ECONOMICS BULLETIN

CIPS test for Unit Root in Panel Data: further Monte Carlo results

Andrea Cerasa
University

Abstract

This paper analyzes, through Monte Carlo experiments, the behaviour of Pesaran's CIPS test for the null of a unit root in panel data when (i) the assumption of a single common factor in the specification of the cross-section dependence is violated and (ii) the autoregressive order of the residuals is estimated. The simulation analysis points to the single common factor as a fundamental assumption for a suitable behaviour of the CIPS test and suggests that the test delivers the best performance when the truncation lag is estimated as a deterministic function of the sample size.

Citation: Cerasa, Andrea, (2008) "CIPS test for Unit Root in Panel Data: further Monte Carlo results." *Economics Bulletin,* Vol. 3, No. 16 pp. 1-13
Submitted: July 12, 2007. **Accepted:** March 19, 2008.
URL: http://economicsbulletin.vanderbilt.edu/2008/volume3/EB-07C20091A.pdf

1 Introduction

In recent years, the issue of testing for unit root in panel data has been a much debated topic. The literature about the development of such tests was initially based upon the assumption of cross-sectional independence between the units and produced the so called "first generation panel unit root tests". However, in several empirical applications, this assumption is likely to be violated and O'Connell (1998) showed that not considering the possible dependence between units could introduce severe bias in the first generation panel unit root tests. Hence researchers were interested in developing tests invariant with respect to the cross-sectional dependence, called "second generation unit root tests".
Among them, Pesaran (2006) proposed the CIPS test, based on a single common factor specification for the cross-correlation structure. Simulation results under the assumption of a single common factor and known autocorrelation order of the residuals, show that the CIPS test performs very well. Objections can be raised to the empirical relevance of both assumptions. While the latter is obvious (in practice the lag order is of course never known), the former may be acceptable in some cases but not in others. For instance, the so-called "convergence clubs" theory implies more than one common factor (see, *inter alia*, Ben-David 1994, and Galor 1996).
The aim of this paper is thus to extend Pesaran's results considering: (i) the presence of more than one common factor in the Data Generating Process (DGP) and (ii) empirical estimation of the truncation lag.
We shall now first outline the CIPS test (Section 2) and then present the results of our Monte Carlo studies (Section 3).

2 Pesaran's CIPS test

Let us consider the dynamic linear heterogeneous panel data model:

$$z_{it} = (1 - \phi_i)\mu_i + \phi_i z_{i,t-1} + u_{it} \tag{1}$$

where u_{it} has the one common factor structure

$$u_{it} = \gamma_i f_t + e_{it} \tag{2}$$

in which $f_t \sim i.i.d.(0, \sigma_f^2)$ is the unobserved common effect, $\gamma_i \sim i.i.d.(0, \sigma_\gamma^2)$ the individual factor loading and e_{it} the idiosyncratic component which can be $i.i.d.(0, \sigma_i^2)$ or, more generally, a stationary autoregressive process. Rewriting (1) and (2) as

$$\Delta z_{it} = \alpha_i + \beta_i z_{i,t-1} + e_{it}$$

Pesaran (2006) proposes to proxy the common factor f_t with the cross section mean of z_{it}, namely $\bar{z}_t = N^{-1}\sum_{i=1}^{N} z_{it}$, and its lagged value(s) $\bar{z}_{t-1}$, $\bar{z}_{t-2}$,... The test for the null of unit root regarding the unit i can now be based on the t ratio of the OLS estimate of b_i in the cross-sectionally augmented Dickey-Fuller (CADF) regression

$$\Delta z_{it} = a_i + b_i z_{i,t-1} + c_i \bar{z}_{t-1} + d_i \Delta \bar{z}_t + e_{it}$$

A natural test of the null $H_0 : \beta_i = 0$ for all i, against the heterogeneous alternative $H_1 : \beta_1 < 0, \ldots, \beta_{N_0} < 0, N_0 \leq N$ in the whole panel data set, is given by the average of the individual CADF statistics:

$$CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$$

The distribution of this test is non-standard, even asymptotically; 1%, 5% and 10% critical values are tabulated by the author for different combinations of N and T.

