HITOTSUBASHI UNIVERSITY

Global COE Hi-Stat Discussion Paper Series 256

Research Unit for Statistical and Empirical Analysis in Social Sciences (Hi-Stat)

Covariate Unit Root Test for Cross-Sectionally Dependent Panel Data

Eiji Kurozumi
Daisuke Yamazaki
Kaddour Hadri

November 2012
(Revised: August 2013)

Hi-Stat Discussion Paper

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

Covariate Unit Root Test for Cross-Sectionally Dependent Panel Data [1,2]

Eiji Kurozumi
Department of Economics
Hitotsubashi University

Daisuke Yamazaki
Department of Economics
Hitotsubashi University

Kaddour Hadri
Queen's University Management School
Queen's University Belfast

August, 2013

Abstract

This paper proposes the use of covariate unit root tests and the exploitation of the information on the cross-sectional dependence when the panel data null hypothesis of a unit root is rejected or when N is relatively small in order to help the interpretation of the test results.

In particular, it investigates the optimal point optimal covariate unit root test by Juhl and Xiao (2003), which is based on the theory by Hansen (1995) and Elliott and Jansson (2003). We first compare the asymptotic power function of the covariate test with those of panel unit root tests and show that the covariate unit root test can be potentially more powerful than panel unit root tests when the cross-sectional dimension is not so large. We also suggest several methods to choose appropriate covariates. The Monte Carlo simulations show that some of our methods work fairly well compared with the simple method of using only one covariate. Using our methods, we investigate the Prebish-Singer hypothesis for nine commodity prices and find that this hypothesis holds except for the price of petroleum. We also examine the PPP hypothesis employing eight real exchange rates of developed economies relative to the US dollar.

JEL classification: C12, C22, C23

Key words: covariate unit root test, cross-sectional dependence in panel data, point optimal test, squared correlation, power

[1]We are grateful to participants at the 17th African Econometric Society meeting and seminar participants at National University of Singapore, Kyoto University, Queen's Universit Belfast, and Hitotsubashi University. All errors are our responsibility.

[2]This research was supported by JSPS and BA under the Japan-Britain Research Cooperative Program, by The Seimei Foundation, the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No.23243038 and by the Global COE program of the Research Unit for Statistical and Empirical Analysis in Social Sciences, Hitotsubashi University. Address correspondence to Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan; e-mail: kurozumi@stat.hit-u.ac.jp

1. Introduction

Testing for a unit root has a long history and its application in economics is well understood. A variety of univariate unit root tests have been proposed in the literature. However, they have been, generally, criticized for their low power, particularly in cases where the series are close to a unit root. To increase the power of univariate tests, panel data unit root tests have been proposed and developed (cf. Baltagi (2008) and Breitung and Pesaran (2008)). A typical example is the investigation of whether the purchasing power parity (PPP) hypothesis holds among OECD countries using panel data unit root tests. Another aspect when employing panel data is that, generally, the cross-sections are correlated, particularly, in macro panel data where T and N are large. O'Connell (1998) was the first to show via simulation that the panel data unit root tests are considerably distorted when the likely presence of the cross-section dependency is not accounted for. The necessary treatment of the cross-sectional dependence has been investigated profusely in the last decade. Bai and Ng (2004) and Moon and Perron (2004) assumed a common factor structure to model strong cross-sectional dependence. They proposed a method for extracting the common factors in order to make the panel data cross-sectionally uncorrelated, and be hence able to apply panel unit root tests such as the IPS test by Im, Pesaran, and Shin (2003), the Fisher test by Maddala and Wu (1999) and Choi (2001), and the inverse normal test by Choi (2001). In contrast, Pesaran (2007) proposed to augment the regressions with the cross-sectional averages, while Chang (2002) and Chang and Song (2009) used nonlinear instrumental variables to mop up the cross-sectional dependence.

While we may be able to overcome the problem of the low power of univariate unit root tests by making use of panel data, it has been pointed out in the literature (cf. Pesaran (2012)) that it is difficult to interpret the results when panel unit root tests reject the null hypothesis. This is because typical panel data tests reject the null of a unit root if some of individuals are stationary and the others have a unit root; thus the rejection of the null hypothesis implies that not all individuals have a unit root, but we do not know which individuals are stationary. In this case, we may partition cross-sectional units into sub-groups and/or estimate the proportion of stationary units, as suggested by Pesaran (2012) and references therein, and we can eventually go back to univariate tests which we shall consider in this paper. See also Elliott and Pesavento (2006) for the discussion of the problem of the panel data approach.

There have also been significant efforts to improve the power of univariate tests. For instance, Hansen (1995) proposed to augment the regression for the ADF test with covariates correlated with the disturbance term of the process for which we want to test the unit root hypothesis. This covariate ADF (CADF) test was further extended to a point-optimal covariate (POC) unit root test by Elliott and Jansson (2003), the power function of which is tangent to the Gaussian power envelope at some point of the alternative. Juhl and Xiao (2003) proposed to modify the POC test by introducing the standard of optimality proposed by Cox and Hinkley (1974). More recently, Fossati (2013) extended covariate unit root tests to models with strucrural break, while Westerlund (2013) allowed for conditional heteroskedasticity. These papers showed that the powers of the ADF and the ADF-GLS tests are much improved if we can find covariates that are highly correlated with the disturbance term. In other words, the power improvement of these covariate unit root tests crucially depends on whether or not we can find appropriate covariates.

Although we may naturally find covariates in some cases, such as in the investigation of a technology shock by Christiano, Eichenbaum, and Vigfusson (2003), it is not always easy to find them in many practical situations. However, in cases where the null hypothesis of a unit root is rejected when investigating macro panel data, one possibility open to us is to test for a unit root in each cross-sectional unit employing covariate tests. In this case, the natural candidates for covariates are the series of individuals other than the one focused on because, macro panel data are typically cross-sectionally correlated. To increase the power of our covariate unit root tests, we must carefully choose the appropriate covariates among the potential candidates. Sometimes, in empirical analysis, only one covariate is used for covariate unit root tests such as in Elliott and Pesavento (2006), Amara and Papell (2006), and Christopoulos and León-Ledesma (2008). However, as was considered by Lee and Tsong (2011), there is no reason to use only one covariate, and we can expect that the covariate tests with several covariates would be more powerful than those with only one covariate as long as we choose an appropriate set of covariates.

In this paper, we first compare the optimal POC (OPOC) test by Juhl and Xiao (2003) with panel unit root tests such as the Fisher and the inverse normal tests under various situations. We do not intend to conclusively determine which test is better than the others; rather, our power comparison will tell us the case in which the OPOC test is preferred to the panel data tests with respect to power while helping us understand the extent to which we can downsize the number of cross-sectional units by keeping the power of the panel tests higher than that of the OPOC test so that we can effectively use either test in applications. The second contribution of this paper is that we propose some selection rules to help us to choose the appropriate covariates from the potential available candidates. In addition to the factor model approach proposed by Lee and Tsong (2011), we propose two other procedures based on the asymptotic power functions and the adjusted squared correlation. We will show that the latter two methods work fairly well in finite samples, while the former approach has a problem of controlling the size of the test in some cases.

In sum, our paper seeks: (a) To propose the use of univariate covariate unit root tests that exploit the cross-sectional dependence information contained in panel data when the null hypothesis of a unit root is rejected in the latter. (b) To compare the power of OPOC test with some well-known panel data unit root tests. (c) To propose new procedures for choosing the "best" covariates from a set of available contenders.

The rest of this paper is organized as follows. We briefly review the covariate unit root tests in Section 2. The asymptotic powers of the OPOC test and the panel tests are compared in Section 3. We propose three selection rules for covariates and investigate their finite sample properties in Section 4. Our methods are applied to the Prebish-Singer hypothesis and the PPP hypothesis in Section 5. Concluding remarks are given in Section 6.

2. Model and Covariate Unit Root Test

Let us consider the following panel model:

$$z_{it} = \beta_{i,0} + \beta_{i,1} t + u_{it} \quad \text{for} \quad i = 1, \cdots, N \quad \text{and} \quad t = 1, \cdots, T. \tag{1}$$

We call model (1) the trend case and the case with no linear trend ($\beta_{i,1} = 0$ for all i) the constant case.

Suppose that our interest is whether or not the first variable z_{1t} has a unit root. To focus on z_{1t}, let $y_t = z_{1t}$ and $x_t = [z_{2t}, \cdots, z_{Nt}]'$. Stacking variables in the cross-sectional direction, model (1) can be expressed in vectorized form as

$$z_t = \beta_0 + \beta_1 t + u_t, \quad A(L)u_t(\rho) = \varepsilon_t, \quad \text{where} \quad u_t(\rho) = \begin{bmatrix} (1-\rho L)u_{y,t} \\ u_{x,t} \end{bmatrix}, \tag{2}$$

$z_t = [y_t, x_t']'$, $\beta_0 = [\beta_{y,0}, \beta_{x,0}']'$, $\beta_1 = [\beta_y, \beta_{x,1}']'$, $u_t = [u_{y,t}, u_{x,t}']'$, $\varepsilon_t = [\varepsilon_{y,t}, \varepsilon_{x,t}']'$, and $A(L)$ is a lag polynomial of order p with L being the lag operator. Since $A(L)$ is supposed to be invertible by Assumption =A1 below, $u_t(\rho)$ is assumed to be statonary. Note that the variables and parameters are decomposed conformably with $z_t = [y_t, x_t']'$. For later use, we define the long-run variance of $u_t(\rho)$ and the long-run squared correlation as

$$\Omega = A^{-1}(1)\Sigma A'^{-1}(1) = \begin{bmatrix} \omega_{yy} & \omega_{yx} \\ \omega_{xy} & \Omega_{xx} \end{bmatrix} \quad \text{and} \quad R^2 = \omega_{yy}^{-1}\omega_{yx}\Omega_{xx}^{-1}\omega_{xy}, \quad \text{respectively.}$$

Model (1)-(2) allows for heterogeneity and cross-sectional dependence in u_{it} through the lag-polynomial $A(L)$ and the innovation variance matrix Σ. Note that the factor structure is also included as a special case by assuming that $A(L)u_t(\rho) = \varepsilon_t$ with $\varepsilon_t = \Lambda f_t + e_t$, where f_t is an r-dimensional common factor, Λ is an $N \times r$ loading matrix, and e_t consists of the idiosyncratic errors.

The following assumption, which is standard in the time series literature, is made throughout this paper.

