting distribution in (13). Under the alternative of $\rho_i \neq \rho$, $\hat{\Sigma}_{u,\text{rma}}$ is a consistent estimator of Σ_u. Hence the second source of inconsistency can be avoided by using $\hat{\Sigma}_{u,\text{rma}}$. Note that the pooled RMA-FGLS estimator based on $\hat{\Sigma}_{u,\text{rma}}$ is also inconsistent. To distinguish between the pooled RMA-FGLS based on $\hat{\Sigma}_{u,\text{rma}}$ from that based on $\hat{\Sigma}_{u,\text{prma}}$, we give a name "PRMA-FGLS" to the pooled RMA-FGLS based on $\hat{\Sigma}_{u,\text{rma}}$. Note that the null hypothesis of the PRMA-FGLS test is $H_1 : \rho < 1$.

Practitioners may want to know at this point which test rejects the null hypothesis of a panel unit root more often. It is not straightforward to compare the asymptotic local power between the Meta and the PRMA-FGLS tests since the alternatives are different across the two tests. However, It is obvious that when $\rho_i = \rho$, the PRMA-FGLS test must be more powerful than the Meta test. When $\rho_i \neq \rho$, the power of the PRMA-FGLS test must be decreasing since the pooled estimators suffers from asymptotic upward bias. Meanwhile the power of the Meta test statistic is invariant to the heterogeneity of c_i. To investigate this issue, we consider $N = 3$ by setting $c_1 = c, c_2 = c + d$ and $c_3 = c + 2d$. Hence the average of local to unity parameters becomes $c + d$. Table 4 reports the asymptotic local powers of the two tests. For the case of unknown constant (M1), as long as $d < 5$, the power of the PRMA-FGLS test is higher than that of Meta test. Moreover for linear trend case (M2), even when $d = 5$, the power of the PRMA-FGLS test is still higher than that of Meta test. When $T = 40$, the equivalent values of (ρ_1, ρ_2, ρ_3) for $d = 5$ and 3 are $(0.975, 0.85, 0.725)$ and $(0.975, 0.9, 0.825)$, respectively. This implies that as long as the degree of heterogeneity in ρ_i is not that significant, the PRMA-FGLS test usually rejects the null hypothesis of a panel unit root more often than the Meta test.

4 Monte Carlo Simulation

We consider three sets of simulations. In the first set, we compare the finite sample performance between the PRMA-FGLS and the pooled ADF-FGLS tests. The second set reports the finite sample performance between the PRMA-FGLS and BN's tests. The last set reports the finite sample performance of CRMA tests with cross sectional average data against that of individual CRMA tests. For all three cases, we consider $T \in [50, 100, 150, 200]$.

RMA-FGLS vs ADF-FGLS The data generating process is given by

$$\begin{aligned} y_{it} &= \rho_i y_{it-1} + u_{it}, \quad u_{it} = \begin{cases} \phi u_{it-1} + \epsilon_{it} & \text{for AR error} \\ \phi \epsilon_{it-1} + \epsilon_{it} & \text{for MA error} \end{cases} \\ \epsilon_{it} &= \lambda_i f_t + m_{it} \end{aligned}$$

where $m_{it} - iidN(0, 1)$, $f_t - iidN(0, 1)$, $\lambda_i - U[0, 5]$. The pooled RMA-FGLS tests are based on $\hat{\Sigma}_{urma}$ to improve the power of the tests under heterogeneity of ρ_i. Here we report only homogenous ρ case since the simulation results with heterogeneous ρ_i are very similar to those of the homogenous case. Various cases are studied to compare these two tests but only three cases are reported since the results are similar. The parameters values for each case are as follows: Case I: AR(1), $\phi = 0$; Case II: AR(2), $\phi \in [0.2, -0.2]$; Case III: ARMA(1,1), $\phi = -0.2$. For cases II and III, BIC criteria is used for the choice of the optimal lag. For all cases, we consider $N = (3, 5, 7, 10)$ and set $\rho = 0.95$ under the alternative. Tables 5, 6 and 7 report the simulation results. Usually there is a trade-off between the size and the power. However, the PRMA-FGLS tests reduce the size distortion but increase the power of the tests for all cases. Meanwhile the pooled ADF-FGLS tests suffer from a somewhat serious size distortion with small T and provide lower power of the tests than PRMA-FGLS tests.

RMA-FGLS vs Bai and Ng's Tests The data generating process is given by

$$y_{it} = \lambda_i F_t + M_{it},\ F_t = \phi F_{t-1} + f_t,\ \ M_{it} = \rho M_{it-1} + \varepsilon_{it}$$

where $\phi = \rho,\ v_t - iidN\left(0,1\right),\ \varepsilon_{it} - iidN\left(0,I_N\right),\ \lambda_i - U\left[0,5\right].$ Under the alternative, we set $\phi = \rho = 0.95,\ N = (5,10,15,20,30)$ and consider only the AR(1) case. The number of lag is assumed to be known. The number of common factors are estimated by IC_{p1} which is one of the best criteria suggested by Bai and Ng (2002). The maximum number of common factors is set to be 8. Table 8 reports the results. When N is small, as BN and MP find, BN's test suffers somewhat from a size distortion. However, the size distortion disappears very quickly as N increases. Meanwhile, the PRMA-FGLS test does not suffer from any size distortion as long as T is relatively large compared to N. For fixed effects, the power of BN's tests are lower than those of the PRMA-FGLS test when $N < 20$. Interestingly, for incidental linear trends, the power of the PRMA-FGLS test is higher even when $N \geq 20$.

Individual CRMA vs CRMA with cross sectional aggregated data The data generating process is given by

$$y_{it} = \lambda_i F_t + m_{it},\ F_t = \phi F_{t-1} + v_t,\ \ m_{it} = \rho m_{it-1} + u_{it}$$

$$v_t = e_t + w_t,\ e_t = 0.4 e_{t-1} + \varepsilon_t$$

where

$$\begin{pmatrix} w_t \\ e_t \end{pmatrix} - N\left(0, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right),$$

and $\rho = 0.95,\ v_t - iidN\left(0,1\right)$ and $u_{it} - iidN\left(0,I_N\right).$ We consider $N = [5,15]$ and set $\phi = 0.95$ under the alternative. The factor loading coefficient λ_i is drawn from a unifor distribution, viz. $\lambda_i - U\left(a,5\right)$. Two different lower bounds, a, are considered; 0 and 1. Since the variance of v_t is set to be unity, the size of individual test varies depending on the value of the lowest λ_i. As λ_i is lower, the size distortion for an individual CRMA test should increase. Meanwhile the cross section average eliminates the stationary idiosyncratic errors as $N \to \infty$ so that CRMA tests with the cross sectional average must not suffer from any size distortion regardless of the value of a.