In case of serial correlation of the individual-specific error terms, the testing procedure can be easily extended by adding a suitable number of lagged values of $\bar{z}_t$ and Δz_{it} in the CADF regression [1] without any change in the distribution of the statistic.

3 Monte Carlo simulations

3.1 More than one common factor in the DGP

Consider first the presence of more than one common factor in the data. Adopting the same Data Generating Process (DGP) and simulation design used by Pesaran (2006), we can replace (2) with

$$u_{it} = \sum_{j=1}^{K} \gamma_{ij} f_{jt} + e_{it} \tag{3}$$

where the idiosyncratic component is given by

$$e_{it} = \rho_i e_{i,t-1} + \varepsilon_{it} \tag{4}$$

$$\varepsilon_{it} \sim N(0, \sigma_i^2) \quad \text{with} \quad \sigma_i^2 \sim Uni[0.5, 1.5]; f_{jt}, \mu_i \sim N(0,1)$$

[1] For example, for an AR(p) error specification, the suitable CADF regression is $\Delta z_{it} = a_i + b_i z_{i,t-1} + c_i \bar{z}_{t-1} + \sum_{j=1}^{p} d_{ij} \Delta \bar{z}_{t-j} \sum_{j=1}^{p} \delta_{ij} \Delta z_{i,t-j}$

and K=1, 2, 3 and 5.While the number of common factor is admittedly likely to be generally small, it is nevertheless interesting to investigate the behaviour of the test for a wide range of cases. The cross-section dependence can be low ($\gamma_{ij} \sim Uni[0, 0.2]$) or high ($\gamma_{ij} \sim Uni[-1, 3]$), and the serial correlation absent ($\rho_i = 0$), positive ($\rho_i \sim Uni[0.2, 0.4]$), high and positive ($\rho_i \sim Uni[0.4, 0.8]$[2]) and negative ($\rho_i \sim Uni[-0.2, -0.4]$). Finally, to evaluate the size of the test $\phi_i = 1$, whereas for the power $\phi_i \sim Uni[0.85; 0.95]$. These settings are combined to generate the following DGPs:

DGP 1: no serial correlation, low cross-section dependence;
DGP 2: no serial correlation, high cross-section dependence;
DGP 3: (low) positive serial correlation, high cross-section dependence;
DGP 4: (low) negative serial correlation, high cross-section dependence;
DGP 5: (high) positive serial correlation, high cross-section dependence.

Since the aim of the experiment is to analyze the behaviour of the test when the assumption of a single common factor is violated, the autoregressive order of the residuals L is always fixed to the true value for each DGP (that is, 0 for DGP 1 and DGP 2 and 1 for DGP 3, DGP 4 and DGP 5).

As expected, the results (table I) suggest that, when K=1, no or very little size bias affects the CIPS test, even with T=50. But when $K > 1$, this happens only for DGP 1 (low cross-section dependence). In the other cases, large biases appear even with K=2, which do not worsen as K grows, but unfortunately do not vanish as the sample size and the number of cross-sectional units are raised. However, the power does grow with T and N.

3.2 Empirically-selected lag length

Let us now examine the performance of the test when there is a single common factor and the lag order is empirically estimated. The DGP is now given by (1), (2) and (4). The settings do not change whereas the autoregressive order of the residuals is estimated (i) with Campbell and Perron (1991) lag length selection procedure [3]; (ii) maximizing the Bayesian Information Criteria (BIC); (iii) as a deterministic function of the sample size, following the Zivot and Andrews criterion (1992)[4]. In the first two cases, the maximum

[2] Actually, this last setting was not taken into consideration by Pesaran in his experiments. It has been added because it allows to obtain a more complete description of the behaviour of the test.