Assumption A1 *(a)* $\{\varepsilon_t\}$ *is a martingale difference sequence with respect to* $\mathcal{F}_t = \sigma\{\varepsilon_t, \varepsilon_{t-1}, \cdots\}$ *with* $E[\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}] = \Sigma > 0$ *for all* t. *(b)* $\sup_t E\|\varepsilon_t\|^{2+\kappa} < \infty$ *for some* $\kappa > 0$. *(c)* $|A(z)| = 0$ *implies* $|z| > 1$. *(d)* $u_0, u_{-1}, \cdots, u_{-p}$ *are* $O_p(1)$ *and independent of* T.

Since we are interested in whether or not y_t is a unit root process, we consider the following testing problem:

$$H_0: \ \rho = 1 \quad \text{vs.} \quad H_1: \ |\rho| < 1.$$

In this case, the common practice is to consider a univariate model for y_t given by

$$\Delta y_t = \beta_{y,0} + \beta_{y,1}t + \rho y_{t-1} + \psi_1 \Delta y_{t-1} + \cdots + \psi_p \Delta y_{t-p} + u_{y,t} \tag{3}$$

and conduct the ADF test by Dickey and Fuller (1979) and Said and Dickey (1984) and the ADF-GLS test by Elliott, Rothenberg, and Stock (1996).

While these well-known tests concentrate on a univariate process y_t, Hansen (1995) proposed to make use of the covariates x_t to improve the power of the ADF tests and considered augmenting (3) by x_t:

$$\Delta y_t = \beta_{y,0} + \beta_{y,1}t + \rho y_{t-1} + \psi_1 \Delta y_{t-1} + \cdots + \psi_p \Delta y_{t-p} + \gamma' x_t + e_{y,t}. \tag{4}$$

The CADF test is based on the t-statistic for ρ in (4). The improvement of power comes from the fact that the covariate vector x_t is correlated with $u_{y,t}$; part of the fluctuation in $u_{y,t}$ is explained by x_t and the variance of $e_{y,t}$ in (4) becomes smaller than that of $u_{y,t}$, so we can estimate ρ more efficiently by (4) than by (3).

For further refinement of the CADF test, Elliott and Jansson (2003) proposed a point optimal covariate unit root (POC) test by considering the local-to-unity system for (2). More precisely, they assumed that $\rho = 1 - c/T$ for $c \geq 0$ and proposed to construct the likelihood ratio (LR) test statistic $\Lambda(1, \bar{\rho})$ assuming that $c = \bar{c}$ (or $\bar{\rho} = 1 - \bar{c}/T$). This test has been shown to depend on only c, $\bar{c}$, and R^2 asymptotically, and the asymptotic local power function can then be written as $h_{poc}(c, \bar{c}, R^2)$. By the Neyman-Pearson lemma, the LR test is a most powerful test against $c = \bar{c}$ under the assumption of normality, and the Gaussian power envelope, which was also investigated by Hansen (1995), is then given by $h_{poc}(c, c, R^2)$. We can see that the power function of the LR test is tangent to the power envelope at $c = \bar{c}$, but it is generally lower than the envelope at $c \neq \bar{c}$. Note that we need to prespecify $\bar{c}$ to construct the test statistic $\Lambda(1, \bar{\rho})$. Elliott and Jansson (2003) recommended using $\bar{c} = 7$ in the constant case and $\bar{c} = 13.5$ in the trend case.

Although Elliott and Jansson (2003) showed that the power function of the POC test is close to the power envelope in a wide range of alternatives for different values of R^2, Juhl and Xiao (2003) pointed out that there are other possibilities for the choice of $\bar{c}$. This is because the value of $\bar{c}$ suggested by Elliott and Jansson (2003) is based on the choice of Elliott, Rothenberg, and Stock (1996), which implies that the power function of the POC test with the suggested $\bar{c}$ is tangent to the 50% point of the power envelope only when $R^2 = 0$, so that $h_{poc}(\bar{c}, \bar{c}, 0) = 0.5$. However, because the power function depends on the true value of R^2, Juhl and Xiao (2003) concluded that the choice of $\bar{c}$ should also depend on R^2. Moreover, there is no theoretical reason why we choose the value of $\bar{c}$ at which the power function is tangent to the 50% point of the power envelope. Rather, they proposed to choose $\bar{c}$ at which

$$\int_0^\infty \left[h_{poc}(c, c, R^2) - h_{poc}(c, \bar{c}, R^2)\right] dc \tag{5}$$

is minimized; that is, for a given value of R^2, the averaged loss of power compared to the power envelope is minimized at $\bar{c}$. This criterion of optimality was originally proposed by Cox and Hinkley (1974) and also adopted by Kurozumi (2003) in a different situation. Juhl and Xiao (2003) called the POC test with $\bar{c}$ minimizing (5) the optimal point optimal covariate unit root (OPOC) test. The optimal values of $\bar{c}$ for a given value of R^2 are given in Table 1 in Juhl and Xiao (2003).

According to Elliott and Jansson (2003), the OPOC test is constructed as follows. First, we estimate $\beta = [\beta_0', \beta_1']'$ from the quasi-differenced series under the null and alternative hypotheses, that is,

$$\tilde{\beta}(r) = \left[S\left(\sum_{t=1}^T d_t(r)\hat{\Omega}^{-1}d_t'(r)\right)S\right]^- \left[S\sum_{t=1}^T d_t(r)\hat{\Omega}^{-1}z_t(r)\right]$$

for $r = 1$ and $r = \bar{\rho}$, where $z_1(r) = [y_1, x_1']'$ for $t = 1$ and $z_t(r) = [(1 - rL)y_t, x_t']'$ for $t > 1$, $d_1'(r) = [I_N, I_N]$ for $t = 1$, and

$$d_t'(r) = \begin{bmatrix} 1-r & 0 & (1-rL)t & 0 \\ 0 & I_{N-1} & 0 & tI_{N-1} \end{bmatrix}$$

for $t > 1$, $S = diag\{I_N, 0\}$ in the constant case and $S = I_{2N}$ in the trend case, B^- is the Moore-Penrose inverse of a matrix B, and $\hat{\Omega}$ is the estimator of the long-run variance Ω under the null hypothesis. We next construct the detrended series given by

$$\tilde{u}_t(r) = z_t(r) - d_t'(r)\tilde{\beta}(r)$$

for $r = 1$ and $\bar{\rho}$. Using $\tilde{u}_t(r)$, we estimate the VAR model of order p and obtain the estimated residual $\tilde{\varepsilon}_t(r)$ and the estimator of variance $\tilde{\Sigma}(r)$. Then, the test statistic is given by

$$\Lambda(1, \bar{\rho}) = T\left[tr\left(\tilde{\Sigma}^{-1}(1)\tilde{\Sigma}(\bar{\rho})\right) - (N - 1 + \bar{\rho})\right].$$

The rejection region is the left-hand tail of the distribution of $\Lambda(1, \bar{\rho})$.

3. Asymptotic Power Comparison

As discussed in the introduction, when in panel data the null hypothesis of a unit root is rejected, we may then rely on univariate unit root tests to ascertain which cross-section variable is stationary and which one is not. In this situation, we may use for a given cross-section variable the OPOC test and employ the other cross-sectional variables as covariates and hence use the information contained in the cross-sectional dependence. We are also interested in this paper to evaluate the power of the OPOC test and find out if it is potentially more powerful than panel unit root tests. Theoretically, we expect that the former would be more powerful than the latter for relatively small N but as N gets larger, the latter would dominate the former in terms of power. Therefore, the power comparison may help us understand the extent to which we can decrease the number of cross-sectional units for panel unit root tests while keeping the power of the panel tests higher than the univariate test as well as ascertain the maximum number of cross-section units from which we can rely more on the univariate OPOC test than panel unit root tests.

In particular, we compare the OPOC test with the following panel unit root tests: the Fisher test by Maddala and Wu (1999) and Choi (2001) and the inverse normal test by Choi (2001). The most commonly used testing procedures for cross-sectionally dependent panel data involves first estimating the factor model, as suggested by Bai and Ng (2004), and then applying the Fisher test or the inverse normal test to the estimated idiosyncratic errors, which are asymptotically cross-sectionally uncorrelated. These tests are desirable for the power comparison with univariate unit root tests because their asymptotics do not have to rely on the joint limit theory. They only need T-asymptotics with a fixed N, so that the same local-to-unity system can be assumed for the power comparison. [3] In this sense,the Fisher and the inverse normal tests seem to be suitable as benchmarks.

Before proceeding with the comparison, we should note that the premises for the panel unit root tests are different from those of the covariate unit root tests; the former assumes

[3] Exactly speaking, The Fisher and the inverse normal tests can be applied with a fixed N whereas the factor estimation based on Bai and Ng (2004) requires both N and T go to infinity. Since Bai and Ng's method works well even for samll N as reported by Tanaka and Kurozumi (2012), we proceed with the comarison using the Fihser and the inverse normal tests.

that all individuals have a unit root under the null hypothesis while at least a part of them is stationary under the alternative. This implies that the processes for all individuals may be characterized as a local-to-unity system. On the other hand, only the variable of interest has a unit root for the covariate unit root tests under the null hypothesis and covariates must be stationary under both the null and the alternative. These premises are different and it seems there is no meaning for comparing the powers of these tests. However, we sometimes encounter the case where both tests can be applied in empirical analysis. Suppose that we have a set of panel data, $z_{1,t}, z_{2,t}, \cdots, z_{N,t}$, and implement univariate unit root and stationarity tests for each of individuals to check the I(0)/I(1) property. Typically, we may find that some individuals are stationary and some others have a unit root, but we cannot determine the I(0)/I(1) property for the rest of the variables. We denote those groups as $z^0_{i,t}$, $z^1_{i,t}$, and $z^*_{i,t}$, respectively. In this case, we may apply panel unit root tests for $z^*_{i,t}$ because they satisfy the premises for panel unit root tests. On the other hand, if we want to pick up one variable, $y_{1,t}$, from $z^*_{i,t}$ and focus on it, we can implement the covariate unit root tests for $y_{1,t}$ with $z^0_{i,t}$ and $\Delta z^1_{i,t}$ as covariates. We may further use the first differences of $z^*_{i,t}$ as covariates because they are possibly strongly serially correlated. Then, we have a situation in which both panel unit root tests and the covariate unit root tests are valid and can be applied in practical analysis.

The Fisher and the inverse normal tests are constructed as follows. First, we construct the ADF-GLS test statistic t_i^{gls} for $i = 1, \cdots, N$. Then, the asymptotic p-value for t_i^{gls} is given by $\pi_i = G_0(t_i^{gls})$, where $G_0(\cdot)$ is the distribution function of t_i^{gls} under the null hypothesis. The Fisher and the inverse normal test statistics are defined as

$$F = -2\sum_{i=1}^{N} \log(\pi_i) \quad \text{and} \quad Z = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Phi^{-1}(\pi_i),$$

respectively, where $\Phi(\cdot)$ is the distribution function of a standard normal random variable. Since $\pi_i \sim U(0,1)$ asymptotically under the null hypothesis where $U(a,b)$ denotes a uniform distribution on the interval (a,b), it can be shown that $F \xrightarrow{d} \chi^2_{2N}$ and $Z \xrightarrow{d} N(0,1)$ as $T \to \infty$ with a fixed N under the null hypothesis, where $\xrightarrow{d}$ signifies convergence in distribution. Note that the Fisher and the inverse normal tests requires cross-sectional independence.