The following regression is used to obtain CMRA test statistics

$$\begin{aligned} y_{it} &= a_i + b_i t + \rho_i y_{it-1} + \gamma_1 g_t + \gamma_2 g_{t-1} + e_{it} \\ \bar{y}_t &= a + bt + \rho \bar{y}_{t-1} + \gamma_1 g_t + \gamma_2 g_{t-1} + e_t \end{aligned}$$

Table 9 shows the results. To access the summary measure of an individual CRMA test, we use cross sectional average of the size and power of CRMA tests with individual time series data. When the lower bound of λ_i, a, includes zero (but is never zero), the size distortion of an individual CRMA test increases as T increases. Since we take the cross sectional mean as the summary measure of an individual CRMA test, as N increases, the size distortion appears to decrease. When the restriction $a \geq 1$ is imposed, this abnormal behavior disappears. However this is very strong restriction. In contrast to the individual CRMA test, the CRMA test with cross sectional average time series data does not suffer from any size distortion. Also the size adjusted power of the test for the cross sectional average is always higher than that of the individual CRMA test.

5 Practical Issues and Application

This section consists of three subsections. The first subsection provides step by step procedures of how to test for a panel unit root under cross section dependence. The second subsection discusses several practical issues. The third section illustrates an empirical example: testing for long run purchasing power parity.

5.1 Step by Step Procedure

Consider the following panel AR(p) models

$$y_{it} = \alpha_i + \beta_i t + \rho_i y_{it-1} + \sum_{j=1}^{p_i} \phi_{ij} \Delta y_{it-j} + u_{it}, \quad u_{it} = \sum_{s=1}^{K} \lambda_{is} F_{st} + m_{it}$$

where T is large but N is relatively small (say $N < 20$). The proposed test in the paper consists of two tests; the PRMA-FGLS tests and the covariate RMA tests. The first test is the PRMA-FGLS test which examines if the idiosyncratic errors are stationary or not. If one cannot reject the null hypothesis, (s)he does not need to proceed to the second test. Otherwise, one should perform the second test since the rejection of the panel unit root does not imply that y_{it} is stationary if the common factors to y_{it} are I(1). Since the number of common factors are unknown due to the small N, it is hard to extract the common factors precisely. However the cross sectional average of y_{it} provides a proxy of the cross sectional average of common factors since $\bar{y}_t = N^{-1} \sum_{i=1}^{N} y_{it} = \bar{F}_t + o_p\left(N^{-1/2}\right)$ where $\bar{F}_t = K^{-1} \sum_{s=1}^{K} F_{st}$. If one can find a covariate for y_{it}, then (s)he can use the covariate RMA (CRMA) test. If there is no available covariate, then apply the RMA or DFGLS test with $\bar{y}_t$. Here we provide step by step procedures for the PRMA-FGLS test.

Step 1: Run the following regression for each i.

$$\begin{array}{rll} y_{it} - c_{it-1} = \rho_i \left(y_{it-1} - c_{it-1}\right) + \sum_{j=1}^{p_i} \phi_{ij} \Delta y_{it-j} + u_{it} & \text{for constant} \\ y_{it} - 2c_{it-1} = \beta_i + \rho_i \left(y_{it-1} - 2c_{it-1}\right) + \sum_{j=1}^{p_i} \phi_{ij} \Delta y_{it-j} + u_{it} & \text{for linear trend} \end{array} \tag{16}$$

where $c_{it-1} = (t-1)^{-1} \sum_{s=1}^{t-1} y_{is}$. Define $\hat{\rho}_i^{rk}$ as the least squares estimator in (16) where $k = c$ or τ for constant and linear trend, respectively.

Step 2: Treat $\hat{\rho}_i^{rk}$ as if it were the true value of ρ_i. Run the following regression to obtain the estimates for ϕ_{ij} for each i. If $\hat{\rho}_i^{rk} > 1$, then set $\hat{\rho}_i^{rk} = 1$.

$$\begin{array}{rll} y_{it} - \hat{\rho}_i^{rk} y_{it-1} = a_i + \sum_{j=1}^{p} \phi_{ij} \Delta y_{it-j} + u_{it} & \text{for constant} \\ y_{it} - \hat{\rho}_i^{rk} y_{it-1} = a_i + \beta_i t + \sum_{j=1}^{p} \phi_{ij} \Delta y_{it-j} + u_{it} & \text{for linear trend} \end{array}$$

Construct the sample covariance and variance matrix, $\hat{\Sigma}_{u,\mathrm{rma}}$ where the ith and jth element of $\hat{\Sigma}_{u,\mathrm{rma}} = (T-p-1)^{-1} \sum_{t=1}^{T-p-1} \hat{u}_{it} \hat{u}_{jt}$.

Step 3: Run the following two projection regressions for each i and obtain the regression residuals.

$$\begin{array}{rl} \left.\begin{array}{r} y_{it} - c_{it-1} = \sum_{j=1}^{p} \varphi_{ij} \Delta y_{it-j} + \xi_{it}, \\ y_{it-1} - c_{it-1} = \sum_{j=1}^{p} \zeta_{ij} \Delta y_{it-j} + \xi_{it-1} \end{array}\right\} & \text{for constant} \\ \left.\begin{array}{r} y_{it} - 2c_{it-1} = g_i + \sum_{j=1}^{p} \varphi_{ij} \Delta y_{it-j} + \xi_{it} \\ y_{it-1} - 2c_{it-1} = g_i + \sum_{j=1}^{p} \zeta_{ij} \Delta y_{it-j} + \xi_{it-1} \end{array}\right\} & \text{for linear trend} \end{array}$$

Step 4: Define $\hat{\omega}_{ij}$ as the ith and jth element of $\hat{\Sigma}_{u,\mathrm{rma}}^{-1}$. Obtain the pooled FGLS estimator and construct t-statistic

$$\hat{\rho}_{\text{rma-fgls}} = \frac{\sum_{i=1}^{N} \sum_{j=i}^{N} \hat{\omega}_{ij} \sum_{t=1}^{T} \hat{\xi}_{it-1} \hat{\xi}_{jt}}{\sum_{i=1}^{N} \sum_{j=i}^{N} \hat{\omega}_{ij} \sum_{t=1}^{T} \hat{\xi}_{it-1}^2}, \quad t_N^{rk} = \frac{\hat{\rho}_{\text{rma-fgls}} - 1}{\sqrt{V\left(\hat{\rho}_{\text{rma-fgls}}\right)}} \tag{17}$$

where $V\left(\hat{\rho}_{\text{rma-fgls}}\right) = \sum_{i=1}^{N} \sum_{j=i}^{N} \hat{\omega}_{ij} \sum_{t=1}^{T} \hat{\xi}_{it-1}^2$.

5.2 Practical Issues

In this subsection we provide a couple of practical guidelines for the selection of panel data and of a covariate.

5.2.1 Use high frequency data if you can

As discussed in the previous section, the size distortion of panel unit root tests arises due to a small number of time series observations. One may think that the use of high frequency data can avoid this problem. For example, suppose that approximately 30 annual real exchange rates time series observations are available. At a quarterly frequency, the total number of time series observations becomes 120, and for a monthly frequency it becomes 360. Sounds like $T = 30$ is a small number while $T = 360$ is a large number. Hence one may think that the small sample bias with 30 annual observations is much more serious than that of monthly observations. Moreover, one may further think that 360 is not a small sample any more.

The above conjecture is not entirely true for univariate unit root tests since the bias of AR(1) coefficient increases as the degrees of frequency increases.[3] However it does work very well for the RMA tests. The bias of the pooled RMA estimator approaches zero as $T \to \infty$. Moreover, under heterogeneity in ρ_i, the upward bias of the PRMA-FGLS estimator becomes smaller as the degrees of the frequency increases. Table 10 shows the results. We set $N = 18$ which is a typical number of cross sectional units for empirical purchasing power parity studies. Evidently, for all ranges of ρs, using a monthly frequency yields much higher power of the tests.

