An Intersection Test for Panel Unit Roots

Christoph Hanck[*]

April 28, 2008

Preliminary.

Abstract

This paper proposes a new panel unit root test based on Simes' [*Biometrika* 1986, "An Improved Bonferroni Procedure for Multiple Tests of Significance"] classical intersection test. The test is robust to general patterns of cross-sectional dependence and yet straightforward to implement, only requiring p-values of time series unit root tests of the series in the panel, and no resampling. Monte Carlo experiments show good size and power properties relative to existing panel unit root tests. Unlike previously suggested tests, the new test allows to identify the units in the panel for which the alternative of stationarity can be said to hold. We provide two empirical applications to panels of real gross domestic product (GDP) and real exchange rate data.

Keywords: Multiple Testing, Panel Unit Root Test, Cross-Sectional Dependence

JEL classification: C12, C23

1 Introduction

The problem of testing for unit roots in heterogeneous panels has attracted much attention in recent years.[1] So called 'first generation' tests [Maddala and Wu, 1999; Im, Pesaran, and Shin, 2003; Levin, Lin, and Chu, 2002] rely on the assumption that the individual time series in the panel are cross-sectionally independent. It is, however, now widely recognized that this assumption is not met in typical macroeconometric panel data sets. For instance, common global shocks induce cross-sectional dependence among the test statistics [see, e.g., O'Connell, 1998].

The aim of 'second generation' panel unit root tests therefore is to provide reliable inference in the presence of cross-sectional dependence. Phillips and Sul [2003], Moon and Perron [2004], and Bai and Ng [2004] assume the dependence to be driven by (multiple) factors in the error terms. Suitably 'de-factoring' the observations, e.g. by the principal component method, asymptotically removes the common factors, then allowing for the application of standard panel unit root tests. Breitung and Das [2005], in turn, propose a feasible generalized least-squares approach that can be applied when $T > n$, where T denotes the number of time series observations on each of the n series. Pesaran [2007] adds the cross-section averages of lagged levels and of first-differences of the individual series to Augmented Dickey-Fuller 1979 (ADF) regressions. Panel unit root tests can then be based on the simple averages of the individual cross-sectionally augmented ADF statistics. Chang [2004] uses a bootstrap approach to condition on the dependence structure in the dataset. The approach most closely related to the one to be put forward here is by Demetrescu, Hassler, and Tarcolea [2006], who draw on the meta-analytic literature to derive their p-value combination tests.

In view of the additional complexities induced by the cross-sectional panel dimension, many of these tests require non-trivial decisions by the user, which complicate implementation and often impact the test outcome. These concern, for instance, the choice of the number of factors, of a joint lag length in pooled regressions, or of a suitable resampling scheme in bootstrap tests. In addition, as pointed out in the review article of Breitung

---

[1]Banerjee [1999] provides an early review.

and Pesaran [2008], available tests do not provide guidance as to the size of the fraction or the identity of the cross section units that are stationary. Hence, it can be difficult to properly interpret the rejection of a panel test.

Our goal therefore is to provide a new panel unit root test that alleviates these potential shortcomings. The test is based on Simes' [1986] classical intersection test of the 'global' null hypothesis H_0 that all individual null hypotheses $H_{i,0}$, $i = 1, \ldots, n$, are true. (Here, that all n time series are unit root processes.) Simes' [1986] test is widely applied in, among many other areas, genetical micro-array experiments [e.g., Dudoit, Shaffer, and Boldrick, 2003]. See Hochberg and Tamhane [1987] for a classical reference. The test is robust to general patterns of cross-sectional dependence and yet straightforward to implement, only requiring p-values of n time series unit root tests on the series in the panel. It compares the ordered p-values to suitably increasing critical points and rejects the panel unit root null whenever at least one of the p-values is smaller than the corresponding critical point.

Importantly, the new test allows to identify the units in the panel for which the alternative of stationarity appears to hold. Doing so, it still controls the 'Familywise Error Rate' (FWER), i.e. the probability to falsely reject at least one true individual time series null hypothesis, at some chosen level α. This would not be achieved by the widely applied strategy to reject for all those time series unit root tests statistics that exceed some fixed level-α critical value, as this latter approach ignores the multiple testing nature of the problem.[2]

The next section develops the new test. Section 3 reports results of a Monte Carlo study of the new test as well as some other popular panel unit root tests. Section 4 presents two applications of the tests to GDP data of OECD countries and to a panel of real exchange rates. The last section concludes.

---

[2]Recently, procedures taking multiplicity into account have begun to find their way into the econometrics literature. Romano and Wolf [2008] provide a survey of available methods and Hanck [2008] an application.

2 The Panel Unit Root Test

This section develops the Simes [1986]-type panel unit root test. We consider the following dynamic panel model:

$$y_{i,t} = \mu_i(1 - \phi_i) + \phi_i y_{i,t-1} + \epsilon_{i,t} \qquad (i \in \mathbb{N}_n,\ t \in \mathbb{N}_T), \tag{1}$$

where $j \in \mathbb{N}_a$ is shorthand for $j = 1, \ldots, a$, $\phi_i \in (-1, 1]$, $i \in \mathbb{N}_n$, and n denotes the number of series in the panel. Equation (1) says that the time series $\{y_{i,0}, \ldots, y_{i,T}\}$ are generated by a simple first-order autoregressive process for each cross section unit i. The panel unit root null hypothesis states that all time series are unit-root nonstationary [Breitung and Pesaran, 2008]. Formally,

$$H_0 : \phi_1 = \phi_2 = \ldots = \phi_n = 1$$

Put differently, H_0 states that *all* single time series hypotheses $H_{i,0} : \phi_i = 1$ are true,

$$H_0 = \bigcap_{i \in \mathbb{N}_n} H_{i,0}, \tag{2}$$

where $\bigcap_{i \in \mathbb{N}_n}$ denotes the intersection over the n individual time series hypotheses. Simes [1986] provides a simple test for testing the 'global' or 'intersection' null hypothesis (2). Suppose p-values p_i, $i \in \mathbb{N}_n$, of suitable test statistics for the individual time series hypotheses $H_{i,0}$ are available. Denote by $p_{(1)}, \ldots, p_{(n)}$ the ordered p-values $p_{(1)} \leqslant \ldots \leqslant p_{(n)}$. Then, Simes' intersection test (henceforth S test, for short) rejects H_0 at level α if and only if

$$p_{(j)} \leqslant j \cdot \alpha/n \quad \text{for some}\ \ j \in \mathbb{N}_n. \tag{3}$$

That is, one sorts the p-values from most to least significant and compares these to gradually less challenging critical points $j\alpha/n$. If there exists at least one p-value sufficiently small so as to be smaller than the corresponding critical point, the S test rejects the panel unit root null. Simes [1986, Thm. 1] proves that the S test has type I error probability equal to α when the test statistics are independent. As argued in the Introduction, the assumption of independence is unlikely to be met in most, if not all, applications of panel unit root tests. Fortunately, Sarkar [1998] shows that the assumption of independence is not necessary and can, in fact, be weakened substantially. The following is adapted from Sarkar [1998, Prop. 3.1]

PROPOSITION 1.