To compare the asymptotic powers, we consider a simple model (1) with

$$u_{it} = \left(1 - \frac{c_i}{T}\right) u_{it-1} + \varepsilon_{it} \quad \text{where} \quad c_i \geq 0. \tag{6}$$

Using the above local-to-unity system, Elliott and Jansson (2003) showed that the POC test statistic $\Lambda(1, \bar{\rho})$ weakly converges to $h_{poc}(c, \bar{c}, R^2)$, where

$$\begin{aligned} h_{poc}(c, \bar{c}, R^2) = (\bar{c}^2 - 2c\bar{c}) \int (W_1^c)^2 + 2\bar{c} \int W_1^c dW_1 \\ + (\bar{c}^2 - 2c\bar{c})Q \int (\widetilde{W}_1^c)^2 + g(W_1^c) - 2\bar{c}\sqrt{Q} \int \widetilde{W}_1^c dW_2, \end{aligned} \tag{7}$$

where $W_1^c(s)$ is an Ornstein-Uhlenbeck process defined by $dW_1^c(s) = -cW_1^c(s)ds + dW_1(s)$ with $W_1(s)$ being a standard Brownian motion, $W_2(s)$ is an $N-1$-dimensional standard Brownian motion independent of $W_1(s)$, $Q = R^2/(1-R^2)$, $\widetilde{W}_1^c(s) = W_1^c(s) - \int W_1^c(r)dr$ in the constant case while $\widetilde{W}_1^c(s) = W_1^c(s) + (6s-4)\int W_1^c(r)dr - (12s-6)\int rW_1^c(r)dr$ in the trend case, and $g(W_1^c) = 0$ in the constant case while $g(W_1^c) = W_1^c(1)^2 - 1/(1 + \bar{c}^2/3 + \bar{c})\left\{(1+\bar{c})W_1^c(1) + \bar{c}^2 \int rW_1^c\right\}^2$ in the trend case. This power function is calculated by approximating Brownian motions using the normalized partial sums of 2,000 independent standard normal random variables.

The asymptotic local power functions of the two panel unit root tests are obtained as follows. We first note that under the local alternative $\rho_i = 1 - c_i/T$, $t_i^{gls} \sim G_{c_i}(\cdot)$, where $G_{c_i}(\cdot)$ is the distribution function of t_i^{gls}. Then, we can see that

$$\pi_i = G_0(t_i^{gls}) = G_0(G_{c_i}^{-1}(G_{c_i}(t_i^{gls}))) = G_0(G_{c_i}^{-1}(U_i)),$$

where $U_i \sim U(0,1)$. Hence, the asymptotic power function of the Fisher test at significance level α can be expressed as

$$\begin{aligned} h_f(c_1, \cdots, c_N) &= P\left(-2\sum_{i=1}^{N} \log \pi_i \geq \chi^2_{2N,\alpha}\right) \\ &= P\left(\prod_{i=1}^{N} \pi_i \leq \exp\left[-\frac{\chi^2_{2N,\alpha}}{2}\right]\right) \\ &= P\left(\prod_{i=1}^{N} G_0(G_{c_i}^{-1}(U_i)) \leq \exp\left[-\frac{\chi^2_{2N,\alpha}}{2}\right]\right), \end{aligned}$$

where $\chi^2_{2N,\alpha}$ denotes the $(1-\alpha)$ quantile of a χ^2_{2N} distribution. Similarly, for the inverse normal test, the power function is given by

$$h_{in}(c_1, \cdots, c_N) = P(Z \leq z_\alpha) = P\left(\sum_{i=1}^{N} \Phi^{-1}(G_0(G_{c_i}^{-1}(U_i))) \leq \sqrt{N} z_\alpha\right),$$

where z_α is the α quantile of $N(0,1)$. Both of these power functions can be calculated by generating N independent $U(0,1)$ distributions for given $c_1, \cdots, c_N$, while the distribution function of t_i^{gls} for a given c_i is obtained by the similar approximation to the case of the OPOC test.

Figures 1 and 2 show the asymptotic limiting powers of the OPOC test when $R^2 = 0.3$, 0.6, and 0.9, and of the Fisher and the inverse normal tests when $c_1 = \cdots = c_N = c$ (referred to as "all") and when $c_1 = \cdots = c_{N/2} = c$ and $c_{N/2+1} = \cdots = c_N = 0$ (referred to as "half"). That is, the former corresponds to a case where all individuals do not have a unit root and hence higher power is expected for the panel unit root tests, while the latter case assumes that only half the individuals have a unit root. When $N = 4$ in the constant case as in Figure 1(i), the OPOC test is more powerful than the panel unit root tests if only half the individuals are stationary even when the long-run squared correlation is weak ($R^2 = 0.3$), whereas if none of the individuals has a unit root, the powers of the panel unit root tests increase dramatically and the OPOC test is superior only when R^2 is close to one. When $N = 8$, the powers of the panel tests increase, as expected, and they are close to the power of the OPOC test with $R^2 = 0.6$ and $R^2 = 0.9$ in the half and all cases, respectively. As seen in Figure 1(iii)-(v), when $N \geq 12$, the panel unit root tests in the half case dominate the OPOC test with moderate and weak long-run correlations, whereas the latter test with large R^2 is still more powerful than the former in the half case.

Figure 2 draws the power functions when a linear trend is included as a regressor. In this case, the relative performance of the OPOC test becomes better compared to the constant case. For example, when $N = 4$, the OPOC test with $R^2 = 0.6$ ($R^2 = 0.3$) is more powerful in a wide range of alternatives than the panel unit root tests even if all (half of) individuals are stationary. The OPOC test with the moderate long-run correlation is still preferred even when $N = 16$ if half the individuals have a unit root. Moreover, if the long-run squared correlation is close to one, the OPOC test almost dominates the panel tests even for $N = 20$.

On the other hand, if $R^2 = 0.3$, the former test is useful for only small N and the panel unit root tests are preferred.

To summarize the comparison, the relative performance depends on N, R^2, and the number of I(0) individuals, as expected. Roughly speaking, the OPOC test in the constant case is useful if N is at most around 8 except for the case of a large value of R^2, whereas we may rely on it in the trend case for a larger N such as 16 and 20.

4. Selection of Covariates

4.1. Three different selection rules

In practice, we need to find the appropriate covariates, but this is not necessarily an easy task. We may be able to find several candidates for covariates, as in the case of panel unit root tests, but we need a guideline for choosing the covariates. Intuitively, using as many covariates as possible may result in the highest R^2 (of course, this is not always the case), but the increase in the number of covariates implies a decrease in the degree of freedom in a sample, which may result in a loss of power. In this section, we propose three different approaches to select the appropriate covariates amongst the candidates by taking the degree of freedom into account.

The first method of choosing the covariates amongst the candidates is basically the same as in Lee and Tsong (2011), in which a factor model is assumed for a set of variables. Because the common factors play a key role for cross-sectional dependence, the natural candidates for covariates are the common factors in this case. We thus propose to estimate the common factors by the principal component method proposed by Bai (2003) and to use the estimated common factors as covariates. The advantage of this method is that it is computationally easy to obtain the covariates even if N is relatively large. However, there is no guarantee that the factor structure is the correct specification or that the long-run squared correlation becomes highest when using the selected covariates. We call this selection procedure the factor model rule.

The second method we propose is to make use of the asymptotic local power of the OPOC test. From its definition, we can see that when the true value of ρ equals $\rho^* = 1 - c^*/T$, the corresponding limiting power for a given R^2 is $h_{poc}(c^*, \bar{c}, R^2)$. Now suppose that we have a set of covariates with $R^2 = R_1^2$ and that the total number of regressors is k_1. In this case, the effective sample size is $T - k_1$, and ρ^* can be expressed as $\rho^* = 1 - c^*/T = 1 - c_1/(T - k_1)$, where $c_1 = c^*(T - k_1)/T$. Then, the corresponding asymptotic power against $\rho = \rho^*$ becomes $p_1 = h_{poc}(c_1, \bar{c}_1, R_1^2)$, where $\bar{c}_1$ is the pre-specified value of $\bar{c}$ required to construct the OPOC test. Similarly, if we use another set of covariates with $R^2 = R_2^2$ and the total number of regressors equals k_2, the asymptotic power is given by $p_2 = h_{POC}(c_2, \bar{c}_2, R_2^2)$, where $c_2 = c^*(T - k_2)/T$. For these two sets of covariates, we choose the first set (the second set) if $p_1 > p_2$ ($p_1 < p_2$). The key feature of this procedure is that even if $R_2 > R_1$, it is possible for p_1 to be greater than p_2, so that the first set of covariates is preferred. The illustration is also given in Figure 3. In the figure, even though the power function for R_2^2 dominates that for R_1^2, if we use too many covariates to attain $R^2 = R_2^2$, then the effective sample size is decreased, and p_2, the corresponding asymptotic power against $\rho = \rho^*$, can be smaller than p_1.

In this procedure, we have to determine a specific value of alternative c^* as a benchmark. Note that if c^* is too small or too large, the difference between power functions for different values of R^2 is small and we may obtain similar results for any set of covariates. Thus, we propose to choose a c^* at which the difference between the power functions for $R^2 = 0$ and 0.9 is maximized. According to our calculation, $c^* = 4.0$ in the constant case and $c^* = 5.8$ in the trend case. We call this selection procedure the asymptotic power rule.

The third selection rule is to mimic the well-known adjusted R^2. That is, we define

$$\bar{R}^2 = 1 - \frac{T-1}{T-k}(1 - R^2),$$

where k is the total number of regressors, and choose a set of covariates that attains the highest $\bar{R}^2$. Although this is an ad hoc rule and there is no theoretical support, this is the easiest rule among the three to be applied in empirical analysis.

4.2. Finite sample properties

To see the finite sample properties of the above three selection procedures, we conduct Monte Carlo simulations. The main purpose of the simulations is to see whether or not the selection rules proposed in this paper are more useful than the conventional rule of choosing only one covariate that attains the maximum value of R^2 among the candidates.