[3] The procedure starts by estimating the regression with the maximum number p of lags admitted. If the coefficient on the last lag considered is significant, the procedure stops, otherwise the equation is estimated again with (p-1) lags, etc..., until only significant lags are included.

[4] Zivot and Andrews (1992) proposed to choose the truncation lag as the largest integer less than $[(T/100)^{2/9}]$. Thus with T=50 Zivot and Andrews criteria selects $\hat{L}$=3 and when

number of lags admitted is eight, suitable if we are dealing with quarterly data.
The simulation results (table II) suggest that the introduction of the estimation of the truncation lag in the calculation of the CIPS test has a negative effect on the size when the DGP is characterized by autocorrelated residuals. With the Campbell and Perron and BIC selection methods, depending on the sign of the autocorrelation, we can have problems of over-rejection (negative autocorrelation) or under-rejection (positive autocorrelation). The Zivot and Andrews version of the test, instead, under-rejects the null even when there is not autocorrelation in the residuals.
Consider first the results on the CIPS test for DGP 1 (low cross-section dependence) and DGP 2 (high cross-section dependence). The comparison of the two rows points to what are the consequences on the performance of the test of an increase in the cross-sectional correlation degree. As we can see, such consequences are not remarkable, at least when T=100 and $N \leq 50$. In general, when T is not large, an increase of the number of cross-sectional units N has negative effects on the size of the test. In case of low dependence (DGP 1), this happens only for the Campbell and Perron version of the test, but in case of high dependence (DGP 2) the bias in the size of every version significantly increases as N increases. In particular, when T=50, the 5% size of the Campbell and Perron version of the test runs from 3.6 (N=20) to 1.6 (N=80), that of the BIC version from 5.3 (N=20) to 2.9 (N=80)and that of Zivot and Andrews version from 2.7 (N=20) to 1.4 (N=80).
These results appear to be in contrast with the expected asymptotic behaviour of the test. However, we should remember that the asymptotic distribution of the CIPS test is non-standard and the values of the size are calculated on the basis of the tabulated critical values. Then, the simulation results suggest that, when the truncation lag is estimated and T is not large enough, an increasing of N causes a departure of the distribution of the test from the simulated small sample distribution.
Differently from what happens for the size, when T=50 power improves with N.
When T=100, the estimated size of the BIC version is not affected anymore by the number of cross-sectional units. The same does not happen for Zivot and Andrews and for the Campbell and Perron version in the case of DGP 2. Furthermore, power is satisfactory value in most of cases.
Analyzing and comparing the performance of Campbell and Perron and BIC procedure, it emerges that both methods estimate the autocorrelation order (which is, for these two DGPs, zero) with high precision, at least when

T=100, $\hat{L}$=4.

T=100. However, the selection procedure based on the Bayesian Information Criteria is slightly better than the Campbell and Perron Criteria. It is seen, in fact, that the BIC version of the CIPS test has more power when T=50 than the Campbell and Perron version has and it is less affected by the size distortion deriving from the growth of N. This is not unexpected since, as Ng and Perron (1995) showed, the BIC procedure is more parsimonious than the methods based on sequential t-test as Campbell and Perron, and it works better when the residuals are not autocorrelated.

Consider next the results of table II devoted to DGP 3, DGP 4 and DGP 5. The Zivot and Andrews version again under-rejects the null and the size bias does not improve with N and T. Nevertheless, the power is positively affected by an increase of the sample size and the number of cross-sectional units.

For what regards the other two versions, a positive autocorrelation of the residuals (DGP 3 and DGP 5) introduces a clear problem of under-rejection associated to a very low level of power (at least when T is only 50). It is interesting to note, when there is high positive serial correlation (DGP 5), how the size of the test appears to be slightly better compared with the case of low positive serial correlation (DGP 3). This can be due to the fact that both Campbell and Perron and BIC procedures perform better when the level of the autocorrelation is higher. However, this better behaviour in term of size is joined to a lower level of power. If the residuals are negatively correlated (DGP 4), the estimated size of the CIPS test is much higher than the nominal level. In general, an increasing of N leads again to a greater bias in term of size and to an improvement in term of power, which is somewhat expected from what we observed for DGP 1 and DGP 2. But it is not expected the fact that the same tendency is present also in the "true L" version of the test when the autocorrelation of the residuals is high and positive (DGP 5).