If the test statistics for testing the $H_{i,0}$, $i \in \mathbb{N}_n$, are multivariate totally positive of order 2 (MTP$_2$), then

$$P_{H_0}(\exists\, j \in \mathbb{N}_n : p_{(j)} \leqslant j\alpha/n) = P_{H_0}(S \ rejects) \leqslant \alpha,$$

where P_{H_0} denotes the probability under (2).

A vector of random variables $T = (T_1, \ldots, T_n)'$ is said to be MTP$_2$ if its joint density f satisfies

$$f\big(\min(T_1, U_1), \ldots, \min(T_n, U_n)\big) \cdot f\big(\max(T_1, U_1), \ldots, \max(T_n, U_n)\big) \geqslant$$
$$f(T_1, \ldots, T_n) \cdot f(U_1, \ldots, U_n),$$

for any two points $(T_1, \ldots, T_n)$ and $(U_1, \ldots, U_n)$. The MTP$_2$ class is rather large, including the multivariate normal with nonnegative correlations, the absolute-valued multivariate normal with some specific covariance structures, multivariate gamma, absolute-valued central multivariate t, and central multivariate F distributions. Sarkar [1998] further verifies that even the MTP$_2$ assumption of Proposition 1 is not necessary. The following Section shows that the S test controls size for patterns of cross-sectional dependence that are typically assumed in dynamic panel data models.

As emphasized for instance by Breitung and Pesaran [2008], it is important to properly interpret test outcomes if the panel unit root null is rejected. Some tests have the alternative of a completely stationary panel [e.g., Levin et al., 2002], i.e. $\max_i |\phi_i| < 1$, suggesting that a rejection of the null allows to conclude that the entire panel is stationary. However, these panel tests also have power against 'mixed' panels, where only some fraction of the units is actually stationary [see Taylor and Sarno, 1998; Boucher Breuer, McNown, and Wallace, 2001]. To caution against unwarranted conclusions, it may therefore be more appropriate to formulate panel unit root tests that have the more conservative alternative that some nonzero fraction of the panel is stationary,

$$H_A : |\phi_{(\ell)}|, \ \ell \in \mathbb{N}_{n_1}, \ n_1 \leqslant n$$

The S test belongs to this latter category, as is easily seen from the elementary relationship

$$H_A = H_0^c = \left(\bigcap_{i \in \mathbb{N}_n} H_{i,0}\right)^c = \bigcup_{i \in \mathbb{N}_n} H_{i,0}^c = \bigcup_{i \in \mathbb{N}_n} H_{i,A},$$

4

where c denotes the complement and $H_{i,A}$ the individual alternative $|\phi_i| < 1$.

Importantly, the S test is consistent, as $T \to \infty$.

LEMMA 1.
If, for any given $n < \infty$, $|\phi_i| < 1$ for at least one $i \in \mathbb{N}_n$, then $\lim_{T \to \infty} \mathrm{P}(S \text{ rejects}) = 1$.

Proof. Let I_A the index set of the wrong hypotheses. Under the maintained assumptions, $I_A \neq \varnothing$. Using consistent unit root test statistics to construct the p-values in S, there exists some p_j, $j \in I_A$, such that $p_j \equiv p_{(1)} \to_\mathrm{p} 0$. Hence, $\lim_{T \to \infty} \mathrm{P}(p_{(1)} < \alpha/n) = 1$. $\square$

Remark 1. Unlike other panel unit root tests [e.g., Im et al., 2003; Pesaran, 2007], the S test does not require that $n_1/n \to \kappa > 0$ as $n \to \infty$ for test consistency, see the proof of Lemma 1. It does, however, require $p_{(n_\ell)} = o_p(n_\ell/n)$ as $n, T \to \infty$, for some $\ell \in \mathbb{N}_{n_1}$. That is, the fraction of stationary series is allowed to tend to zero, but not too fast.

Existing panel unit root tests are not informative about the size of the fraction or the identity of the cross section units that are stationary. Using the p-values from the S test, one can easily determine those units in the panel for which the alternative can be said to hold, once the null hypothesis is rejected. Hommel [1988, Sec. 2] proves that the following procedure controls the FWER, i.e. the probability to falsely reject at least one true $H_{i,0}$, at multiple level α whenever the S test is a level-α test for the intersection hypothesis.

HOMMEL'S PROCEDURE

   A. Compute

$$j = \max\{i \in \mathbb{N}_n : p_{(n-i+k)} > k\alpha/i \ \text{ for } \ k \in \mathbb{N}_i\}. \tag{4}$$

  B1. If the maximum does not exist, reject all $H_{i,0}$ $(i \in \mathbb{N}_n)$.

  B2. If the maximum exists, reject all $H_{i,0}$ with $p_i \leqslant \alpha/j$.

For concreteness, consider an illustrative example where $n = 3$. We find j as the largest i such that all adjacent conditions in Table I hold. If j does not exist, then even $p_{(n)} \leqslant \alpha$, so that we can 'safely' reject all hypotheses. If the p-values are given by, say, $3\alpha/5$, 2α, $\alpha/5$, then $j = 2$ such that we only reject $H_{3,0}$.