In order to apply the common factor rule, we have to determine the number of factors. According to de Silva, Hadri and Tremayne (2009), we adopt the Hannan-Quinn-type criterion HQ_4 with the maximum number of factors set to 4, which is the modified version of the information criteria proposed by Bai and Ng (2002). For the asymptotic power rule and the adjusted long-run squared correlation, we estimate R^2 under the null hypothesis as suggested by Elliott and Jansson (2003) with the order of lag selected by the Bayesian information criterion with the maximum lag length set to 4. In addition, we need the power functions for different values of R^2 for the asymptotic power rule. The asymptotic power functions for $R^2 = 0, 0.1, 0.2, \cdots, 0.9, 0.92, 0.94, 0.96, 0.98, 0.99$ and $c = 0, 1, \cdots, 15$ are calculated as in the previous section, and the power corresponding to the given c_1 and R^2 is obtained by interpolation.

The first data-generating process (DGP) we consider is the same as (2), where

$$\varepsilon_t \sim i.i.d.N(0, \Sigma) \quad \text{with} \quad \Sigma = \begin{bmatrix} 1 & \theta & \cdots & \theta \\ \theta & 1 & \ddots & \theta \\ \vdots & \ddots & \ddots & \vdots \\ \theta & \cdots & \theta & 1 \end{bmatrix},$$

$A(L) = diag\{1 - a_1 L, \cdots, 1 - a_N L\}$, and $\beta_0 = \beta_1 = 0$. We set $a_1 = \cdots = a_N = 0$ because they do not affect the value of R^2. We choose the values of θ so that R^2, which depends on θ and N, equals 0.3, 0.6, and 0.9 for a given N. The null hypothesis corresponds to the case of $\rho = 1$, while ρ is set to 0.98, 0.96, 0.94, 0.92, and 0.90 under the alternative. Since we consider a relatively small number of cross-sectional units, we set $N = 5$ and 10 while $T = 100$ and 200. All simulations are conducted using the GAUSS matrix language with the number of replications equal to 1,000.

Table 1 shows the empirical sizes and powers of the tests in the constant case. In the table, Λ_1 corresponds to the OPOC test using only one covariate that maximizes the estimate R^2 among $N-1$ variables in x_t, whereas the tests with the covariates selected by the factor model rule, the asymptotic power rule, and $\bar{R}^2$ are denoted by Λ_{fac}, Λ_{pow}, and $\Lambda_{\bar{R}^2}$, respectively. We also tabulate the rejection frequencies of the ADF-GLS test, t^{gls}, for the purpose of comparison. When $N = 5$, we cannot control the size of Λ_{fac} even when $T = 200$, whereas the size of this test becomes stable when $N = 10$. The other OPOC tests are slightly oversized when $T = 100$, but the size distortion is mitigated when $T = 200$. As expected, all the OPOC tests are more powerful than the ADF-GLS test, particularly, for a large R^2 or N. Although we have to carefully compare the nominal powers because of the size distortion, our selection rules, except for the factor model rule with $N = 5$, work well, and are more powerful than the OPOC test with one covariate. In general, the power differences tend to be larger as either the true value of R^2 or N gets larger. This is a natural result because when R^2 or N is large, there is room to improve the power of Λ_1 by using several covariates amongst the candidates. As in the case where $T = 200$, $N = 10$ and $R^2 = 0.6$ or 0.9, the largest power difference is about 10% in the constant case.

Table 2 reports the simulation results in the trend case. Again, the OPOC test with the factor model rule fails to control the empirical size when $N = 5$. As a whole, the relative performance of the tests is similar to the constant case, but it seems that the power differences are larger in the trend case than the constant case. For example, the largest power difference in the trend case is about 20% as in the case where $T = 200$, $N = 10$, and $R^2 = 0.9$, whereas it is about 10% in the constant case.

The second DGP we consider is the factor model given by

$$z_t = \beta_0 + \beta_1 t + u_t, \quad A(L)u_t(\rho) = \lambda f_t + \varepsilon_t,$$

where $f_t \sim i.i.d.N(0, \sigma_f^2)$ is a one-dimensional common factor, $\lambda = [\lambda_1, \lambda_c']'$, where $\lambda_c' = [\lambda_2, \cdots, \lambda_N]'$ is an N-dimensional loading vector and $\varepsilon_t \sim i.i.d.N(0, I_n)$ is independent of f_t. In this model, the long-run squared correlation can be expressed as

$$R^2 = \frac{\sigma_f^4 \lambda_1^2 (\lambda_c' \lambda_c)}{(1 + \sigma_f^2 \lambda_1^2) \left\{ 1 + \sigma_f^2 (\lambda_c' \lambda_c) \right\}}.$$

In order to control the value of R^2, we set $\lambda_1 = 1$ and $\lambda_c' \lambda_c = N - 1$ and choose the value of σ_f^2 so that $R^2 = 0.3$, 0.6, and 0.9. More precisely, we first generate $\lambda_i^* \sim U(0.5,\ 1.5)$ for $i = 2, \cdots, N$ independently and normalize them as $\lambda_c = \sqrt{(N-1)/(\lambda_c^{*\prime} \lambda_c^*)} \lambda_c^*$, where $\lambda_c^* = [\lambda_2^*, \cdots, \lambda_N^*]'$, so that the restriction $\lambda_c' \lambda_c = N - 1$ holds. In this case, σ_f^2 is the positive solution of the quadratic function of k given by $(N-1)(1-R^2)k^2 - NR^2 k - R^2 = 0$.

Tables 3 and 4 report the empirical sizes and powers of the tests in the constant case and the trend case, respectively. For DGP2 with $N = 5$, the size distortion of the OPOC test with the factor model rule is mitigated compared to DGP1, but still it suffers from serious oversize distortion. Again, we can see from the tables that our selection rules work better than the OPOC test with the conventional rule of selecting only one covariate; the powers of the tests are improved by using our selection rules.

5. Empirical Application

5.1. Prebish-Singer hypothesis

We recall that the Prebish-Singer (PS) hypothesis states that real commodity prices follow a downward secular trend. Prebish (1950) and Singer (1950) claimed that there had been a downward long-term trend in these prices and that the decline in these prices was likely to continue. The main theoretical explanations given for this negative long-term trend are: (a) income elasticities of demand for primary commodities that are lower than those for manufactured commodities; (b) an absence of differentiation among commodity producers leading to highly competitive markets; (c) productivity differentials between North and South; (d) asymmetric market structures: the presence of oligopolistic rents for the North and zero economic profit for competitive commodity producers in the South; (e) the inability of wages to grow in the presence of an "unlimited" supply of labor at the subsistence wage in primary commodity-producing countries (Lewis, 1954), and (f) a decline of demand from industrial countries. However, recently, this effect has been lessened by the growing demand from emerging market countries such as China, India, and Brazil. The consequences of this hypothesis are very important for developing countries because many of them depend on only a few primary commodities to generate most of their export earnings. If we assume that y_{it}, the real commodity price i, is generated by a stationary process around a time trend (I(0)):

$$y_{it} = \beta_{i,0} + \beta_{i,1}t + u_{it}, \quad t = 1, \cdots, T, \tag{8}$$

where the random variable u_{it} is stationary with mean 0 and variance $\sigma^2_{i,u}$. The parameter of interest is the slope $\beta_{i,1}$ which is predicted to be negative under the PS hypothesis. However, if commodity prices were generated by a so called difference-stationary (DS or I(1)) model, implying that y_{it} is non-stationary, then

$$\Delta y_{it} = \beta_{i1} + \zeta_{it}, \quad t = 1, \cdots, T, \tag{9}$$

where ζ_{it} is stationary. It is now well known that if y_{it} is an I(1) process, then using equation (8) to test the null hypothesis $\beta_{i,1} = 0$ will result in severe size distortions, leading to a wrong rejection of the null when no trend is present, even asymptotically. Alternatively, if the true generating process is given by equation (8) and we base our test on equation (9), then our test becomes inefficient and less powerful than the one based on the correct equation. Therefore, when testing the PS hypothesis we must first test the order of integration of our relative commodity prices in order to use the right equation.

In this subsection, we investigate the I(0)/I(1) properties of nine real commodity prices (zinc, tin, oil, wool, iron, aluminum, beef, coffee, and cocoa) relative to the U.S. CPI index using the annual data from 1960 to 2007. We first treat the data set as panel data and apply the PANIC test by Bai and Ng (2004). We assume the common factor structure in $u_{i,t}$ such that $u_{it} = \lambda_i f_t + \varepsilon_{it}$, where f_t is an r-dimensional common factor, λ_{it} is a $1 \times r$ loading vector, and ε_{it} is an idiosyncratic error, and we estimate ε_{it} by the method of principal component analysis. We then apply the Fisher test and the inverse normal test to the estimates of ε_{it}. The results are given in Panel (a) in Table 5. The number of common factors is estimated as 4 by the HQ_4 proposed by de Silva, Hadri, and Tremayne (2009). Both tests reject the null hypothesis of a unit root for the idiosyncratic errors using the size-adjusted critical values. We also apply the panel trend stationarity tests ZA_{spc} and ZA_{la} proposed by Hadri and Kurozumi (2012) and the test by Harris, Leybourne, and McCabe (2005). The results of the tests are consistent and imply that some of the prices can be characterized as trend stationary processes. However, they do not tell us which prices are trend stationary.

We next conduct univariate tests. We test for the null hypothesis of a unit root for each price by the ADF-GLS test with the lag length selected by the modified AIC by Ng and Perron (2001), while the null of trend stationarity is checked by the bias-corrected version of the KPSS test with the boundary condition equal to 0.95, which was developed by Kurozumi and Tanaka (2010) by correcting the bias in the test statistic previously proposed by Kwiatkowski, Phillips, Schmidt, and Shin (1992). The results are given in the second and third columns in Panel (b) in Table 5. We find strong evidence of stationarity for the prices of wool and aluminum. The bias-corrected version of the KPSS test rejects the null of trend stationarity for the price of tin at the 5% significance level and for the prices of petroleum, tin, and beef at the 10% significance level. However, we should carefully interpret the results of the KPSS test because this test is known to suffer from size distortions when a process is strongly serially correlated.