Differently from what happened for the first two DGPs, an increasing of T has not always a positive effect on the size. In particular, the estimated size of the BIC version of the test is further from the nominal level when T raises to 100. One possible explanation for this negative effect centers on the fact that the BIC method chooses the truncation lag L which maximizes: $\ell_L(\hat{\theta}_L) - \frac{1}{2}\kappa_L log(T)$ where the first expression denotes the maximized value of the log-likelihood function for the model with autoregressive order L whereas the second term represents the penalty for having an additional parameter (κ_L is, indeed, the number of the parameters that needs to be estimated in the model). Thus, as T increases, the penalty for having an additional parameter becomes larger and the method becomes more parsimonious. Hence, even though asymptotically valid, we can think that BIC method, for T large but not enough to guarantee a good asymptotical approximation, has the

tendency to under-parameterize the model. This has adverse effects in the size of the CIPS test.
Contrary to what happens for the size, the power is always positively affected by the increase of the time dimension.
Finally, now the Campbell and Perron version seems to suffer less from the size distortions than the BIC version, both when the autocorrelation of the residuals is positive and when is negative. However, this advantage is compensated with a loss in term of power.
The strange behaviour of Campbell and Perron and BIC versions of the CIPS test is not unexpected. Ng and Perron (2001) indeed show by simulation that, in case of positive autocorrelation of the residuals, an under-estimation of the true autocorrelation order leads to under-rejection of the true null for three individual unit root tests. Whereas, in case of negative autocorrelation, the under-estimation of the true autocorrelation order causes a strong over-rejection problem.
Hence, an explanation for the strange behaviour of the test can be provided by an experiment similar to the one performed by Ng and Perron (2001) and then estimating the size of the CIPS test with L=0,1,...,8. The results (table III) confirm that when the true autoregressive order is imposed (that is, 0 for DGP 1 and DGP 2 and 1 for DGP 3, DGP 4 and DGP 5). The estimated size of the test is close to the nominal level and in line with the estimates obtained by Pesaran. When the autoregressive order is under-estimated, the same problems highlighted by Ng and Perron (2001) arise: under-rejection in case of positive autocorrelation and strong over-rejection in case of negative autocorrelation.
Therefore, the size problems of Campbell and Perron and BIC versions of the CIPS test seem to depend on the performance of the two lag length selection criteria. The probability for both methods to under-estimate the autocorrelation order can be investigated through a new Monte Carlo experiment. As expected, the results (table IV), confirm that the BIC procedure is more parsimonious than that of Campbell and Perron, in accordance with Ng and Perron (1995). When the true lag L is 0, this means that the BIC method outperforms that of Campbell and Perron. Moreover, also the previous hint regarding the penalty term of the BIC procedure is confirmed: on average, when T is raised, the procedure tends to select a lower order of autocorrelation and then to under-parameterize the model when there is autocorrelation in the residuals.
The mean truncation lag selected by Campbell and Perron method is around 3 and it is not remarkably influenced by the growth of time dimension. However, even though on average Campbell and Perron procedure seems to over-parameterize the model, it selects $\hat{L}$=0 when the true autoregressive order is

1 in a consistent number of cases.
Finally, comparing the rows of table IV devoted to DGP 3 and DGP 5, it is confirmed that the higher is the level of autocorrelation in the residuals, the higher the probability is to select the correct truncation lag (that is 1) and the lower the probability is to under-parameterize the model (that is to select $\hat{L}$=0). Despite these improvements both methods continue to select $\hat{L}$=0 in a consistent number of cases also when there is high serial correlation in the residuals.
Concluding, both methods have the tendency to select $\hat{L}$=0 too often when the residuals are autocorrelated and this causes large distortions in the CIPS test size.