5

TABLE I.—HOMMEL'S [1988] PROCEDURE FOR $n = 3$

$i = 1$:	$k = 1$	$p_{(3-1+1)} = p_{(n)} > \alpha$
$i = 2$:	$k = 1$	$p_{(3-2+1)} = p_{(n-1)} > \alpha/2$
	$k = 2$	$p_{(3-2+2)} = p_{(n)} > \alpha$
$i = 3$:	$k = 1$	$p_{(3-3+1)} = p_{(1)} > \alpha/3$
	$k = 2$	$p_{(3-3+2)} = p_{(n-1)} > 2\alpha/3$
	$k = 3$	$p_{(3-3+3)} = p_{(n)} > \alpha$

3 Monte Carlo Evidence

This section investigates the size and power of the S test discussed in the previous section. For comparison, we also include some other popular 'second generation' panel unit root tests.[3] More specifically, we consider the tests put forward by Demetrescu et al. [2006] $(t_{\hat{\rho}^*,\kappa})$, Moon and Perron [2004] $(t_a^*$ and $t_b^*)$, Pesaran [2007] ($CIPS^*$, denoted C^* here) and Breitung and Das [2005] (t_{rob}). (The Appendix provides a concise description of these tests.) To gauge the effect of cross-sectional dependence on first generation panel unit root tests, we also include the Fisher-type test of Maddala and Wu [1999], defined by $P_{\chi^2} = -2 \sum_{i=1}^{n} \ln(p_i) \overset{H_0}{\sim} \chi^2(2n)$.

We use the following simple data generating process:

$$
\begin{aligned}
y_{i,t} &= \mu_i + x_{i,t} \\
x_{i,t} &= \phi_i x_{i,t-1} + \epsilon_{i,t}
\end{aligned}
\qquad (i \in \mathbb{N}_n,\ t \in \mathbb{N}_T)
$$

We run the recursion for 30 initial observations before using the $y_{i,t}$ to mitigate the effect of initial conditions under H_A. We consider two different schemes to generate cross-sectional correlation among the error terms $\epsilon_{i,t}$.

A. *Equicorrelation*: Let $\boldsymbol{\varepsilon}_t = (\epsilon_{1,t}, \ldots, \epsilon_{n,t})'$. Then, $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \delta \boldsymbol{\imath}_n \boldsymbol{\imath}_n' + (1 - \delta) \boldsymbol{I}_n$ with $\boldsymbol{\imath}_n = (1, \ldots, 1)'$, $(n \times 1)$, and $\boldsymbol{I}_n$ the $(n \times n)$ identity matrix.

B. *Factor Structure*: $\epsilon_{i,t} = \lambda_i \cdot \nu_t + \xi_{i,t}$, where $\xi_{i,t}$ and ν_t are i.i.d. $\mathcal{N}(0,1)$ and $\lambda_i \sim \mathcal{U}(-1, 3)$, with $\mathcal{U}$ denoting the uniform distribution.

---

[3]See, e.g., Breitung and Pesaran [2008] or Gengenbach, Palm, and Urbain [2006] for comprehensive recent surveys of panel unit root tests.

TABLE II.—SIZE RESULTS, $\mu_i = 0$.

	n	8				12				24			
	T	30	50	100	200	30	50	100	200	30	50	100	200
						A. Equicorrelation.							
S		.045	.040	.039	.040	.040	.040	.035	.038	.037	.037	.034	.032
P_{χ^2}		.179	.181	.178	.172	.219	.230	.238	.232	.288	.275	.286	.282
C^*		.059	.057	.064	.059	.056	.068	.066	.063	.073	.071	.072	.073
t_{rob}		.115	.094	.092	.093	.057	.047	.043	.039	.060	.045	.043	.046
$t_{\hat{\rho}^*,\kappa}$		.067	.060	.062	.056	.067	.071	.064	.060	.067	.067	.063	.066
t_a^*		.110	.135	.156	.167	.042	.132	.152	.166	.027	.045	.133	.136
t_b^*		.092	.106	.108	.106	.053	.104	.116	.119	.038	.053	.110	.098
						B. Factor Structure							
S		.055	.056	.054	.048	.050	.055	.049	.052	.059	.053	.049	.052
P_{χ^2}		.068	.066	.066	.065	.084	.086	.088	.085	.123	.123	.125	.127
C^*		.043	.048	.050	.050	.045	.047	.049	.055	.054	.047	.052	.056
t_{rob}		.126	.103	.096	.096	.067	.052	.053	.049	.069	.060	.058	.045
$t_{\hat{\rho}^*,\kappa}$		.063	.061	.061	.060	.064	.066	.070	.064	.079	.075	.073	.074
t_a^*		.052	.079	.115	.145	.021	.045	.077	.105	.012	.028	.053	.078
t_b^*		.068	.077	.084	.094	.051	.062	.069	.080	.048	.055	.061	.068

Rejection rates of the panel unit root tests at nominal level $\alpha = 0.05$, using $M = 5000$ replications. In A, $\delta = 0.98$. S is Simes [1986] test, P_{χ^2} is by Maddala and Wu [1999], C^* is by Pesaran [2007], t_{rob} is from Breitung and Das [2005], $t_{\hat{\rho}^*,\kappa}$ is from Demetrescu et al. [2006] and t_a^* and t_b^* are by Moon and Perron [2004].

When $\boldsymbol{\phi}_n \equiv (\phi_1, \ldots, \phi_n)' = \boldsymbol{\imath}$, $H_0 = \bigcap_{i \in \mathbb{N}_n} H_{i,0}$ is true, allowing us to study the size of the tests. Choosing $\boldsymbol{\phi}_n$ such that $\min_i |\boldsymbol{\phi}_n| < 1$, we analyze power of the tests. More specifically, we let $\boldsymbol{\phi}_n = 0.95 \times \boldsymbol{\imath}$ to analyze a homogenous alternative and $\boldsymbol{\phi}_n = (\boldsymbol{\imath}'_{n/2}, \widetilde{\boldsymbol{\phi}}'_{n/2})'$, $\widetilde{\boldsymbol{\phi}}_n = (\boldsymbol{\imath}'_{3n/4}, \widetilde{\boldsymbol{\phi}}'_{n/4})'$ as well as $\boldsymbol{\phi}_n = \widetilde{\boldsymbol{\phi}}_n$ to investigate heterogeneous alternatives. The components of $\widetilde{\boldsymbol{\phi}}$ are distributed as $(\widetilde{\boldsymbol{\phi}})_i \sim \mathcal{U}(3/4, 1)$. When $\mu_i = 0$ for all $i \in \mathbb{N}_n$, we calculate n Dickey and Fuller [1979] τ-statistics from regressions of $y_{i,t}$ on $y_{i,t-1}$. To investigate the case with non-zero intercept, we simulate $\mu_i \sim \mathcal{U}[0, 0.2]$ and calculate τ-statistics from regressions of $y_{i,t}$ on 1, $y_{i,t-1}$.[4] The p-values are then calculated using MacKinnon [1994]-type response surface regressions.