5.2.2 How to choose a covariate

Practitioners may choose a covariate by estimating a long run correlation between a covariate and the regression error. Here are some guidelines for which a covariate, x_t, should not be chosen.

1. Avoid a covariate which is cointegrated with y_t. When x_t is cointegrated with y_t, then it is hard to distinguish if one is testing for a unit root or for cointegration. To see this, assume $y_t - I(1)$, $x_t - I(1)$ but $e_t = y_t - \beta x_t = I(0)$. That is, $y_t = \beta x_t + e_t$ where $e_t = \rho e_{t-1} + \epsilon_t$. Then the error correction representation exists such that
$$y_t = \rho y_{t-1} - \rho\beta x_{t-1} + \sum \phi_j^y \Delta y_{t-j} + \sum \phi_j^x \Delta x_{t-j} + h_t^y.$$
Since x_{t-1} also can be written as a function of lags of y_{t-1} and x_{t-1}, the error correction representation can be re-written as
$$y_t = \rho y_{t-1} + \sum \psi_j^y \Delta y_{t-j} + \sum \psi_j^x \Delta x_{t-j} + h_t$$
Hence the null hypothesis of H_0: $\rho = 1$ implies the null of no cointegration.

2. Avoid a stationary covariate. If x_t is stationary, then Δx_t is $I(-1)$. In this case, the limiting distribution of the CRMA test statistics become identical to that of the RMA test statistics. Hence make sure if x_t is I(0) or x_t is I(1).[4]

3. From 1 and 2, it is obvious that a variable x_t can be a good covariate if $e_t = y_t - \beta x_t$ is I(1) and $\beta \neq 0$. For example, a log real exchange rate, $q_t = s_t - (p_t - p_t^*)$ where s_t is a log nominal exchange rate while $(p_t - p_t^*)$ is a log relative price. Since $(p_t - p_t^*)$ and s_t are I(1), both of them are good candidates for covariates of q_t.

[3] See Choi and Chung (1995) for a detailed discussion.
[4] See Hansen (1995) for further detailed discussion.

5.3 Empirical Example

Testing long run purchasing power parity (PPP) is the most popular application for panel unit root tests. The most recent empirical evidence is summarized by Moon and Perron (2004b). They use 17 monthly bilateral real exchange rates from 1974 to 1998 and find rather mixed evidence for long run PPP. Since the number of common factors is hard to estimate with small N, Moon and Perron vary the number of common factors from 1 to 8. BN's test rejects the null of a panel unit root for up to 3 factors while other panel unit root tests don't reject the null of panel unit root at all.

Here we use 14 bilateral real exchange rates from the OECD main economic indicators during the period between January 1975 and March 2005. The numeraire currency is the U.S. dollar. Since CPI is not seasonally adjusted, a one year moving average is used for the seasonal adjustment. Out of 14 exchange rates, 8 rates come from EU countries. Hence the final panel includes only 7 bilateral exchange rates. Each exchange rate out of the 8 EU rates is included in each panel so that the total number of panels is 8. The number of lags is selected by BIC and the maximum lag is set to be 60. Table 11 reports the results. The PRMA-FGLS test rejects the null of panel unit root even at the 5% level for all panels. Also the CRMA test with cross sectionally averaged real exchange rates rejects the null. The cross sectional average of nominal spot exchange rates is used as the covariate for the real exchange rate.

6 Conclusion

This paper studies the principle of common recursive mean adjustment and proposes a new detrending method in dynamic panel models. By utilizing recursive mean adjustment, this paper provides three unit root tests: the RMA test, the CRMA and the PRMA-FGLS tests. The first two tests are designed for testing cross sectional averages of panel time series data to examine if common factors in a panel are stationary or not. The third test is designed to test if idiosyncratic errors are stationary or not. The proposed panel unit root test under cross section dependence is precise and powerful especially when T is relatively large compared to N.

The proposed tests are used to examine if real exchange rates are nonstationary and can reject the null of a panel unit root. However, the rejection of a panel unit root does not imply that all real exchange rates are stationary. The development of panel unit root tests under the null of stationarity is a worthwhile topic of future research.

References

[1] Bai, J. and S. Ng (2002). "Determining the number of factors in approximate factor models," *Econometrica*, 191-221

[2] Bai , J. and S. Ng (2004a). "A PANIC Attack on Unit Roots and Cointegration," *Econometrica,* 70:1, 191-221.

[3] Chang, Y. (2002). "Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency," *Journal of Econometrics*, 110, 261-292.

[4] Chang, Y., J.Y. Park and P.C.B. Phillips (2001), "Nonlinear econometric models with cointegrated and deterministically trending regressors," *The Econometrics Journal*, 4, 1—36.

[5] Choi, C.Y., N. C. Mark and D. Sul (2005), ""Bias Reduction by Recursive Mean Adjustment in Panel Data," *mimeo,* University of Auckland.

[6] Choi, I. (2001). "Unit Root Tests for Panel Data," *Journal of International Money and Finance*, 20, 249-272.

[7] Choi, I. (2005) "Unit Root Tests for Cross-Sectionally Correlated Panels," forthcoming in Econometric Theory and Practice: Frontiers of Analysis and Applied Research: Essays in Honor of Peter C.B. Phillips, Cambridge University Press.

[8] Choi, I. and B.S. Chung (1995) "Sampling Frequency and the Power of Unit Root Tests: A Simulation Study," *Economics Letters* 49, 131-136.

[9] Elliott, G., T.J. Rothenberg, and J.H. Stock (1996). "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813-836.

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

[11] Hansen B.E. (1995), "Rethinking the univariate approach to unit root tests: How to use covariates to increase power," *Econometric Theory*, 11, 1148-1171

[12] Im, K.S., M.H. Pesaran, and Y. Shin (2003). "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115, 53-74.

[13] Levin, A., F. Lin, and C. Chu (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 108, 1-24.

[14] Maddala, G.S. and S. Wu (1999). "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

[15] Harris, R.D.F. and E. Tzavalis (1999). "Inference for unit roots in dynamic panels where the time dimension is fixed," *Journal of Econometrics* 201–226.

[16] Moon, H. R. and B. Perron (2004a). "Testing for a Unit Root in Panels with Dynamic Factors," *Journal of Econometrics*, 122, 81-126.

[17] Moon, H.R. and B. Perron, (2004b) "An empirical analysis of nonstationarity in panels of exchange rates and interest rates with factors", mimeo, University of Southern California.

[18] Müeller, U and G. Elliott (2003),"Tests for Unit Roots and the Initial Condition", *Econometrica*, 71, 1269-86

[19] Phillips, P.C.B. and D. Sul (2003). "Dynamic panel estimation and homogeneity testing under cross section dependence," *The Econometrics Journal,* 6, 217-259.

[20] Phillips, P.C.B. and D. Sul (2004). "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence," mimeo, University of Auckland.