We next apply the OPOC test with the asymptotic power rule to each of the prices except for wool and aluminum, which have already been found to be I(0). Because the covariates must be stationary, we take the first difference of the prices except for those of wool and aluminum when using these variables as covariates. If we reject the null of a unit root for some of the prices, then we treat those variables as trend stationary, use them in levels as covariates, and test the other variables again. We repeat the procedure until we cannot find additional evidence of stationarity. The final results are given in Panel (b) in Table 5. The fourth column reports the number of covariates chosen by our selection rule, while the fifth column reports the estimated long-run squared correlation when those covariates are used for testing. From the results given in the sixth column, we can reject the null of a unit root for five prices: zinc, tin, petroleum, iron and coffee. In addition to the prices of wool and aluminum for which we have already rejected the null of a unit root, we find that seven of the nine commodities have trend-stationary prices, although they might be very persistent as shown, inter alia, by Cuddington and Jerret (2008). For the prices of beef and cocoa, we cannot reach a concluding result. Reflecting these results, we test the PB hypothesis based on equation (8) for the seven trend-stationary prices, whereas we estimate the slope for beef and cocoa using (9). The final column reports the p-values of the one-sided tests based on the t-statistics calculated using the autocorrelation-heteroskedasticity consistent standard errors using the quadratic spectral kernel with the bandwidth selected by the method proposed by Andrews (1991). We can see that except for the price of petroleum, whose coefficient is positive, all the estimates of the slope coefficients are significant and negative. We find strong evidence of the PB hypothesis for seven commodity prices and the weak evidence for cocoa; the exception is the price of petroleum.

We also investigated the same data with a shorter sample period from 1960 to 2002, because a structural change might occurred in the early 2000s, as pointed out by Arezki, Hadri, Kurozumi, and Rao (2012). The results of the tests are similar to the above case and we reach the same conclusion.

5.2. PPP hypothesis

The second data set we investigate consists of eight real exchange rates of developed countries and economies[4] relative to the U.S. dollar, spanning from January 2002 to December 2011: the Canadian dollar (CAD), the Danish krone (DKK), the euro (EUR), the Japanese yen (JPY), the Norwegian krone (NOK), the Swedish krona (SEK), the Swiss franc (CHF), and the U.K. pound (GBP). We divide the nominal exchange rates by the CPI indices (the harmonized index for the euro) and take their logarithms. The PPP hypothesis is supported if the series are stationary.

The results are given in Table 6. The PANIC test rejects the null hypothesis of a unit root for the estimated idiosyncratic errors, but the panel stationarity tests reject the null of stationarity, as reported in Panel (a). This happens when some of the cross-sectional units are stationary but the others have a unit root. Then, from the panel tests, we can say that the PPP hypothesis does not hold for some currencies.

We next apply the univariate tests. As seen in Panel (b), the ADF-GLS test does not reject the null of a unit root for any currencies, while the bias corrected version of the KPSS test strongly rejects the null of stationarity for the CAD, DKK, and EUR at the 1% significance level, which implies that the PPP hypothesis does not hold for them. The KPSS test also rejects the null hypothesis for the NOK and CHF at the 5% significance level and for the SEK at the 10% level. We then test for the null of a unit root for the five currencies using the OPOC test, but we cannot reject the null of a unit root for any of them. As a whole, we can say that the PPP hypothesis does not hold in this sample period for the CAD, DKK, and EUR; the NOK and CHF may not be stationary either. On the other hand, we cannot reach a conclusion for the JPY, SEK, and GBP; they might be stationary in this period, but if so,

[4]The data set is obtained from International Financial Statistics. We chose these currencies because the monthly series for both the nominal exchange rates and the CPI are available for these currencies.

they are very persistent, and it is difficult to distinguish between I(0) and I(1). These results are found to be robust to the Lehman shock by using the sample period before September 2008.

6. Concluding Remarks

In this paper, we have investigated the covariate unit root tests. Our analysis has revealed that the optimal point optimal covariate unit root test by Juhl and Xiao (2003) has potentially good power compared to panel unit root tests when N is not so large. Therefore, we can rely on panel unit root tests for data sets with a moderate to large number of cross-sectional units N, whereas the covariate tests may be useful for panels with a relatively small N. When, in panel data the null hypothesis of a unit root is rejected, we propose the use of the OPOC test and the information contained in the cross-sectional dependence to ascertain which cross-section variable is stationary and which one is not. We have also suggested three selection rules for choosing potential covariates. The Monte Carlo simulations have shown that the rules based on asymptotic power and the adjusted R^2 work fairly well, whereas the common factor rule must be used with caution.

We have considered the covariate unit root tests for panel data with only one type of variables generated by the cross-sectional dependency. But, if we can find other type of covariates, then it would be better to include those series as covariates when applying our selection rules. Moreover, in such a situation, we may be able to consider a panel version of the covariate unit root test as considered by Chang and Song (2009) and Westerlund (2012). In any case, combining the use of the covariate test and panel tests could be helpful in applications as shown above.

References

[1] Amara, J. and D. H. Papell (2006). Testing for Purchasing Power Parity Using Stationary Covariates. *Applied Financial Economics* 16, 29-39.

[2] Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59, 817-858.

[3] Arezki, R., K. Hadri, E. Kurozumi and Y. Rao (2012). Breaking the Prebish Singer Hypothesis Using Panel Data Stationarity Tests. Forthcoming in *Economics Letters*.

[4] Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica* 71, 135-171.

[5] Bai, J. and S. Ng (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70, 191-221.

[6] Bai, J. and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica* 72, 1127-1177.

[7] Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* . Chichester, Wiley.

[8] Breitung, J. and Pesaran, M. H. (2008). Unit roots and cointegration in panels. In: L. Matyas and P. Sevestre (eds): *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice. Dordrecht:* Kluwer Academic Publishers, 3rd edition, Chapt. 9, pp. 279-322.

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

[10] Chang, Y. and W. Song (2009). Testing for Unit Roots in Small Panels with Short-Run and Long-Run Cross-Sectional Dependencies *Review of Economic Studies* 76, 903-935.

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

[12] Christiano, L. J., M. Eichenbaum and R. Vigfusson (2003). What Happens After a Technology Shock? *NBER working paper* 769, 1-52.

[13] Christopoulos, D. K. and M. A. León-Ledesma (2008). Time-Series Output Convergence Tests and Stationary Covariates. *Economics Letters* 101, 297-299.

[14] Cox, D. R. and D. V. Hinkley (1974). *Theoretical Statistics*. Chapman and Hall, London.

[15] Cuddington, J. and D. Jerret (2008). Super Cycles in Metal Prices? ÎMF staff Papers 55(4), 541-565.

[16] de Silva, S., Hadri, K. and Tremayne, A. R. (2009). Panel Unit Root Tests in the Presence of Cross-Sectional Dependence: Finite Sample Performance and an Application. *Econometrics Journal* 12, 340-366.

[17] Dickey, D. A. and W. A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association* 74, 427-431.

[18] Elliott, G. and M. Jansson (2003). Testing for Unit Roots with Stationary Covariates. *Journal of Econometrics* 115, 75-89.

[19] Elliott, G. and E. Pesavento (2006). On the Failure of Purchasing Power Parity for Bilateral Exchange Rates after 1973. *Journal of Money, Credit and Banking* 38, 1405-1430.

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

[21] Fossati, S. (2013). Unit Root Testing with Stationary Covariates and a Structural Break in the Trend Function. *Journal of Time Series Analysis* DOI: 10.1111/jtsa.12020.

[22] Hadri, K. and E. Kurozumi (2012). A Simple Panel Stationarity Test in the Presence of Serial Correlation and a Common Factor. *Economics Letters* 115, 31-34.

[23] Hansen, B. E. (1995). Rethinking the Univariate Approach to Unit Root Testing. *Econometric Theory* 11, 1148-1171.

[24] Harris, D., S. Leybourne and B. McCabe (2005). Panel Stationarity Tests for Purchasing Power Parity with Cross-Sectional Dependence. *Journal of Business and Economic Statistics* 23, 395-409.

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

[26] Juhl, T. and Z. Xiao (2003). Power Functions and Envelopes for Unit Root Tests. *Econometric Theory* 19, 240-253.

[27] Kurozumi, E. (2003). Some Properties of the Point Optimal Invariant Test for the Constancy of Parameters. *Journal of the Japan Statistical Society* 33, 169-180.

[28] Kurozumi, E. and S. Tanaka (2010). Reducing the Size Distortion of the KPSS Test. *Journal of Time Series Analysis* 31, 415-426.

[29] Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159-178.

[30] Lee, C. F. and C. C. Tsong (2011). Covariate Selection for Testing Purchasing Power Parity. *Applied Economics* 43:15, 1923-1933.

[31] Lewis, A. (1954). Economic Development with Unlimited Supplies of Labour. *Manchester School of Economics and Social Studies* 22, 139-191.

[32] 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, 631-52.

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

[34] Ng, S. and P. Perron (2001). Lag Length Selection and the Construction of Unit Root Tests With Good Size and Power. *Econometrica* 69, 1519-1554.

[35] O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics 44,* 1-19.

[36] Pesaran, M. H. (2007). A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence. *Journal of Applied Econometrics* 22, 265-312.

[37] Pesaran, M. H. (2012). On the Interpretation of Panel Unit Root Tests. *Economics Letters*, forthcoming.

[38] Prebish, R. (1950). The Economic Development of Latin America and Its Principal Problems. *Economic Bulletin for Latin America* 7, 1-12.

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

[40] Singer, H. (1950). The Distribution of Gains between Investing and Borrowing Countries. *American Economic review, Papers and Proceedings* 40, 473-485.

[41] Tanaka, S. and E. Kurozumi (2012).Investigating Finite Sample Properties of Estimators for Approximate Factor Models When N is small. *Economics Letters* 116, 465-468.

[42] Westerlund, J. (2012). Rethinking the Univariate Approach to Panel Unit Root Testing: Using Covariates to Resolve the Incidental Trends Problem. manuscript, Deakin University.

[43] Westerlund, J. (2013). A Computationally Convenient Unit Root Test with Covariates, Conditional Heteroskedasticity and Efficient Detrending. *Journal of Time Series Analysis* DOI: 10.1111/jtsa.12025.