The major findings of our experiments, reported in Tables II to IV, can be summarized as follows. The P_{χ^2}, being a first generation test, exhibits an expected size distortion. The

---

[4]The results with intercept, which we do not report for brevity but which are available upon request, are qualitatively similar to those to be reported below. As one would expect, fitting a constant results in loss of power for the tests considered.

TABLE III.—POWER RESULTS FOR PARTIALLY STATIONARY HETEROGENEOUS PANELS, $\mu_i = 0$, $\widetilde{\phi}_n = \left(\boldsymbol{\imath}'_{n/2}, \widetilde{\phi}'_{n/2}\right)'$ AND $(\widetilde{\phi}_{n/2})_i \sim \mathcal{U}(.75, 1)$.

	n	8				12				24			
	T	30	50	100	200	30	50	100	200	30	50	100	200
					A. Equicorrelation								
S		.158	.319	.835	1.000	.152	.339	.875	1.000	.118	.257	.801	1.000
C^*		.147	.170	.249	.313	.151	.192	.244	.315	.142	.171	.229	.284
t_{rob}		.689	.914	.995	1.000	.738	.943	.999	1.000	.642	.903	.998	1.000
$t_{\hat{\rho}^*,\kappa}$		.293	.512	.888	.987	.333	.591	.903	.985	.255	.449	.842	.962
t_a^*		.244	.282	.295	.289	.177	.348	.372	.420	.188	.226	.362	.363
t_b^*		.225	.251	.263	.245	.170	.330	.344	.399	.195	.227	.350	.343
					B. Factor Structure								
S		.211	.430	.943	1.000	.252	.547	.989	1.000	.202	.434	.974	1.000
C^*		.190	.259	.336	.386	.210	.272	.381	.459	.221	.276	.373	.474
t_{rob}		.252	.360	.640	.897	.204	.344	.697	.956	.176	.316	.676	.964
$t_{\hat{\rho}^*,\kappa}$		.411	.717	.975	.999	.595	.892	.992	1.000	.554	.845	.983	.999
t_a^*		.194	.263	.339	.364	.153	.253	.348	.386	.154	.253	.395	.463
t_b^*		.195	.232	.272	.279	.189	.253	.315	.329	.235	.293	.388	.430

Rejection rates of the panel unit root tests at nominal level $\alpha = 0.05$, using $M = 5000$ replications. In A, $\delta = 0.98$. S is Simes [1986] test, C^* is by Pesaran [2007], t_{rob} is from Breitung and Das [2005], $t_{\hat{\rho}^*,\kappa}$ is from Demetrescu et al. [2006] and t_a^* and t_b^* are by Moon and Perron [2004].

performance of the Moon and Perron [2004] tests is also not entirely satisfactory in many configurations. In particular, they perform poorly under equicorrelation (for which they are not designed, of course). The S test is a level-α test throughout, though it is slightly conservative under equicorrelation (panel A). Its size is very accurate in the factor model experiment (panel B). C^* and $t_{\hat{\rho}^*,\kappa}$ also perform very well throughout. The t_{rob} test controls size reasonably well, at least for T and n sufficiently large.

Some results for the power experiments are reported in Tables III and IV. First, we report results for $\boldsymbol{\phi}_n = (\boldsymbol{\imath}'_{n/2}, \widetilde{\boldsymbol{\phi}}'_{n/2})'$ in Table III. In view of its substantial upward size distortion, we do not report results for the P_{χ^2} test. The excellent performance of the Breitung and Das [2005] test in the heterogeneous scenario is noteworthy (though to some extent due to its upward size distortion), as it is designed for homogenous alternatives (see the Appendix). Similarly, the test by Demetrescu et al. [2006] delivers high power in the present setup. For small T, the power of the S test is somewhat lower than that of these

TABLE IV.—POWER RESULTS FOR PARTIALLY
STATIONARY HETEROGENEOUS PANELS, $\mu_i = 0$,
$\widetilde{\phi}_n = \left(\boldsymbol{\imath}'_{3n/4}, \widetilde{\phi}'_{n/4}\right)'$ AND $(\widetilde{\phi}_{n/4})_i \sim \mathcal{U}(.75, 1)$.

n		8				12				24		
T	30	50	100	200	30	50	100	200	30	50	100	200
					A. Equicorrelation							
S	.081	.128	.430	.959	.106	.216	.788	1.000	.080	.181	.753	1.000
C^*	.099	.107	.130	.145	.094	.100	.118	.134	.095	.104	.119	.135
t_{rob}	.688	.882	.990	1.000	.528	.765	.944	.996	.447	.722	.936	.995
$t_{\hat{\rho}^*,\kappa}$	.106	.167	.401	.708	.109	.169	.449	.699	.095	.146	.336	.685
t_a^*	.128	.223	.229	.196	.089	.207	.216	.228	.104	.108	.209	.213
t_b^*	.120	.189	.178	.144	.068	.158	.151	.160	.095	.086	.169	.152
					B. Factor Structure							
S	.107	.184	.543	1.000	.142	.273	.818	1.000	.125	.250	.816	.992
C^*	.095	.112	.137	.184	.092	.106	.141	.171	.094	.123	.156	.166
t_{rob}	.139	.167	.267	.456	.104	.130	.223	.430	.101	.125	.222	.482
$t_{\hat{\rho}^*,\kappa}$	.185	.300	.604	.880	.220	.379	.685	.891	.235	.397	.723	.849
t_a^*	.069	.113	.156	.190	.057	.103	.148	.183	.040	.079	.152	.196
t_b^*	.089	.113	.128	.165	.086	.116	.134	.144	.092	.124	.157	.145

Rejection rates of the panel unit root tests at nominal level $\alpha = 0.05$, using $M = 5000$ replications. In A, $\delta = 0.98$. S is Simes [1986] test, C^* is by Pesaran [2007], t_{rob} is from Breitung and Das [2005], $t_{\hat{\rho}^*,\kappa}$ is from Demetrescu et al. [2006] and t_a^* and t_b^* are by Moon and Perron [2004].

two tests, although it outperforms the tests by Pesaran [2007] and Moon and Perron [2004], whose power is rather disappointing here. That the S test has relatively lower power for small T is not surprising since the p-values cannot yet be close to their probability limit of 0 under the alternative. Hence, condition (3) is less likely to be satisfied. We feel that this may well be a price worth paying in view of the S test's additional ability to identify individual false hypotheses. Moreover, for large T, the power of the S test is highest among all tests considered here.