[21] So, B.S. and D.W. Shin (1999), "Recursive mean adjustment in time series inferences," *Statistics & Probability Letters*, 43, 65-73

[22] Shin, D.W. and B. S. So (2001). "Recursive mean adjustment and tests for unit roots," *Journal of Time Series Analysis*, 595–612.

[23] Sul, D, P.C.B. Phillips and C.Y. Choi (2005), "Prewhitening Bias in HAC Estimation" forthcoming in *Oxford Bulletin of Economics & Statistics.*

[24] Taylor R. (2002), "Regression-Based Unit Root Tests with Recursive Mean Adjustment for Seasonal and Non-Seasonal Time Series," 2002, Journal of Business and Economic Statistics 20, 269-281.

[25] Quah, D. (1994), "Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Panels," *Economics Letters* 44, 1–9.

7 Appendix

We restate the models and make the following assumptions.

M1: (Unknown Costant) $\begin{cases} y_{it} = a_i\left(1-\rho_i\right) + \rho_i y_{it-1} + u_{it} \\ y_{it} = a_i + x_{it},\ x_{it} = \rho_i x_{it-1} + u_{it} \end{cases}$

M2: (Linear Trends) $\begin{cases} y_{it} = a_i\left(1-\rho\right) + b_i\rho_i + b_i\left(1-\rho_i\right)t + \rho_i y_{it-1} + u_{it} \\ y_{it} = a_i + b_i t + x_{it},\ x_{it} = \rho x_{it-1} + u_{it} \end{cases}$

Assumption 1 *The u_{it} have zero mean, finite $2+2\nu$ moments for some $\nu > 0$, are independent over i and t with $E(u_{it}^2) = \sigma_i^2$ for all t, and $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \sigma_i^2 = \sigma^2$.*

Assumption 2 $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N a_i = \mu_a < \infty$, $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N b_i = \mu_b < \infty$, $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N a_i^2 = \mu_a^2 < \infty$ and $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N b_i^2 = \mu_b^2 < \infty$.

7.1 Appendix A: Proof of Proposition 3

For notational convenience, we delete the superscript '+' and just denote y_{it}^+ as y_{it}.

7.1.1 The case of unknown constant

Asymptotic Bias It is easy to see that

$$\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N\sum_{t=2}^T\left(y_{it-1}-\bar{y}_{it-1}\right)u_{it} = 0$$

Asymptotic Normality Let $\xi_{iT} = \sum_{t=2}^T\left(y_{it-1}-\bar{y}_{it-1}\right)u_{it}$. From Assumption 1, ξ_{iT} is an independent random series across i which has zero mean and constant variance, σ_ξ^2. Note that

$$\begin{aligned} \sigma_\xi^2 &= \text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N\left[\sum_{t=2}^T\left(y_{it-1}-\bar{y}_{it-1}\right)u_{it}\right]^2 \\ &= \sigma^2\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N\sum_{t=2}^T\left(y_{it-1}-\bar{y}_{it-1}\right)^2 = \sigma^2 Q_x, \quad \text{let say.} \end{aligned}$$

where

$$Q_x = \frac{\sigma^2}{6}\left(T-3\right)\left(T-1\right) + \frac{\sigma^2}{6}\sum_{t=1}^{T-1}\frac{1}{t}$$

The asymptotic probability limit of the denominator of the pooled RMA estimator is less than that of the LSDV estimator. That is,

$$\begin{aligned} & Q_x - \text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T}\left(x_{it-1}-x_{i\cdot-1}\right)^2 \\ = & -\frac{1}{3}T+\frac{1}{2}+\frac{1}{6}\ln\left(T-1\right)+\gamma<0 \text{ for all } t. \end{aligned}$$

where γ is a Euler constant which is approximately 0.56.

Applying the CLT, we have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\xi_{iT}}{\sigma_\xi}\xrightarrow{d}N\left(0,1\right)$$

Hence we have

$$\frac{\hat{\rho}^{rc}_{\text{rma-gls}}-1}{\sqrt{Var\left(\hat{\rho}^{rc}_{\text{rma-gls}}\right)}}\xrightarrow{d}N\left(0,1\right)$$

where the asymptotic limit of $Var\left(\hat{\rho}^{rc}_{\text{rma-gls}}\right)$ is given by

$$\text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}^2\left[\sum_{t=2}^{T}\left(y_{it-1}-\bar{y}_{it-1}\right)^2\right]^{-1}=\sigma^2Q_x^{-1}$$

7.1.2 The case of linear trend:

Asymptotic Bias The bias of $\hat{\rho}^{r\tau}_{\text{rma-gls}}$ is given by

$$\text{plim}_{N\to\infty}\left(\hat{\rho}^{r\tau}_{\text{rma-gls}}-1\right)=\text{plim}_{N\to\infty}\frac{C_{NT}}{D_{NT}}$$

where

$$D_{NT}=\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it-1}-y_{i\cdot-1}-2\left[\bar{y}_{it-1}-\bar{y}_{i\cdot-1}\right]\right)^2$$

and

$$C_{NT}=\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it-1}-y_{i\cdot-1}-2\left[\bar{y}_{it-1}-\bar{y}_{i\cdot-1}\right]\right)\left(u_{it}-u_{i\cdot}\right)$$

From Lemma 2. the probability limit of the numerator term is given by

$$\text{plim}_{N\to\infty}C_{NT}=C_{1T}-\frac{2}{T-1}C_{2T}$$

where

$$\begin{aligned} C_{1T} &= \text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T}\left(y_{it-1}-y_{i\cdot-1}\right)\left(u_{it}-u_{i\cdot}\right) \\ C_{2T} &= \text{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T}\left[\bar{y}_{it-1}-\bar{y}_{i\cdot-1}\right]\left(u_{it}-u_{i\cdot}\right) \end{aligned}$$

From direct calculation,

$$C_{1T} = -\frac{1}{2}T + 1$$

while

$$\begin{aligned}
C_{2T} &= E\left(x_{i1} + \frac{1}{2}(x_{i1} + x_{i2}) + \frac{1}{3}(x_{i1} + x_{i2} + x_{i3}) + ... + \frac{1}{T-1}\sum_{s=1}^{T-1} x_{is}\right)(u_{i2} + ... + u_{iT}) \\
&= E\left\{\frac{1}{2}(x_2 u_2) + \frac{1}{3}(x_2 u_2 + x_3(u_2 + u_3)) + ... \frac{1}{T-1}\left(\sum_{s=1}^{T-1} x_{is}\right)\left(\sum_{t=2}^{T} u_{it}\right)\right\} \\
&= \frac{1}{2}(1) + \frac{1}{3}(1+2) + \frac{1}{4}(1+2+3) + ... + \frac{1}{T-1}(1+2+...+T-2) \\
&= \sum_{t=1}^{T-1}\left(\frac{1}{t}\sum_{k=1}^{t-1} k\right) = \frac{1}{4}T^2 - \frac{3}{4}T + \frac{1}{2}
\end{aligned}$$

Hence

$$\mathrm{plim}_{N\to\infty} C_{NT} = -\frac{1}{2}T + 1 + 2\frac{1}{T-1}\left(\frac{1}{4}T^2 - \frac{3}{4}T + \frac{1}{2}\right) = 0$$