Table 1: Empirical size and power (constant case, DGP1)

N	R^2	ρ_1	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$
					$T=100$					$T=200$		
5	0.3	1.00	0.046	0.077	1.000	0.080	0.080	0.059	0.065	1.000	0.062	0.063
		0.98	0.116	0.202	1.000	0.219	0.221	0.267	0.322	1.000	0.359	0.358
		0.96	0.242	0.374	1.000	0.407	0.413	0.542	0.706	1.000	0.759	0.765
		0.94	0.382	0.537	1.000	0.579	0.586	0.824	0.912	1.000	0.948	0.948
		0.92	0.532	0.691	1.000	0.731	0.739	0.925	0.983	1.000	0.989	0.989
		0.90	0.665	0.805	1.000	0.828	0.835	0.960	0.996	1.000	0.998	0.998
	0.6	1.00	0.046	0.060	1.000	0.065	0.067	0.059	0.056	1.000	0.061	0.061
		0.98	0.116	0.243	1.000	0.270	0.270	0.267	0.420	1.000	0.537	0.544
		0.96	0.242	0.450	1.000	0.525	0.532	0.542	0.844	1.000	0.920	0.922
		0.94	0.382	0.652	1.000	0.734	0.742	0.824	0.975	1.000	0.990	0.991
		0.92	0.532	0.803	1.000	0.869	0.876	0.925	0.995	1.000	1.000	1.000
		0.90	0.665	0.911	1.000	0.948	0.950	0.960	0.999	1.000	1.000	1.000
	0.9	1.00	0.046	0.061	1.000	0.074	0.074	0.059	0.052	1.000	0.056	0.056
		0.98	0.116	0.462	1.000	0.562	0.563	0.267	0.852	1.000	0.940	0.940
		0.96	0.242	0.843	1.000	0.911	0.913	0.542	0.990	1.000	0.999	0.999
		0.94	0.382	0.971	1.000	0.985	0.987	0.824	1.000	1.000	1.000	1.000
		0.92	0.532	0.989	1.000	0.995	0.995	0.925	1.000	1.000	1.000	1.000
		0.90	0.665	0.994	1.000	1.000	1.000	0.960	1.000	1.000	1.000	1.000
10	0.3	1.00	0.063	0.073	0.072	0.076	0.070	0.069	0.072	0.057	0.062	0.066
		0.98	0.131	0.200	0.197	0.195	0.200	0.232	0.300	0.338	0.350	0.359
		0.96	0.251	0.362	0.365	0.371	0.390	0.546	0.682	0.711	0.739	0.744
		0.94	0.405	0.535	0.550	0.551	0.569	0.801	0.889	0.914	0.920	0.920
		0.92	0.555	0.681	0.701	0.710	0.702	0.906	0.969	0.973	0.977	0.974
		0.90	0.643	0.805	0.805	0.817	0.813	0.950	0.995	0.994	0.999	0.998
	0.6	1.00	0.063	0.060	0.058	0.062	0.067	0.069	0.056	0.057	0.054	0.055
		0.98	0.131	0.207	0.228	0.239	0.244	0.232	0.389	0.503	0.492	0.491
		0.96	0.251	0.415	0.500	0.535	0.531	0.546	0.797	0.896	0.902	0.907
		0.94	0.405	0.624	0.739	0.738	0.738	0.801	0.953	0.980	0.979	0.977
		0.92	0.555	0.785	0.862	0.876	0.872	0.906	0.989	0.996	0.995	0.995
		0.90	0.643	0.864	0.933	0.934	0.933	0.950	0.998	0.999	1.000	1.000
	0.9	1.00	0.063	0.054	0.057	0.058	0.058	0.069	0.049	0.043	0.048	0.048
		0.98	0.131	0.377	0.556	0.516	0.517	0.232	0.806	0.936	0.906	0.906
		0.96	0.251	0.775	0.908	0.878	0.881	0.546	0.986	0.996	0.993	0.993
		0.94	0.405	0.931	0.985	0.977	0.977	0.801	0.998	1.000	0.999	0.999
		0.92	0.555	0.984	0.997	0.995	0.995	0.906	1.000	1.000	1.000	1.000
		0.90	0.643	0.994	1.000	1.000	1.000	0.950	1.000	1.000	1.000	1.000

Table 2: Empirical size and power (trend case, DGP1)

N	R^2	ρ_1	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$
					$T = 100$					$T = 200$		
5	0.3	1.00	0.041	0.065	1.000	0.068	0.069	0.036	0.053	1.000	0.063	0.063
		0.98	0.044	0.086	1.000	0.104	0.106	0.066	0.121	1.000	0.151	0.151
		0.96	0.072	0.138	1.000	0.162	0.163	0.171	0.288	1.000	0.360	0.361
		0.94	0.110	0.213	1.000	0.255	0.254	0.347	0.531	1.000	0.637	0.636
		0.92	0.166	0.307	1.000	0.366	0.366	0.547	0.752	1.000	0.818	0.818
		0.90	0.241	0.397	1.000	0.480	0.480	0.720	0.873	1.000	0.924	0.924
	0.6	1.00	0.041	0.061	1.000	0.072	0.074	0.036	0.056	1.000	0.067	0.068
		0.98	0.044	0.112	1.000	0.148	0.149	0.066	0.193	1.000	0.279	0.278
		0.96	0.072	0.206	1.000	0.289	0.292	0.171	0.484	1.000	0.682	0.684
		0.94	0.110	0.342	1.000	0.465	0.468	0.347	0.777	1.000	0.897	0.898
		0.92	0.166	0.492	1.000	0.609	0.614	0.547	0.937	1.000	0.968	0.970
		0.90	0.241	0.611	1.000	0.737	0.744	0.720	0.972	1.000	0.989	0.989
	0.9	1.00	0.041	0.064	1.000	0.076	0.077	0.036	0.045	1.000	0.063	0.063
		0.98	0.044	0.285	1.000	0.371	0.379	0.066	0.596	1.000	0.799	0.801
		0.96	0.072	0.605	1.000	0.748	0.752	0.171	0.961	1.000	0.987	0.988
		0.94	0.110	0.845	1.000	0.931	0.933	0.347	0.994	1.000	0.999	0.999
		0.92	0.166	0.954	1.000	0.978	0.978	0.547	0.998	1.000	1.000	1.000
		0.90	0.241	0.981	1.000	0.993	0.993	0.720	1.000	1.000	1.000	1.000
10	0.3	1.00	0.048	0.057	0.051	0.055	0.059	0.041	0.055	0.046	0.053	0.053
		0.98	0.049	0.076	0.089	0.108	0.108	0.086	0.118	0.128	0.145	0.147
		0.96	0.073	0.118	0.158	0.167	0.170	0.177	0.277	0.359	0.333	0.341
		0.94	0.123	0.191	0.243	0.230	0.231	0.324	0.505	0.600	0.601	0.602
		0.92	0.180	0.289	0.356	0.347	0.345	0.538	0.709	0.799	0.800	0.799
		0.90	0.265	0.382	0.471	0.443	0.448	0.727	0.841	0.899	0.907	0.907
	0.6	1.00	0.048	0.055	0.052	0.055	0.058	0.041	0.049	0.046	0.058	0.058
		0.98	0.049	0.100	0.128	0.135	0.138	0.086	0.178	0.259	0.259	0.269
		0.96	0.073	0.196	0.248	0.272	0.274	0.177	0.448	0.654	0.624	0.635
		0.94	0.123	0.307	0.452	0.442	0.453	0.324	0.719	0.882	0.862	0.868
		0.92	0.180	0.424	0.612	0.617	0.616	0.538	0.875	0.968	0.958	0.959
		0.90	0.265	0.551	0.748	0.732	0.731	0.727	0.951	0.989	0.985	0.985
	0.9	1.00	0.048	0.042	0.050	0.061	0.059	0.041	0.052	0.044	0.058	0.058
		0.98	0.049	0.229	0.349	0.316	0.318	0.086	0.529	0.779	0.737	0.737
		0.96	0.073	0.535	0.763	0.697	0.699	0.177	0.930	0.987	0.978	0.978
		0.94	0.123	0.762	0.932	0.883	0.885	0.324	0.989	1.000	0.998	0.998
		0.92	0.180	0.905	0.982	0.971	0.971	0.538	0.999	1.000	1.000	1.000
		0.90	0.265	0.961	0.994	0.988	0.988	0.727	1.000	1.000	1.000	1.000

Table 3: Empirical size and power (constant case, DGP2)

N	R^2	ρ_1	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$
					$T = 100$					$T = 200$		
5	0.3	1.00	0.042	0.071	0.695	0.071	0.072	0.059	0.072	0.596	0.060	0.061
		0.98	0.119	0.182	0.722	0.176	0.180	0.252	0.341	0.682	0.361	0.361
		0.96	0.241	0.362	0.762	0.387	0.388	0.572	0.715	0.777	0.745	0.744
		0.94	0.388	0.543	0.799	0.577	0.576	0.833	0.926	0.837	0.944	0.950
		0.92	0.548	0.693	0.827	0.731	0.737	0.946	0.986	0.873	0.989	0.989
		0.90	0.672	0.810	0.859	0.831	0.830	0.978	0.998	0.903	0.997	0.997
	0.6	1.00	0.058	0.066	0.489	0.066	0.067	0.064	0.062	0.430	0.056	0.057
		0.98	0.138	0.200	0.532	0.224	0.233	0.260	0.448	0.573	0.526	0.529
		0.96	0.249	0.467	0.619	0.510	0.522	0.572	0.887	0.725	0.928	0.930
		0.94	0.401	0.683	0.683	0.744	0.750	0.845	0.979	0.786	0.987	0.985
		0.92	0.569	0.822	0.728	0.876	0.885	0.947	0.996	0.830	0.998	0.998
		0.90	0.700	0.914	0.774	0.938	0.939	0.981	0.999	0.862	1.000	0.999
	0.9	1.00	0.066	0.060	0.317	0.060	0.062	0.045	0.038	0.234	0.043	0.043
		0.98	0.146	0.479	0.366	0.557	0.562	0.257	0.882	0.516	0.939	0.939
		0.96	0.277	0.836	0.461	0.885	0.892	0.604	0.994	0.660	0.996	0.998
		0.94	0.418	0.956	0.536	0.982	0.985	0.846	1.000	0.688	1.000	1.000
		0.92	0.582	0.985	0.602	0.997	0.996	0.953	1.000	0.728	1.000	1.000
		0.90	0.735	0.994	0.635	0.999	0.999	0.984	1.000	0.780	1.000	1.000
10	0.3	1.00	0.053	0.073	0.073	0.073	0.078	0.064	0.064	0.057	0.067	0.066
		0.98	0.140	0.196	0.189	0.192	0.200	0.245	0.336	0.378	0.374	0.383
		0.96	0.262	0.342	0.363	0.360	0.364	0.582	0.716	0.729	0.729	0.737
		0.94	0.401	0.516	0.554	0.543	0.555	0.829	0.902	0.933	0.924	0.924
		0.92	0.548	0.670	0.730	0.701	0.697	0.937	0.985	0.977	0.984	0.983
		0.90	0.688	0.792	0.843	0.783	0.788	0.981	0.997	0.994	0.997	0.996
	0.6	1.00	0.058	0.061	0.058	0.064	0.064	0.060	0.054	0.058	0.066	0.064
		0.98	0.131	0.211	0.225	0.232	0.241	0.247	0.467	0.533	0.532	0.534
		0.96	0.264	0.436	0.508	0.514	0.519	0.592	0.842	0.901	0.895	0.904
		0.94	0.426	0.657	0.765	0.727	0.735	0.827	0.971	0.986	0.982	0.982
		0.92	0.588	0.812	0.893	0.860	0.858	0.947	0.995	0.999	0.998	0.998
		0.90	0.704	0.900	0.946	0.932	0.930	0.990	0.999	1.000	1.000	1.000
	0.9	1.00	0.062	0.069	0.060	0.066	0.074	0.044	0.045	0.050	0.059	0.059
		0.98	0.141	0.429	0.555	0.548	0.568	0.269	0.866	0.940	0.913	0.920
		0.96	0.241	0.820	0.929	0.904	0.897	0.588	0.991	0.999	0.997	0.997
		0.94	0.432	0.947	0.981	0.979	0.980	0.857	0.999	1.000	1.000	1.000
		0.92	0.599	0.987	0.996	0.993	0.991	0.946	1.000	1.000	1.000	1.000
		0.90	0.714	0.996	1.000	1.000	1.000	0.986	1.000	1.000	1.000	1.000