Table IV increases the fraction of nonstationary series to 3/4 and further corroborates the excellent performance of the S test. Of course, all tests are now uniformly less powerful. However, it is noteworthy that the loss of power relative to Table III is smallest for the S test; its performance is now the best among all tests considered for many practically relevant configurations.[5]

---

[5]Results for the remaining power experiments described above, which were qualitatively similar, are

4   Applications

We now present an application of the tests discussed here to two longstanding questions in empirical macroeconomics. First, we investigate the null hypothesis that there is a unit root in the (logarithms of) GDPs in a panel of OECD countries. Second, we test for stationarity in a panel of real exchange rate indices.

As should be clear from the previous discussion, any unit root test for which p-values are available can be used to compute the S test. In order to maximize the ability of the S test to reject a false null, preference should therefore be given to the most powerful time series unit root tests available. The test regressions in the present PPP and GDP applications are known to require an intercept and an intercept plus time trend, respectively. Thus, the power of the tests can be increased by efficient GLS demeaning and detrending, using the unit root tests proposed by Elliott, Rothenberg, and Stock [1996] (ERS). The limit distributions of the ERS test statistics are functionals of Brownian motions that cannot be evaluated analytically to obtain p-values. We therefore rely on response surface regressions suggested by MacKinnon [1994, 1996] to obtain numerical distribution functions of the test statistics.

4.1   GDP convergence

This subsection applies the panel unit root tests from the last section to logarithms of annual real per capita GDP data of OECD countries. The data are from the Penn World Tables 6.2.[6] Testing for a unit root in real GDPs is important, because acceptance of the null lends empirical support to certain real business-cycle models [see Nelson and Plosser, 1982]. Backus and Kehoe [1992] have established the stylized fact that real GDPs are

---

available upon request. The performance of the C^* test improves markedly in the other scenarios, as the average of the individual augmented ADF statistics is now of course more negative. Similar results obtain for the Moon and Perron [2004] tests. For near-integrated homogenous alternatives, the power of S test is relatively lower. This is unsurprising because we then have $p_{(1)} \approx p_{(n)}$, and it is unlikely that $p_{(n)} \leqslant \alpha$ for ϕ_n close to ι.

[6]We consider those countries for which a complete record of observations on the variable denoted RGDPCH from 1950 to 2004 is available, yielding a total of $n = 23$ series.

TABLE V.—PANEL UNIT ROOT TESTS FOR THE GDP
PANEL

		critical points
S	$p_{\text{Turkey}} = 0.0001$	0.002
t_{rob}	0.358	-1.64
$t_{\hat{\rho}^*,\kappa}$	4.082	-1.64
C^*	-2.674	-2.65

Nominal level is $\alpha = 0.05$. S [Simes' test] reports the p-value and corresponding critical point for which condition (3) holds. In the other rows, the last entry is the α level critical value. t_{rob} from Breitung and Das [2005], $t_{\hat{\rho}^*,\kappa}$ from Demetrescu et al. [2006] and C^* is by Pesaran [2007].

cross-sectionally dependent. Thus, it is expedient to employ panel unit root tests robust to cross-sectional correlation.

Table V reports the results of the panel unit root tests applied to the real GDP series.[7] In line with many findings in the literature, the tests by Demetrescu et al. [2006] and Breitung and Das [2005] do not reject the panel unit root null. The C^* test indicates some stationary series in the panel, as does the S test. Using the former, it would not be clear how pervasive the evidence of stationarity actually is. For the latter, an application of Hommel's [1988] method yields an individual rejection for Turkey only. Thus, according to the S test, we cannot reject the null hypothesis of a unit root in the real GDPs for the vast majority of OECD countries.

4.2 Purchasing Power Parity

We now apply the tests to the relative Purchasing Power Parity (PPP) hypothesis. Let $p_{i,t}$ be the (log) price index in country i and period t, p_t^* the 'foreign' (log) price index of the reference country in the panel and $s_{i,t}$ the (log) nominal exchange rate between the currencies of country i and the reference country. The real exchange rate is then given by

$$r_{i,t} = p_{i,t} - p_t^* - s_{i,t} \qquad (i \in \mathbb{N}_n, \ t \in \mathbb{N}_T)$$

---

[7]In view of the finding of Moon and Perron [2004] that their tests only have trivial (local) power in the presence of incidental trends we waive to report results for these.

TABLE VI.—DF-GLS$^{\mu}$-p-VALUES FOR THE REAL
EXCHANGE RATE SERIES

country	p-values	Simes criterion
Argentina	0.0001	0.0026
Sweden	0.0001	0.0053
Norway	0.0001	0.0079
Mexico	0.0001	0.0105
Italy	0.0001	0.0132
Finland	0.0001	0.0158
France	0.0050	0.0184
Germany	0.0050	0.0211
Belgium	0.0050	0.0237
UK	0.0050	0.0263
Brazil	0.0175	0.0289
Australia	0.0175	0.0316
Netherlands	0.0200	0.0342
Portugal	0.0250	0.0368
Canada	0.0400	0.0395
Spain	0.0500	0.0421
Denmark	0.0575	0.0447
Switzerland	0.2375	0.0474
Japan	0.2475	0.0500

'p-values' are the marginal significance levels from
DF-GLS$^{\mu}$ time series unit root tests [Elliott et al., 1996].
'Simes criterion' is the value that needs to exceed the adjacent p-value in at least one case for a rejection of H_0, i.e. (3).

Testing the PPP hypothesis is naturally formulated [see Rogoff, 1996] as a unit root test on the real exchange rate. We revisit the dataset used by Taylor [2002], which includes annual data for the nominal exchange rate, CPI and the GDP deflator. The countries contained in our panel are given in Table VI. We use the United States as the reference country and report results using CPI price series. See Taylor [2002] for further details on data sources and definitions.