Asymptotic Normality Let $\zeta_{iT} = \sum_{t=1}^{T}(y_{it-1} - y_{i\cdot-1} - 2[\bar{y}_{it-1} - \bar{y}_{i\cdot-1}])(u_{it} - u_{i\cdot})$. The asymptotic variance of ζ_{it} is given by

$$\begin{aligned}
\sigma_\zeta^2 &= \mathrm{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left[\sum_{t=1}^{T}(y_{it-1} - y_{i\cdot-1} - 2[\bar{y}_{it-1} - \bar{y}_{i\cdot-1}])(u_{it} - u_{i\cdot})\right]^2 \\
&= \sigma^2 \mathrm{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it-1} - y_{i\cdot-1} - 2[\bar{y}_{it-1} - \bar{y}_{i\cdot-1}])^2 = \sigma^2 Q_z, \quad \text{let say.}
\end{aligned}$$

where

$$Q_z = q_{1T} + 4q_{2T} - 4q_{3T},$$

$$\begin{aligned}
q_{1T} &= \frac{1}{6}(T-2)T \\
q_{2T} &= \sum_{t=2}^{T}\left(\frac{1}{t-1}\right)^2\left(\frac{1}{6}t + \frac{1}{3}t^3 - \frac{1}{2}t^2\right) - \frac{1}{T-1}\sum_{j=1}^{T-1} j\left(\sum_{s=j}^{T-1}\frac{1}{s}\right)^2 \\
&\quad -\frac{2}{T-1}\sum_{k=2}^{T-1}\sum_{j=1}^{T-k} j\left(\sum_{s=j}^{T-1}\frac{1}{s}\right)\left(\sum_{s=j+(k-1)}^{T-1}\frac{1}{s}\right) \qquad (18) \\
q_{3T} &= -\frac{1}{4}T^2 + \frac{3}{4}T - \frac{1}{2} + \frac{1}{T-1}\left(\sum_{k=1}^{T-2}\sum_{j=1}^{T-k-1} j\sum_{s=k}^{T-1}\frac{1}{s}\right)
\end{aligned}$$

Applying the CLT, we have

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\zeta_{iT}}{\sigma_{\zeta}}\xrightarrow{d}N\left(0,1\right)$$

Next we have

$$\frac{\hat{\rho}_{\text{rma-gls}}^{r\tau}-1}{\sqrt{Var\left(\hat{\rho}_{\text{rma-gls}}^{r\tau}\right)}}\xrightarrow{d}N\left(0,1\right)$$

where the asymptotic limit of $Var\left(\hat{\rho}_{\text{rma-gls}}^{r\tau}\right)$ is given by

$$\text{plim}_{N\rightarrow\infty}\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}^{2}\left[\sum_{t=1}^{T}\left(y_{it-1}-y_{i\cdot-1}-2\left[\bar{y}_{it-1}-\bar{y}_{i\cdot-1}\right]\right)^{2}\right]^{-1}=\sigma^{2}Q_{z}^{-1}$$

QED.

Table 1: Bias, variance and mean square error (MSE) of RMA and OLS estimators.

	Constant				Linear Trend			
	(A)	(B)	(C)	(D)	(A)	(B)	(C)	(D)
ρ=0.92,T=30	-0.046	-0.144	0.97	0.51	-0.012	-0.254	0.68	0.28
ρ=0.94,T=30	-0.050	-0.149	0.95	0.49	-0.023	-0.265	0.63	0.26
ρ=0.96,T=30	-0.054	-0.155	0.93	0.47	-0.036	-0.277	0.60	0.25
ρ=0.98,T=30	-0.060	-0.161	0.93	0.46	-0.051	-0.292	0.56	0.24
ρ=1.00,T=30	-0.065	-0.167	0.92	0.44	-0.068	-0.309	0.52	0.24
ρ=1.00,T=50	-0.040	-0.104	0.85	0.43	-0.042	-0.193	0.45	0.22
ρ=1.00,T=70	-0.029	-0.075	0.82	0.42	-0.030	-0.141	0.40	0.20
ρ=1.00,T=100	-0.020	-0.053	0.81	0.41	-0.022	-0.100	0.39	0.19
ρ=1.00,T=200	-0.010	-0.027	0.77	0.41	-0.011	-0.051	0.36	0.18

Note: (A) =Bias of recursive mean adjusted estimator; (B) = Bias of OLS estimator; (C) = Variance ratio of recursive mean adjusted estimator to OLS estimator; (D) = MSE ratio of recursive mean adjusted estimator to OLS estimator.

Table 2: Critical Values for the Pooled RMA Tests with various N.

T	N=1	N=2	N=3	N=5	N=10	N=20	N=50	N=100
				Constant Case 5%				
50	-1.90	-1.86	-1.84	-1.81	-1.77	-1.73	-1.71	-1.69
100	-1.89	-1.86	-1.83	-1.81	-1.77	-1.73	-1.71	-1.69
150	-1.88	-1.86	-1.83	-1.81	-1.77	-1.73	-1.71	-1.69
200	-1.88	-1.86	-1.83	-1.81	-1.77	-1.73	-1.71	-1.69
				Linear Trend Case 5%				
50	-1.87	-1.85	-1.81	-1.79	-1.76	-1.74	-1.72	-1.71
100	-1.86	-1.83	-1.80	-1.78	-1.75	-1.74	-1.72	-1.70
150	-1.86	-1.82	-1.80	-1.78	-1.75	-1.73	-1.70	-1.68
200	-1.86	-1.82	-1.80	-1.78	-1.75	-1.71	-1.69	-1.68
				Constant Case 10%				
	N=1	N=2	N=3	N=5	N=10	N=20	N=50	N=100
50	-1.54	-1.50	-1.48	-1.45	-1.40	-1.37	-1.34	-1.33
100	-1.54	-1.50	-1.48	-1.45	-1.41	-1.37	-1.34	-1.33
150	-1.54	-1.50	-1.48	-1.45	-1.41	-1.37	-1.34	-1.33
200	-1.54	-1.50	-1.48	-1.45	-1.41	-1.36	-1.34	-1.33
				Linear Trend Case 10%				
50	-1.51	-1.48	-1.45	-1.42	-1.40	-1.37	-1.35	-1.33
100	-1.51	-1.48	-1.45	-1.42	-1.39	-1.37	-1.35	-1.33
150	-1.51	-1.48	-1.44	-1.42	-1.39	-1.36	-1.33	-1.31
200	-1.51	-1.48	-1.44	-1.41	-1.37	-1.35	-1.33	-1.31

Table 3: Bias and mean square error (MSE) of the pooled RMA and LSDV estimators
$\rho = 1, N = 1000$, AR(2) with incidental trends

Regression: $y_{it} = a_i + b_i t + \rho y_{it-1} + \phi \Delta y_{it-1} + u_{it}$

	$\phi = 0.4$				$\phi = -0.4$			
T	(A)	(B)	(C)	(D)	(A)	(B)	(C)	(D)
5	-1.600	1.461	6.493	0.839	-2.401	0.802	5.766	0.113
10	-0.661	0.266	0.698	0.164	-1.139	0.263	1.298	0.054
15	-0.392	0.075	0.118	0.037	-0.738	0.107	0.545	0.022
20	-0.276	0.032	0.031	0.014	-0.546	0.055	0.299	0.010
25	-0.212	0.018	0.011	0.008	-0.433	0.033	0.188	0.006
30	-0.172	0.011	0.005	0.005	-0.359	0.022	0.129	0.004
35	-0.145	0.008	0.003	0.003	-0.307	0.015	0.094	0.003
40	-0.125	0.006	0.002	0.002	-0.268	0.012	0.072	0.002
45	-0.110	0.004	0.001	0.002	-0.237	0.009	0.056	0.002
50	-0.098	0.003	0.001	0.001	-0.214	0.007	0.046	0.001

Note: (A) = Bias of LSDV estimator for ρ; (B) =Bias of the pooled recursive mean adjusted estimator for ρ; (C) = MSE of LSDV; (D) = MSE ratio of recursive mean adjusted estimator to OLS estimator.