Table 4: Empirical size and power (trend case, DGP2)

N	R^2	ρ_1	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$	t^{gls}	Λ_1	Λ_{fac}	Λ_{pow}	$\Lambda_{\bar{R}^2}$
					$T = 100$					$T = 200$		
		1.00	0.033	0.050	0.715	0.056	0.056	0.053	0.053	0.599	0.060	0.059
		0.98	0.038	0.070	0.738	0.081	0.082	0.091	0.127	0.651	0.152	0.150
	0.3	0.96	0.060	0.112	0.756	0.131	0.129	0.190	0.315	0.711	0.355	0.356
		0.94	0.091	0.186	0.775	0.218	0.218	0.369	0.557	0.757	0.601	0.599
		0.92	0.147	0.272	0.802	0.314	0.313	0.552	0.755	0.802	0.807	0.807
		0.90	0.221	0.369	0.832	0.455	0.458	0.716	0.895	0.840	0.924	0.924
		1.00	0.033	0.058	0.510	0.062	0.065	0.043	0.060	0.436	0.058	0.058
		0.98	0.040	0.110	0.542	0.126	0.127	0.080	0.212	0.507	0.279	0.280
5	0.6	0.96	0.064	0.207	0.573	0.240	0.248	0.201	0.550	0.621	0.660	0.660
		0.94	0.089	0.325	0.618	0.440	0.446	0.373	0.817	0.697	0.901	0.902
		0.92	0.150	0.485	0.657	0.601	0.608	0.569	0.933	0.742	0.967	0.967
		0.90	0.226	0.639	0.701	0.731	0.735	0.744	0.978	0.779	0.991	0.992
		1.00	0.035	0.061	0.320	0.061	0.061	0.046	0.054	0.252	0.055	0.053
		0.98	0.038	0.288	0.366	0.352	0.360	0.082	0.693	0.414	0.791	0.798
	0.9	0.96	0.067	0.637	0.425	0.739	0.744	0.184	0.973	0.569	0.985	0.985
		0.94	0.106	0.862	0.488	0.907	0.911	0.385	0.998	0.609	0.999	1.000
		0.92	0.168	0.941	0.522	0.964	0.965	0.583	1.000	0.638	1.000	1.000
		0.90	0.259	0.972	0.570	0.988	0.988	0.749	1.000	0.696	1.000	1.000
		1.00	0.043	0.070	0.067	0.070	0.070	0.036	0.056	0.054	0.052	0.052
		0.98	0.059	0.095	0.098	0.102	0.101	0.076	0.116	0.153	0.151	0.153
	0.3	0.96	0.087	0.146	0.164	0.164	0.161	0.173	0.297	0.367	0.377	0.378
		0.94	0.138	0.188	0.238	0.244	0.239	0.362	0.521	0.606	0.601	0.598
		0.92	0.199	0.263	0.342	0.352	0.354	0.550	0.739	0.810	0.782	0.781
		0.90	0.266	0.366	0.468	0.457	0.455	0.734	0.879	0.914	0.898	0.901
		1.00	0.041	0.064	0.061	0.073	0.080	0.037	0.050	0.052	0.049	0.050
		0.98	0.060	0.117	0.134	0.141	0.148	0.072	0.204	0.287	0.287	0.283
10	0.6	0.96	0.086	0.196	0.279	0.289	0.288	0.172	0.523	0.668	0.633	0.641
		0.94	0.140	0.308	0.460	0.468	0.473	0.361	0.782	0.890	0.864	0.871
		0.92	0.202	0.454	0.636	0.616	0.616	0.565	0.916	0.972	0.962	0.962
		0.90	0.285	0.606	0.774	0.754	0.749	0.728	0.967	0.997	0.995	0.997
		1.00	0.042	0.068	0.071	0.085	0.087	0.034	0.060	0.056	0.053	0.052
		0.98	0.060	0.254	0.364	0.352	0.364	0.059	0.669	0.796	0.780	0.780
	0.9	0.96	0.093	0.610	0.779	0.736	0.740	0.166	0.955	0.993	0.988	0.990
		0.94	0.133	0.828	0.925	0.912	0.911	0.365	0.994	1.000	1.000	1.000
		0.92	0.199	0.935	0.971	0.970	0.968	0.564	1.000	1.000	1.000	1.000
		0.90	0.297	0.962	0.990	0.990	0.989	0.745	1.000	1.000	1.000	1.000

Table 5: Prebish-Singer hypothesis

(a) Panel unit root and stationarity tests

estimated # of the common factors:		4
PANIC test for the idiosyncratic errors		
	(Fisher test)	33.783**
	(inverse normal test)	−2.821***
panel stationarity tests		
	(ZA_{spc})	−1.892
	(ZA_{la})	−2.518
	(HLM test)	−0.707

(b) univariate unit root and stationarity tests

	ADF-GLS	KPSS	OPOC test			$t_{\beta_{i1}}$
			# of cov.	R^2	test stat.	p-value
zinc	−2.416	0.070	6	0.779	−0.143**	0.000
tin	−1.573	0.148**	2	0.954	−6.905**	0.000
petro.	−1.675	0.129*	4	0.981	−0.300**	1.000
wool	−3.467***	0.076	-	-	-	0.000
iron	−1.384	0.136*	5	0.862	−2.306**	0.000
aluminum	−3.905***	0.061	-	-	-	0.000
beef	−1.536	0.122*	5	0.998	7.797	0.000
coffee	−2.495	0.113	2	0.919	−3.388**	0.000
cocoa	−2.082	0.112	4	0.979	−0.273	0.058

Note: Rejections at 10%, 5%, and 1% significance level are denoted by *, **, and **, respectively.

Table 6: PPP hypothesis

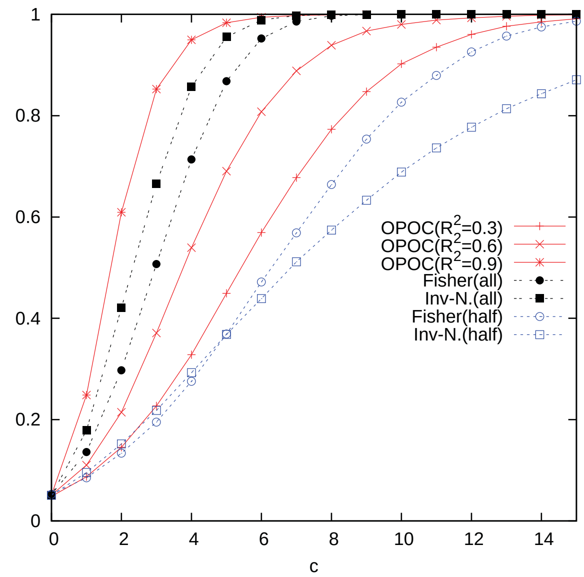
(a) Panel unit root and stationarity tests

estimated # of the common factors:		4
PANIC test for the idiosyncratic errors		
	(Fisher test)	55.017***
	(inverse normal test)	-4.697^{***}
panel stationarity tests		
	(ZA_{spc})	3.606***
	(ZA_{la})	1.008
	(HLM test)	3.261***

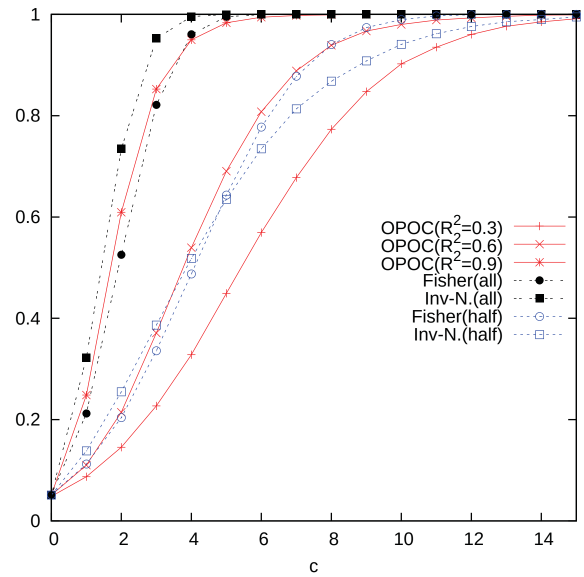
(b) univariate unit root and stationarity tests

	ADF-GLS	KPSS	OPOC test		
			# of cov.	R^2	test stat.
CAD	-0.037	0.968***	-	-	-
DKK	-0.502	0.771***	-	-	-
EUR	-0.320	0.751***	-	-	-
JPY	-0.370	0.298	1	0.310	14.768
NOK	-0.736	0.734**	1	0.813	9.263
SEK	-0.919	0.366*	1	0.860	8.949
CHF	-0.268	0.478**	1	0.786	23.249
GBP	-1.308	0.172	1	0.792	4.703

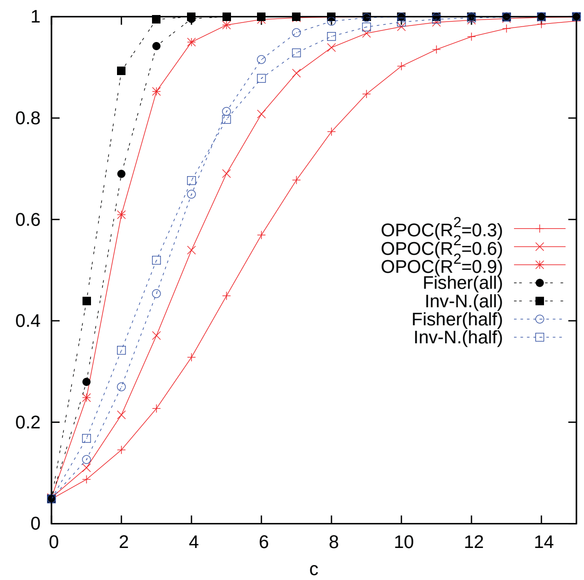
Note: Rejections at 10%, 5%, and 1% significance level are denoted by *, **, and **, respectively.