We report results for the S test in Table VI. Obviously, condition (3) is satisfied, so that the panel unit root null can be rejected. Table VII reports similar findings for the other panel unit root tests considered here.

Applying Hommel's [1988] method now allows us to identify those countries for which the alternative of a real exchange rate can be said to hold, while controlling for multiplicity. Using the results from Table VI and (4) we find $j = 10$. Thus, we can reject $H_{i,0}$ for

TABLE VII.—PANEL UNIT ROOT TESTS FOR THE REAL
EXCHANGE RATE PANEL

		critical points
t_{rob}	-3.703	-1.64
$t_{\hat{\rho}^*,\kappa}$	-7.578	-1.64
$CIPS$	-3.084	-2.16
C^*	-3.077	-2.16
t_a^*	-0.032	-1.64
t_b^*	-0.290	-1.64

Nominal level is $\alpha = 0.05$. The last column gives
the α level critical value. t_{rob} from Breitung and
Das [2005], $t_{\hat{\rho}^*,\kappa}$ from Demetrescu et al. [2006],
$CIPS$ and C^* are by Pesaran [2007] and t_a^* and t_b^*
are from Moon and Perron [2004].

Argentina, Sweden, Norway, Mexico, Italy, Finland, France, Germany, Belgium and the
United Kingdom.

Let us compare these conclusions to those when employing the traditional strategy of
rejecting all those individual hypotheses whose corresponding p-value is less than α, which
ignores the multiple nature of the testing problem. From Table VI we see that this method
would have rejected all individual null hypotheses except for Denmark, Switzerland and
Japan. The results of the Hommel [1988] procedure suggest that several of the additional
rejections may be spurious, i.e. solely due to the fact that the traditional approach does
not control the FWER.

5   Conclusion

This paper proposes a new test for a panel unit root against the alternative of a partially
stationary panel, making use of Simes' [1986] classical test of the intersection null hypoth-
esis. The test is intuitive, straightforward to implement and yet robust to general patterns
of cross-sectional dependence. Monte Carlo simulations show that the new test compares
well in terms of finite sample size and power with recently proposed second generation
panel unit root tests. Importantly, unlike other tests, Simes' [1986] approach allows to
shed light on the important question for how many and also which of the units in the
panel the alternative can be said to hold when the null hypothesis is rejected. Hence, the

test suggested here allows researchers to decide, for each unit individually, whether it is advisable to model the respective time series in levels or first differences. An application to the questions of GDP convergence and real exchange rate stationarity demonstrates the practical usefulness of the test.

Since implementation of the test only requires p-values of individual time series test statistics, the framework put forward here is very flexible. For instance, it can be extended rather easily to test for panel cointegration, the analysis of which is still at an earlier stage of its development [Breitung and Pesaran, 2008]. As such, it could furthermore be applied to other macroeconomic questions such as savings-investment correlation or spot and forward exchange rates, that have hitherto been dealt with using other panel techniques. This issue is currently under investigation by the author.

Appendix

Here, we briefly sketch the panel unit root tests studied in section 3. For more detailed descriptions we refer to the original contributions.

Pesaran [2007]

The basic setup of Pesaran [2007] is given by

$$y_{i,t} = \mu_i(1 - \rho_i) + \rho_i y_{i,t-1} + \sum_{j=1}^{\ell_i} \phi_{ij} \Delta y_{i,t-j} + u_{i,t}, \tag{A.1}$$

where $t \in \mathbb{N}_T$, $i \in \mathbb{N}_n$ and $\rho_i \in (-1, 1]$. He models cross-sectional dependence via a standard normal common time effect θ_t which is allowed to affect the units of the panel heterogeneously:

$$u_{i,t} = \delta_i \theta_t + \epsilon_{i,t}. \tag{A.2}$$

Pesaran [2007] shows that augmenting the usual ADF regressions by cross-sectional means of the lagged level and of the lagged differences is sufficient for filtering out the effect of both serial correlation and of the common factor θ_t. That is, compute the t-statistic on b_i, $t_i(n, T)$, in

$$\Delta y_{i,t} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{i,t-1} + \sum_{j=0}^{p} d_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{p} \gamma_{ij} \Delta \bar{y}_{i,t-j} + \varepsilon_{i,t}$$

for all i. Next, a panel test statistic obtains as $C = n^{-1} \sum_{i=1}^{n} t_i(n, T)$. For our purposes, the relevant critical values lie between -1.8 for $(n, T) = (10, 10)$ and -1.51 for $(n, T) = (200, 200)$. Pesaran [2007], motivated by concerns about the finiteness of the moments of $t_i(n, T)$, also suggests a panel test using a truncated version of the $t_i(n, T)$, $C^* = n^{-1} \sum_{i=1}^{n} t_i^*(n, T)$. This test has the same critical values (in our application). Both tests reject for large negative values.

Moon and Perron [2004]

The tests put forward by Moon and Perron [2004] are able to accommodate patterns of cross-sectional dependence which are driven by more that one common unobserved factor. That is, the error from (A.2) in the panel model

$$y_{i,t} = \mu_i(1 - \rho_i) + \rho_i y_{i,t-1} + u_{i,t} \tag{A.3}$$

becomes

$$u_{i,t} = \boldsymbol{\lambda}_i' \boldsymbol{\theta}_t + \epsilon_{i,t}, \tag{A.4}$$

where $\boldsymbol{\lambda}_i$ and $\boldsymbol{\theta}_t$ are now $(K \times 1)$ vectors. Moon and Perron [2004] provide tests of the null

$$H_0 : \rho_i = 1 \text{ for all } i \quad \text{against} \quad H_A : \rho_i < 1 \text{ for some } i.$$