Table 4: Local Power Envelops for Meta and Pooled RMA
Tests under Heterogeneous local to unity parameters. ($N = 3$)

	Constant					Linear Trend				
	P	PRMA-FGLS				P	PRMA-FGLS			
$c + d$		d=0	d=1	d=3	d=5		d=0	d=1	d=3	d=5
6	0.68	0.82	0.82	0.76	0.61	0.27	0.35	0.35	0.32	0.30
7	0.79	0.91	0.90	0.86	0.75	0.35	0.46	0.46	0.43	0.38
8	0.87	0.96	0.95	0.93	0.86	0.45	0.58	0.57	0.53	0.47
9	0.93	0.98	0.98	0.96	0.92	0.55	0.69	0.68	0.64	0.58
10	0.96	0.99	0.99	0.98	0.96	0.65	0.79	0.78	0.75	0.68
11	0.98	1.00	1.00	0.99	0.98	0.74	0.86	0.86	0.83	0.77
12	0.99	1.00	1.00	1.00	0.99	0.81	0.91	0.91	0.90	0.85
13	1.00	1.00	1.00	1.00	1.00	0.87	0.95	0.95	0.94	0.91
14	1.00	1.00	1.00	1.00	1.00	0.92	0.97	0.97	0.96	0.94
15	1.00	1.00	1.00	1.00	1.00	0.95	0.99	0.99	0.98	0.97

Table 5: Size and power
Case I – AR(1), $\rho = 0.95$ under the alternative

	Size for ADF tests (5%)							
	Constant				Linear Trend			
T	N=3	N=5	N=7	N=10	N=3	N=5	N=7	N=10
50	5.67	6.51	10.29	14.50	5.71	7.88	11.53	19.36
100	5.25	5.75	6.88	8.63	5.35	6.40	7.75	10.48
150	5.02	5.42	6.22	7.32	4.89	5.76	6.81	8.42
200	5.42	5.16	5.12	6.61	5.51	5.33	6.61	7.11
	Size for PRMA-FGLS tests (5%)							
50	5.45	5.65	6.60	6.26	5.64	6.14	5.97	5.98
100	5.02	5.35	6.25	5.86	5.29	5.77	5.82	5.56
150	4.94	5.37	5.30	5.89	4.78	5.40	5.27	5.39
200	5.51	5.07	4.91	5.57	5.31	5.08	5.21	5.11
	Size Adjusted Power for ADF tests (5%)							
50	12.2	18.1	19.8	25.9	6.8	7.8	9.0	10.1
100	31.7	53.2	66.1	80.9	14.2	19.9	24.8	31.5
150	62.7	87.1	95.8	99.4	29.9	45.5	57.2	72.4
200	86.3	98.8	99.9	100	48.8	76.6	86.5	96.6
	Size Adjusted Power for PRMA-FGLS tests (5%)							
50	21.0	33.5	43.2	56.1	8.7	09.8	12.2	13.7
100	57.6	83.6	93.4	98.7	20.3	28.4	36.4	49.5
150	89.1	98.9	99.9	100	41.9	63.1	77.2	89.8
200	98.4	100	100	100	65.6	88.2	96.4	99.4

Table 6: Size and power

Case II – AR(2), $\phi = [-0.2, 0.2]$ $\rho = 0.95$ under the alternative

	Size adjusted power for ADF tests (5%)							
	Constant				Linear Trend			
T	N=3	N=5	N=7	N=10	N=3	N=5	N=7	N=10
50	5.97	6.51	8.81	12.92	6.82	7.81	10.04	15.93
100	5.71	5.40	5.82	6.93	5.26	5.07	5.57	6.62
150	4.85	4.87	4.55	5.64	5.09	4.89	4.65	5.24
200	5.19	5.07	4.51	5.25	4.94	4.60	4.20	4.21
	Size adjusted power for PRMA-FGLS tests (5%)							
50	5.63	6.00	7.63	7.83	5.30	5.29	5.95	5.92
100	5.37	5.48	6.36	6.40	4.85	5.13	5.57	5.16
150	5.17	5.45	5.56	5.83	5.19	5.35	5.20	4.88
200	5.37	5.14	5.19	5.72	4.56	5.17	4.87	5.20
	Size adjusted power for ADF tests (5%)							
50	11.2	16.4	16.3	19.2	6.7	7.3	7.8	7.4
100	26.5	43.3	52.4	65.2	13.2	18.0	19.6	26.0
150	56.8	77.2	87.8	94.3	25.9	37.1	46.2	55.4
200	81.2	94.3	97.8	99.5	45.0	62.9	71.8	84.2
	Size adjusted power for PRMA-FGLS tests (5%)							
50	20.4	30.1	33.7	43.4	7.7	8.8	10.0	10.6
100	52.1	75.3	85.4	95.0	18.6	25.0	29.9	39.6
150	84.2	96.8	99.1	99.9	34.6	52.9	65.6	79.9
200	97.0	99.8	100	100	61.3	80.7	91.0	97.4

Table 7: Size and power
Case V – MA(1) error $\phi = -0.2$,
$\rho = 0.95$ under the alternative

	Size adjusted power for ADF tests (5%)							
	Constant				Linear Trend			
T	N=3	N=5	N=7	N=10	N=3	N=5	N=7	N=10
50	5.72	6.58	10.50	13.94	4.80	6.32	9.98	18.62
100	4.76	5.42	6.88	8.60	4.36	5.50	7.10	10.46
150	4.78	5.06	6.12	7.80	4.66	5.56	6.48	8.20
200	5.68	5.40	5.34	7.02	5.34	5.26	6.18	7.16
	Size for PRMA-FGLS tests (5%)							
50	4.98	5.80	7.60	7.98	1.94	1.76	2.24	2.32
100	4.88	5.24	6.00	6.30	3.12	2.94	3.30	3.16
150	5.06	5.16	5.52	6.18	3.48	3.68	3.62	3.84
200	5.92	4.98	5.06	5.96	4.06	3.62	4.14	3.66
	Size adjusted power for ADF tests (5%)							
50	10.0	16.1	18.6	21.2	6.7	7.4	8.5	8.0
100	28.7	48.3	58.6	71.8	14.3	19.0	21.7	26.6
150	57.6	82.4	92.3	98.0	27.4	38.9	49.2	64.3
200	79.3	96.8	99.7	100	41.6	67.4	80.0	92.7
	Size adjusted power for FRMA-FGLS tests (5%)							
50	19.8	27.1	31.4	40.9	7.9	9.0	9.9	10.2
100	50.6	76.9	87.8	96.3	17.0	25.3	29.4	39.7
150	83.2	97.5	99.8	100	37.5	54.9	67.3	81.6
200	95.8	99.9	100	100	58.9	83.1	92.0	98.4