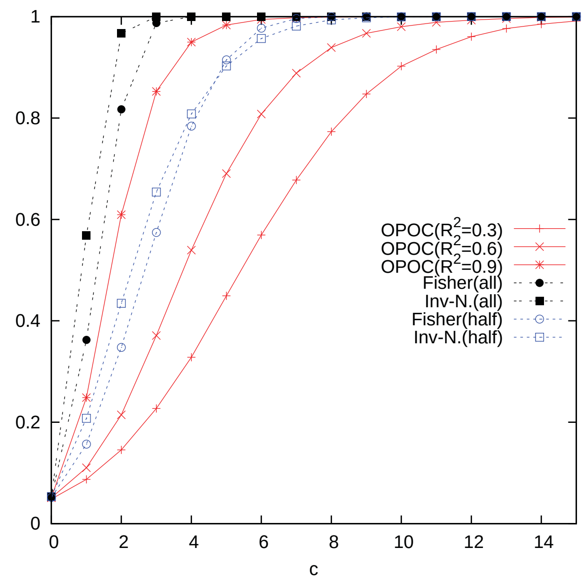
(i) $N = 4$



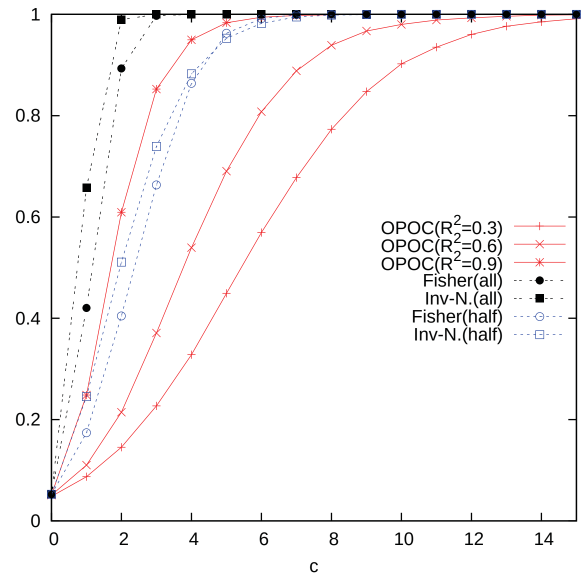
(ii) $N = 8$



(iii) $N = 12$

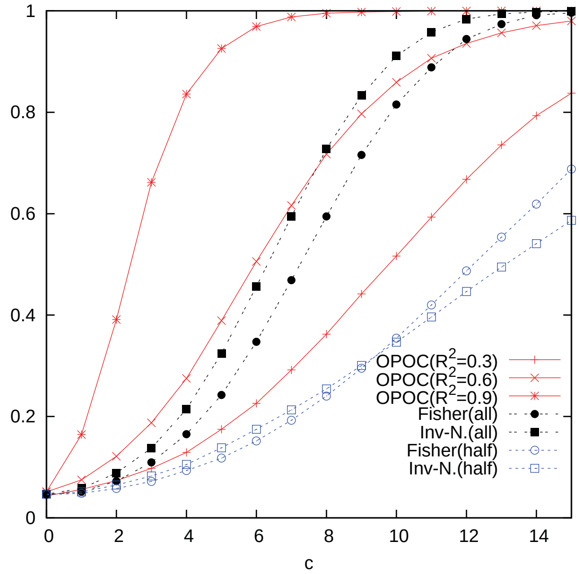


(iv) $N = 16$

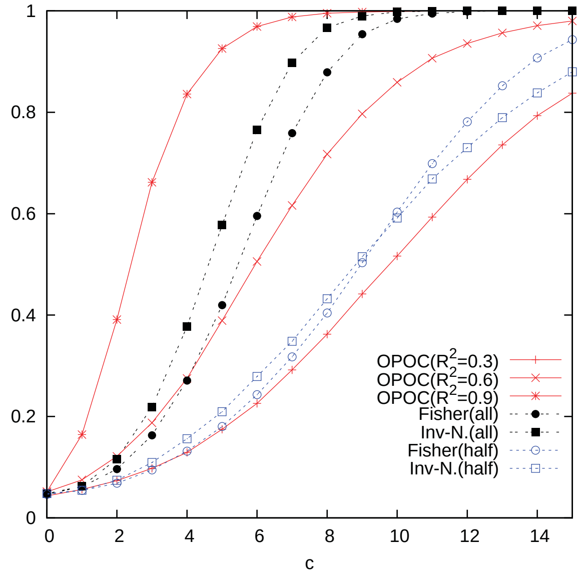


(v) $N = 20$

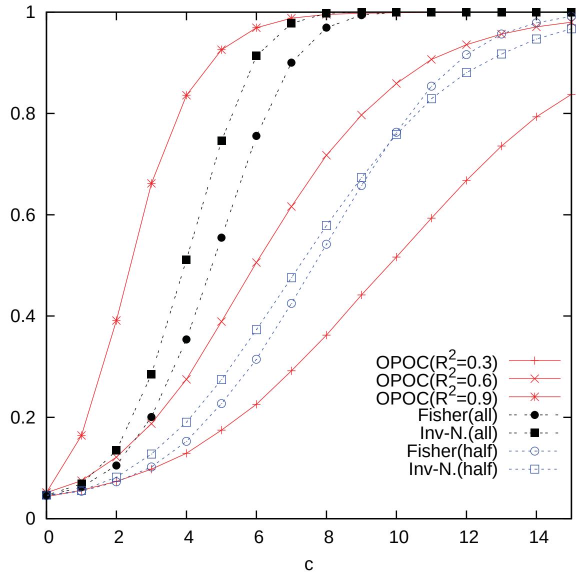
Figure 1: The limiting power functions (constant case)



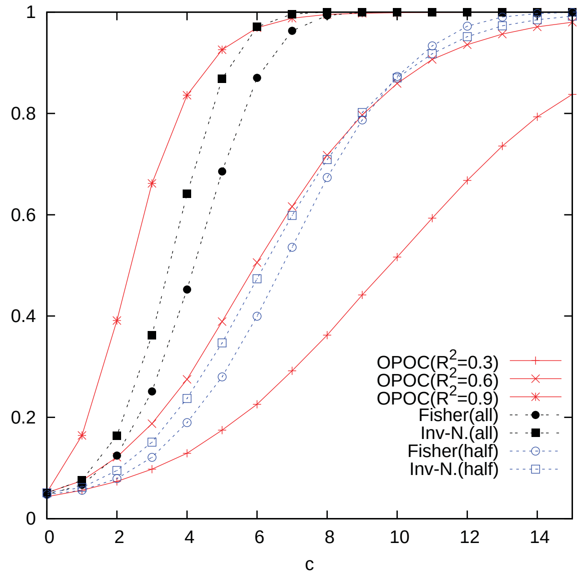
(i) $N = 4$



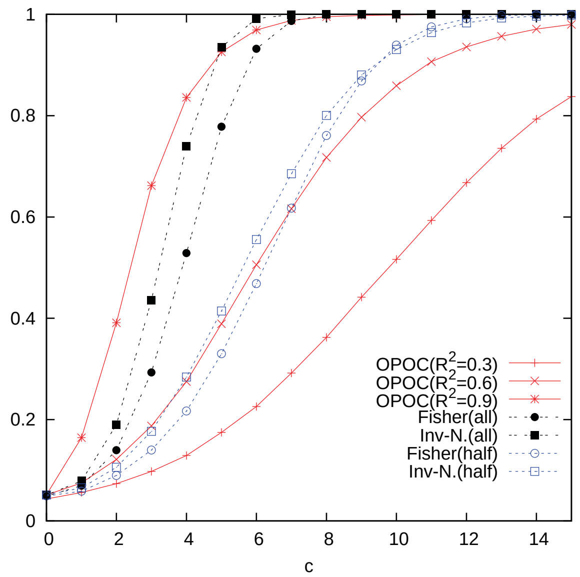
(ii) $N = 8$



(iii) $N = 12$



(iv) $N = 16$



(v) $N = 20$

Figure 2: The limiting power functions (trend case)

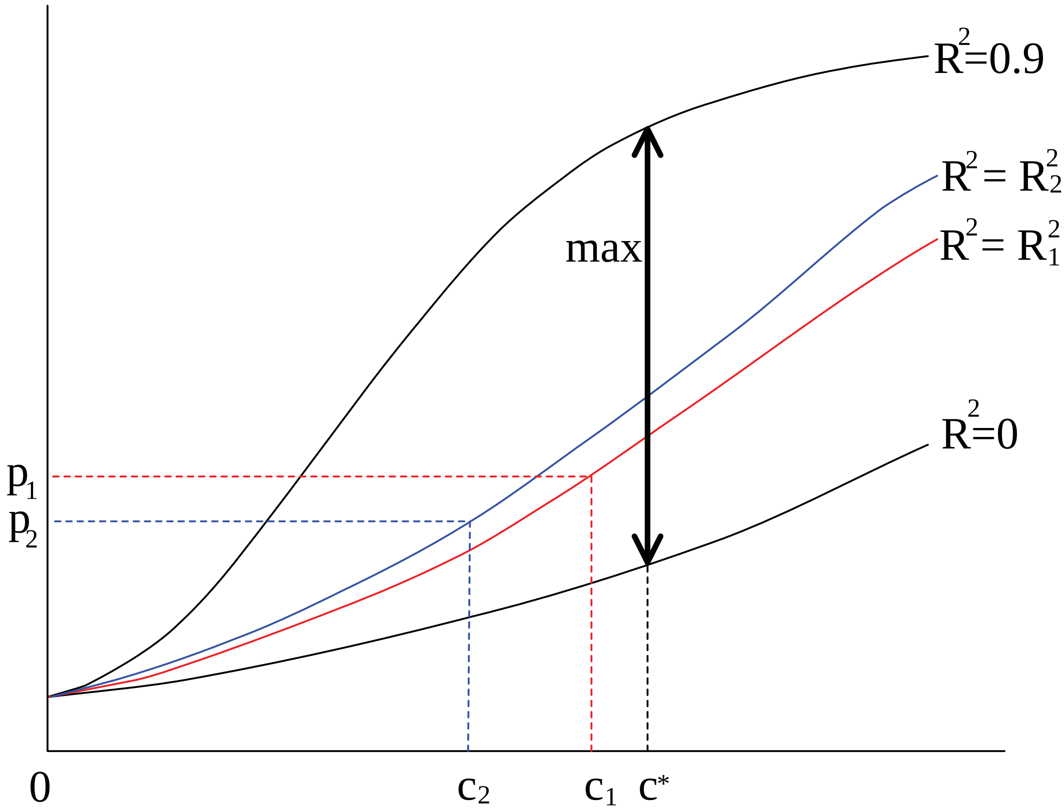


Figure 3: The selection rule of covariates