The algorithm to compute the tests is summarized as follows. First, run a pooled OLS regression on (A.3) under the null (setting $\mu_i(1 - \rho_i) = 0$ for all i). Use the residuals $\hat{u}_{i,t}$ to estimate the $(n \times K)$ matrix of factor loadings $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_n)'$ by principal components. Denote this estimate by $\hat{\boldsymbol{\Lambda}}$. We assume the number of factors to be known in our simulations, i.e. $K = 1$. Let $\hat{\boldsymbol{u}} = ((\hat{u}_{11}, \ldots, \hat{u}_{1,T})', \ldots, (\hat{u}_{n,1}, \ldots, \hat{u}_{n,T})')$. By projecting $\hat{\boldsymbol{u}}$ onto the column space orthogonal to the space spanned by the common factors, $\boldsymbol{Q}_{\hat{\boldsymbol{\Lambda}}} = \boldsymbol{I} - \hat{\boldsymbol{\Lambda}}(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}})^{-1}\hat{\boldsymbol{\Lambda}}'$, obtain de-factored residuals $\hat{\boldsymbol{\epsilon}} = \hat{\boldsymbol{u}}\boldsymbol{Q}_{\hat{\boldsymbol{\Lambda}}}$. Moon and Perron [2004] suggest the following adjusted pooled estimator

$$\hat{\rho}^*_{pool} = \frac{\text{tr}(\boldsymbol{Y}_{-1}\boldsymbol{Q}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Y}') - NT\hat{\varphi}^e_n}{\text{tr}(\boldsymbol{Y}_{-1}\boldsymbol{Q}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Y}'_{-1})}.$$

Here, tr denotes the trace and $\boldsymbol{Y}$ and $\boldsymbol{Y}_{-1}$ are defined similarly as $\hat{\boldsymbol{u}}$, while $\hat{\varphi}^e_n$ is an estimator of the average one-sided long-run variance of $\epsilon_{i,t}$ computed from the de-factored residuals. Similarly, $\hat{\omega}^2_e = n^{-1}\sum_{i=1}^n \hat{\omega}^2_{e,i}$ is some consistent estimator of the long-run variance of $\epsilon_{i,t}$ and $\hat{\phi}^4_e = n^{-1}\sum_{i=1}^n \hat{\omega}^4_{e,i}$. The test statistics put forward by Moon and Perron [2004] are then defined as

$$t^*_a = \frac{\sqrt{n}T(\hat{\rho}^*_{pool} - 1)}{\sqrt{2\hat{\phi}^4_e/\hat{\omega}^4_e}}$$

and

$$t^*_b = \sqrt{n}T(\hat{\rho}^*_{pool} - 1)\sqrt{\frac{1}{NT^2}\text{tr}(\boldsymbol{Y}_{-1}\boldsymbol{Q}_{\hat{\boldsymbol{\Lambda}}}\boldsymbol{Y}'_{-1})} \cdot \frac{\hat{\omega}_e}{\hat{\phi}^2_e}.$$

Moon and Perron [2004] show that, as $n, T \to \infty$ with $n/T \to 0$, $t^*_a, t^*_b \to_d \mathcal{N}(0,1)$ under the null, while $t^*_a, t^*_b \to_p -\infty$ under the alternative. Hence, we reject the null at 5% for realizations smaller than -1.645.

Breitung and Das [2005]

Consider a panel model where the time series $y_{i,t}$ are generated, for $i \in \mathbb{N}_n$, by

$$\Delta y_{i,t} = \mu_i + \rho y_{i,t-1} + \epsilon_{i,t} \tag{A.5}$$

Breitung and Das [2005] work within a Seemingly Unrelated Regression framework to allow for a very general type of cross-sectional error dependence. Stacking the individual equations yields

$$\begin{pmatrix} \Delta y_{1,t} \\ \vdots \\ \Delta y_{n,t} \end{pmatrix} = \begin{pmatrix} \mu_{1,t} \\ \vdots \\ \mu_{n,t} \end{pmatrix} + \rho \begin{pmatrix} y_{1,t-1} \\ \vdots \\ y_{n,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{n,t} \end{pmatrix}$$

or $\Delta y_t = \mu + \rho y_{t-1} + \epsilon_t$, where the vectors are now $(n \times 1)$. The fairly unrestricted error covariance matrix $E(\epsilon_t \epsilon'_t) = \Omega$ admits general patterns of cross-sectional dependence, essentially only being required to satisfy weak moment restrictions and to have bounded eigenvalues. Hence, the use of the Breitung and Das [2005] test allows to investigate whether results are sensitive to the factor structure assumption on the error dependence implied by the tests outlined above. The null

hypothesis of a panel unit root is $H_0 : \rho = 0$ and is tested against the homogenous alternative $H_A : \rho < 0$. Breitung and Das [2005] propose the following statistic to test H_0:

$$t_{rob} = \frac{\sum_{t=1}^{T} y_{t-1}' \Delta y_t}{\sqrt{\sum_{t=1}^{T} y_{t-1}' \hat{\Omega} y_{t-1}}},$$

where

$$\hat{\Omega} = \frac{1}{T} \sum_{t=1}^{T} (\Delta y_t - \hat{\rho} y_{t-1})(\Delta y_t - \hat{\rho} y_{t-1})'.$$

In turn, $\hat{\rho}$ is obtained by first subtracting y_0 (the best estimate of μ under the null) from y_t and running a pooled OLS regression of $\Delta y_{i,t}$ on $(y_{i,t-1} - y_{i0})$. To account for serial correlation in the time series, Breitung and Das [2005] suggest to 'prewhiten' the demeaned series by running a preliminary regression of both $\Delta y_{i,t}$ and $y_{i,t-1}$ on p lagged differences $\Delta y_{i,t-1}, \ldots, \Delta y_{i,t-p}$ and to use the respective residuals $\widetilde{\Delta y_{i,t}}$ and $\widetilde{y_{i,t-1}}$ to construct $\hat{\rho}$ and t_{rob} as outlined above. The lag order can be chosen by some information criterion. Breitung and Das [2005] show that, as $T \to \infty$ followed by $n \to \infty$, $t_{rob} \to_d \mathcal{N}(0,1)$ and $t_{rob} \to_p -\infty$ under the (non-local) alternative $\rho < 0$. Hence, we reject the null at the 5% for realizations of $t_{rob} < -1.645$.

Demetrescu et al. [2006]

Most of the tests outlined above first remove cross-sectional dependence from the data by de-factoring and then apply some standard panel unit root tests. As an alternative, Demetrescu et al. [2006] suggest reweighing the single-unit evidence for (non-)stationarity so as to take cross-sectional dependence between the units into account when forming the panel test statistic. More specifically, they model dependence among the time series test statistics. Their approach is an extension of the inverse normal method used by Choi [2001] and, as such, tests the panel unit root null against the alternative of a non-zero fraction of stationary series in the panel.