Table 8: Comparison between RMA-FGLS and Bai & Ng Tests
AR(1), $\rho = 0.95$ under the alternative

	Size for Bai and Ng tests (5%)									
	Constant					Linear Trend				
T	N=5	N=10	N=15	N=20	N=30	N=5	N=10	N=15	N=20	N=30
50	6.30	13.74	3.58	3.02	3.16	0.44	14.64	3.70	3.26	3.20
100	7.88	13.44	3.54	3.14	2.40	0.40	14.38	3.38	3.46	2.92
150	8.16	13.42	3.48	3.24	2.64	0.80	14.76	3.20	3.60	3.34
200	8.74	13.62	2.98	2.86	2.88	0.96	13.90	3.24	3.48	2.66
	Size for PRMA-FGLS Test (5%)									
50	6.16	5.52	5.96	4.02	1.10	6.08	6.02	4.80	3.08	0.58
100	5.76	5.32	5.94	5.48	4.58	5.76	5.24	5.02	4.26	3.60
150	4.94	5.26	5.36	5.90	4.86	5.34	4.78	5.38	5.02	4.20
200	5.14	5.12	5.40	4.82	5.34	5.40	4.86	4.90	4.78	4.64
	Size adjusted power for Bai and Ng tests (5%)									
50	8.3	12.9	85.8	95.0	99.4	5.7	5.2	16.0	18.6	24.7
100	12.0	41.0	100	100	100	7.5	11.7	60.8	73.6	88.9
150	15.6	73.0	100	100	100	9.2	23.4	96.2	98.9	100
200	19.0	90.8	100	100	100	11.4	37.8	99.9	100	100
	Size adjusted power for PRMA-FGLS tests (5%)									
50	49.1	80.4	88.6	93.1	92.8	11.6	15.4	18.2	22.2	21.6
100	90.2	99.8	100	100	100	36.4	65.2	79.9	88.6	93.4
150	99.7	100	100	100	100	74.3	97.4	99.5	100	100
200	100	100	100	100	100	94.7	100	100	100	100

Table 9: Rejection Rates of CRMA Tests

	Average of individual sizes (5%)							
	Constant				Linear Trend			
	$\lambda_i - U\,[0,5]$		$\lambda_i - U\,[1,5]$		$\lambda_i - U\,[01,5]$		$\lambda_i - U\,[1,5]$	
T	N = 5	N=15	N = 5	N=15	N = 5	N=15	N = 5	N=15
50	5.78	5.32	4.32	4.46	5.60	5.03	5.08	4.64
100	7.06	5.68	3.58	3.91	5.55	4.95	3.88	4.12
150	9.61	5.99	3.82	3.25	6.55	5.13	3.96	3.69
200	12.1	7.19	3.20	3.24	8.24	5.65	3.77	3.45
	Size of tests with $\bar{y}_t$ (5%)							
50	3.96	4.40	3.90	4.48	5.22	4.28	5.24	4.26
100	3.56	4.30	3.62	4.32	4.04	4.30	4.02	4.24
150	4.16	3.54	4.24	3.58	4.20	3.66	4.24	3.62
200	3.40	4.08	3.42	4.14	3.82	3.64	3.88	3.66
	Average individual size adjusted powers (5%)							
50	40.1	42.0	47.7	48.2	25.4	27.7	29.5	31.6
100	64.5	69.9	78.4	79.9	51.6	54.8	62.7	63.3
150	74.7	82.4	91.5	93.1	65.7	71.7	80.0	82.2
200	78.9	87.4	97.1	97.7	72.6	79.6	89.1	90.7
	Size adjusted powers with $\bar{y}_t$ (5%)							
50	55.7	56.4	56.0	56.4	35.6	39.6	36.1	39.7
100	84.4	84.9	84.4	84.7	72.2	71.2	72.5	71.4
150	94.2	94.7	94.1	94.7	85.8	87.2	86.1	87.3
200	98.0	98.3	98.0	98.3	92.7	93.1	92.7	93.0

Table 10: Higher Frequency, Higher Power
– $N = 18$, AR(1) error, PRMA-FGLS Tests

Annual $T = 30$			Month $T = 360$		
ρ	Constant	Trend	$\rho^{1/12}$	Constant	Trend
0.70	99.9	83.9	0.971	100	100
0.75	99.5	70.7	0.976	100	99.9
0.80	97.4	51.3	0.982	100	96.9
0.85	88.7	31.3	0.987	100	75.9
0.90	64.1	15.7	0.991	98.3	35.0
0.91	59.6	14.1	0.992	95.7	29.5
0.93	44.7	10.4	0.994	85.8	18.6
0.95	27.5	7.7	0.996	66.2	12.0

Table 11: Are Real Exchange Rates Nonstationary?
7 OECD Monthly Panels from 1975, January – 2005, March.

Country	CADF	$\hat{\rho}_{\text{ADF-flgs}}$	ADF-FGLS	$\hat{\rho}_{\text{rma-flgs}}$	PRMA-FGLS
Germany	-1.999	0.981	-5.475	0.990	-2.909
Austria	-2.770	0.980	-5.560	0.989	-3.001
Belgium	-3.978	0.982	-5.368	0.991	-2.549
Finland	-2.842	0.982	-5.152	0.991	-2.531
France	-4.980	0.981	-5.409	0.988	-3.126
Italy	-4.963	0.975	-6.187	0.986	-3.112
Netherland	-3.094	0.980	-5.588	0.990	-2.927
Spain	-5.219	0.976	-4.990	0.986	-2.360

Note: The following 6 countries are always included in the panel:
Canada, Japan, Norway, Sweden, Switzerland and UK.
One of EU countries is added into the panel. (Numeraire: USA)

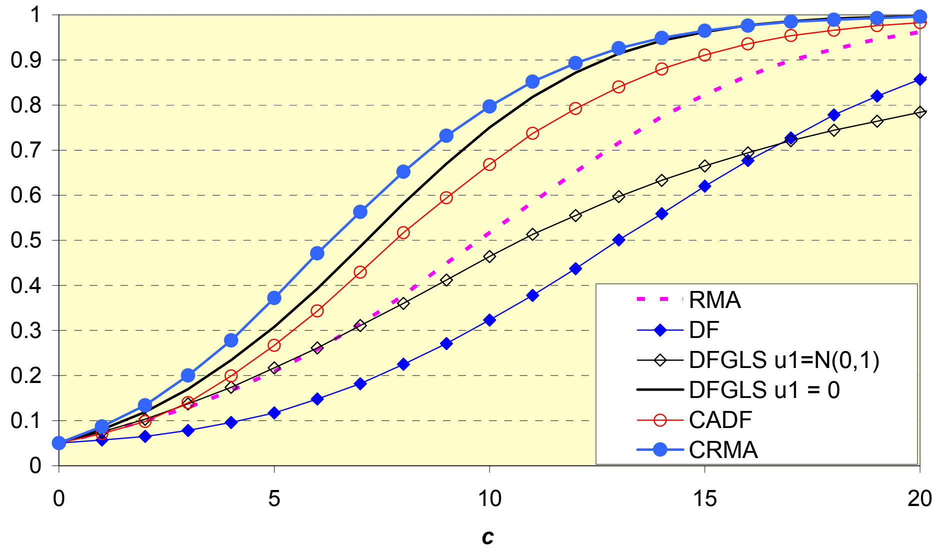

Figure 1: Asymptotic local powers for unknown constant $R^2 = \lambda^2 = 0.8$, $u_1 - N(0, 1)$