Consider the probits $\Phi^{-1}(p_i) =: t_i$. The p-values are obtained from the t-statistic on ϕ_i in (heterogeneous) DF-regressions applied to (A.5). Demetrescu et al. [2006] use the approach of Hartung [1999], who assumes constant correlation across the the panel units, i.e.

$$Cov(t_i, t_j) = \rho, \qquad \text{for } i \neq j, \ i, j \in \mathbb{N}_n.$$

Hartung [1999] proposes to consistently estimate ρ by $\hat{\rho}^* = \max(-1/(n-1), \hat{\rho})$, where $\hat{\rho} = 1 - 1/(n-1) \sum_{i=1}^{n} (t_i - n^{-1} \sum_{i=1}^{n} t_i)^2$ to form the following panel test statistic:

$$t_{\hat{\rho}^*, \kappa} = \frac{\sum_{i=1}^{n} t_i}{\sqrt{n + (n^2 - n)\left(\hat{\rho}^* + \kappa \sqrt{\frac{2}{n+1}}(1 - \hat{\rho}^*)\right)}}$$

Here, $\kappa = 0.1 \cdot (1 + 1/(n+1) - \hat{\rho}^*)$ is a parameter designed to improve the small sample behaviour of the test statistic. Importantly, Demetrescu et al. [2006] show that the test's performance is robust to certain deviations from the constant correlation assumption. $t_{\hat{\rho}^*, \kappa} \to_d \mathcal{N}(0,1)$ under the null. The test rejects for large negative values.

References

Backus, David K., and Patrick J. Kehoe, "International Evidence of the Historical Properties of Business Cycles," *American Economic Review* 82:4 (1992), 864–88.

Bai, Jushan, and Serena Ng, "A PANIC Attack on Unit Roots and Cointegration," *Econometrica* 72:4 (2004), 1127–1177.

Banerjee, Anindya, "Panel Data Unit Roots and Cointegration: An Overview," *Oxford Bulletin of Economics and Statistics* 61:S1 (1999), 607–629.

Boucher Breuer, Janice, Robert McNown, and Myles S. Wallace, "Misleading Inferences from Panel Unit Root Tests with an Illustration from Purchasing Power Parity," *Review of International Economics* 9:3 (2001), 482–493.

Breitung, Jörg, and Samarjit Das, "Panel Unit Root Tests Under Cross Sectional Dependence," *Statistica Neerlandica* 59:4 (2005), 414–433.

Breitung, Jörg, and M. Hashem Pesaran, "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," chap. 9 (Dordrecht: Kluwer Academic Publishers, 2008), 3rd ed., pp. 279–322.

Chang, Yoosoon, "Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency," *Journal of Econometrics* 120:2 (2004), 263–293.

Choi, In, "Unit Root Tests for Panel Data," *Journal of International Money and Finance* 20:2 (2001), 249–272.

Demetrescu, Matei, Uwe Hassler, and Adina-Ioana Tarcolea, "Combining Significance of Correlated Statistics with Application to Panel Data," *Oxford Bulletin of Economics and Statistics* 68:5 (2006), 647–663.

Dickey, David A., and Wayne A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* 74:366 (1979), 427–431.

Dudoit, Sandrine, Juliet Popper Shaffer, and Jennifer C. Boldrick, "Multiple Hypothesis Testing in Microarray Experiments," *Statistical Science* 18:1 (2003), 71–103.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock, "Efficient Tests for an Autoregressive Unit Root," *Econometrica* 64:4 (1996), 813–836.

Gengenbach, Christian, Franz C. Palm, and Jean-Pierre Urbain, "Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling," *Oxford Bulletin of Economics and Statistics* 68:S1 (2006), 683–719.

Hanck, Christoph, "For Which Countries did PPP hold? A Multiple Testing Approach," *Empirical Economics* forthcoming.

Hartung, Joachim, "A Note on Combining Dependent Tests of Significance," *Biometrical Journal* 41 (1999), 849–855.

Hochberg, Yosef, and Ajit C. Tamhane, *Multiple Comparison Procedures* (Wiley, 1987).

Hommel, Gerhard, "A Stagewise Rejective Multiple Tests Procedure Based on a Modified Bonferroni Test," *Biometrika* 75:2 (1988), 383–386.

Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin, "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics* 115:1 (2003), 53–74.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu, "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics* 108:1 (2002), 1–24.

MacKinnon, James G., "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests," *Journal of Business & Economic Statistics* 12:2 (1994), 167–176.

———, "Numerical Distribution Functions for Unit Root and Cointegration Tests," *Journal of Applied Econometrics* 11:6 (1996), 601–618.

Maddala, G.S., and Shaowen Wu, "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics* 61:S1 (1999), 631–652.

Moon, Hyungsik Roger, and Benoit Perron, "Testing for a Unit Root in Panels with Dynamic Factors,"

Journal of Econometrics 122 (2004), 81–126.

Nelson, Charles R., and Charles R. Plosser, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics* 10:2 (1982), 139–162.

O'Connell, Paul G.J., "The Overvaluation of Purchasing Power Parity," *Journal of International Economics* 44:1 (1998), 1–19.

Pesaran, M. Hashem, "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence," *Journal of Applied Econometrics* 22:2 (2007), 265–312.

Phillips, Peter Charles Bonest, and Donggyu Sul, "Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence," *The Econometrics Journal* 6 (2003), 217–259.

Rogoff, Kenneth, "The Purchasing Power Parity Puzzle," *Journal of Economic Literature* 34:2 (1996), 647–668.

Romano, Joseph P., and Michael Wolf, "Formalized Data Snooping Based on Generalized Error Rates," *Econometric Theory* 24 (2008), 404–447.

Sarkar, Sanat K., "Probability Inequalities for Ordered MTP_2 Random Variables: A Proof of the Simes Conjecture," *The Annals of Statistics* 26:2 (1998), 494–504.

Simes, R. John, "An Improved Bonferroni Procedure for Multiple Tests of Significance," *Biometrika* 73:3 (1986), 751–754.

Taylor, Alan M., "A Century of Purchasing-Power Parity," *The Review of Economics and Statistics* 84:1 (2002), 139–150.

Taylor, Mark P., and Lucio Sarno, "The Behavior of Real Exchange Rates During the Post-Bretton Woods Period," *Journal of International Economics* 46:2 (1998), 281–312.