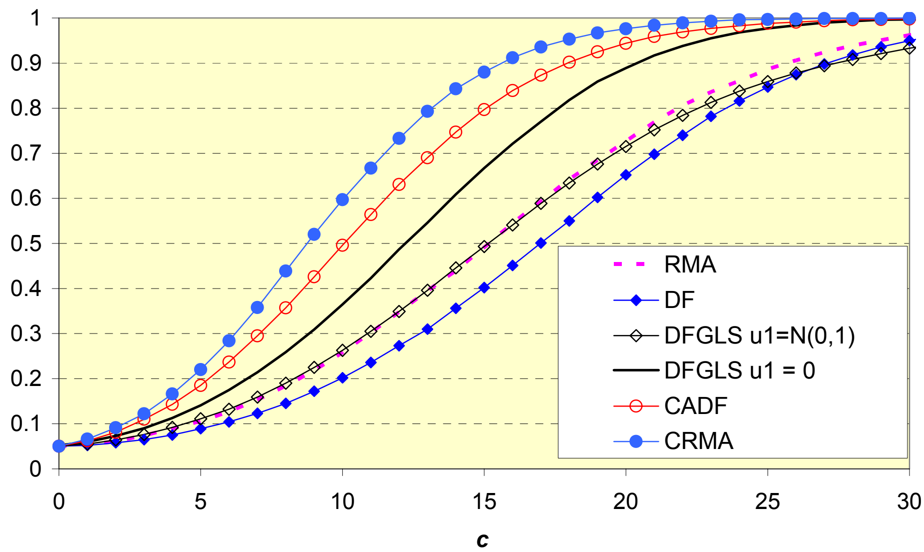

Figure 2: Asymptotic local powers for linear trend: $R^2 = \lambda^2 = 0.8$, $u_1 - N(0, 1)$

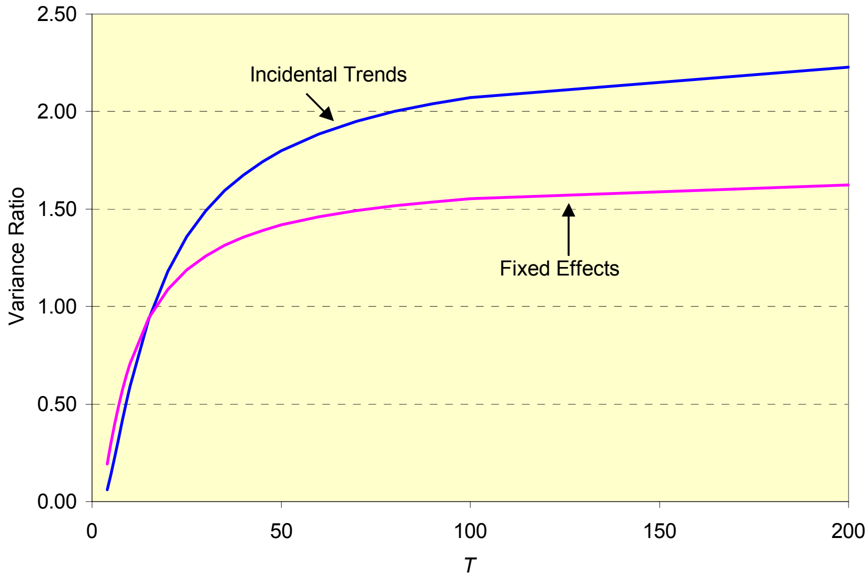

Figure 3: Asymptotic variance ratio of the pooled mean unbiased estimator to the pooled recursive mean adjusted estimator for panel AR(1) case.

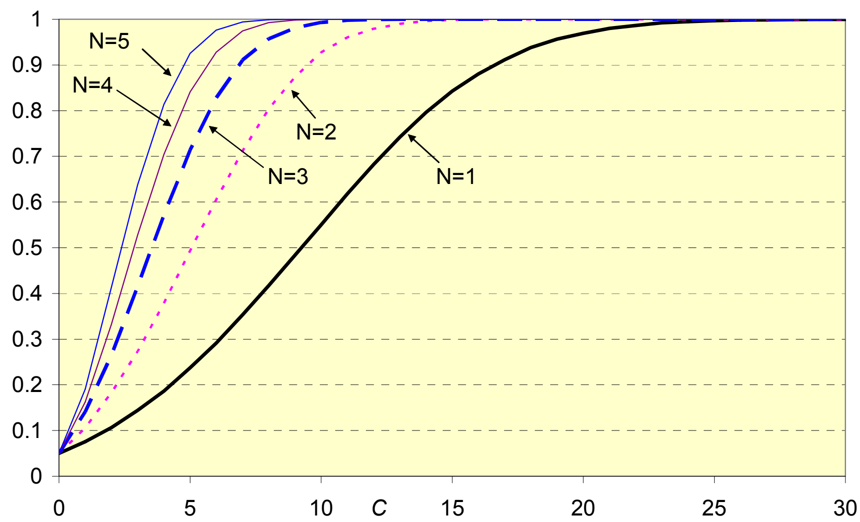

Figure 4: Asymptotic local power of the pooled RMA-FGLS test for fixed effects

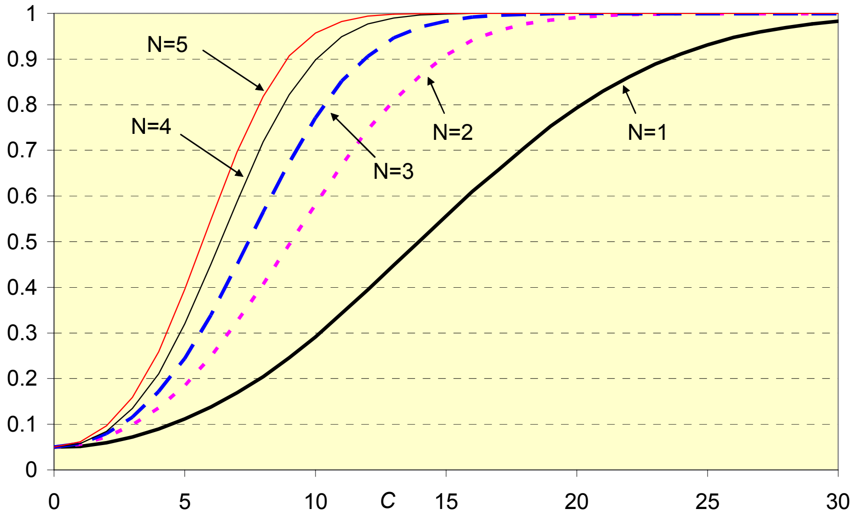

Figure 5: Asymptotic local power of the pooled RMA-FGLS test for incidental linear trends