4 Conclusion

Monte Carlo evidence suggests that the behavior of the CIPS test is not satisfactory when the assumption of a single common factor in the specification of the cross-section dependence is violated. Furthermore, the estimation of the truncation lag L is an important issue: the selection of a wrong L dramatically affects the good properties of the test. Experimental results have shown how the version of the test based on selecting the autoregressive order according to a deterministic function of T (as that of Zivot and Andrews) seems to be the most suitable if compared with selection criteria based on Information criteria (as BIC) or on sequential t-test (as Campbell and Perron). These two methods tend to select $\hat{L}$=0 too often causing under-rejection in case positive autocorrelation and strong over-rejection in case of negative autocorrelation.

Table 1: Size and power of a 5% CIPS test when the autoregressive order is known and the number of common factors K is greater than one.

			K=1		K=2		K=3		K=5	
N	**T**	**DGP**	size	power	size	power	size	power	size	power
20	**50**	**1**	5.3	61.6	4.9	59.7	5.6	60.0	5.0	60.7
		2	5.7	59.3	15.0	58.5	17.2	55.2	14.7	56.2
		3	4.7	40.0	9.4	39.4	10.0	41.9	10.3	41.9
		4	4.8	50.7	16.0	48.0	16.6	46.1	15.5	46.6
		5	3.8	25.0	2.1	16.1	2.3	13.9	2.2	16.9
	100	**1**	5.3	99.9	4.9	99.9	5.4	100.0	5.0	99.9
		2	5.5	99.9	14.8	96.7	16.0	95.7	15.5	96.6
		3	5.5	98.9	9.6	93.4	9.8	92.2	10.6	93.5
		4	5.6	99.7	17.0	90.9	18.2	90.9	15.4	92.8
		5	4.2	91.6	1.6	72.0	1.3	67.3	1.7	71.4
50	**50**	**1**	5.1	85.3	5.0	84.9	5.4	85.2	5.3	85.6
		2	4.9	85.4	19.1	69.6	21.2	67.4	19.5	68.4
		3	4.4	56.6	10.4	53.4	12.8	52.9	13.7	54.3
		4	4.5	73.9	20.6	59.3	21.2	55.2	18.5	54.6
		5	3.4	34.7	1.9	19.8	1.9	18.7	2.6	22.7
	100	**1**	4.6	100.0	4.5	100.0	4.6	100.0	4.9	100.0
		2	5.0	100.0	19.1	98.0	20.6	98.4	18.5	98.8
		3	4.6	100.0	10.9	97.8	12.1	96.7	13.2	97.7
		4	4.2	100.0	19.4	95.0	21.1	94.8	19.5	96.9
		5	3.6	99.9	1.4	90.6	1.1	84.3	1.6	87.4
80	**50**	**1**	4.7	92.2	4.7	92.1	4.9	92.3	5.1	92.3
		2	4.9	92.8	21.2	71.7	22.3	70.7	20.7	70.2
		3	4.2	64.0	11.6	57.4	13.5	56.0	15.1	56.6
		4	4.5	82.9	22.3	61.6	22.3	57.8	19.6	58.9
		5	2.9	38.5	2.1	20.8	1.4	20.2	2.6	25.9
	100	**1**	5.4	100.0	4.8	100.0	4.7	100.0	5.2	100.0
		2	5.0	100.0	20.7	98.7	23.8	98.6	20.7	99.5
		3	4.8	100.0	11.8	98.2	13.6	97.8	14.9	98.7
		4	5.4	100.0	22.8	95.9	23.5	95.8	20.8	97.4
		5	3.7	100.0	1.4	94.1	1.1	88.9	1.6	92.1

Notes: the DGPs are generated according to the expressions (1), (3) and (4) with ϕ_i=1 $\forall i$ in the size case and $\phi_i \sim Uni[0.85, 0.95]$ for the power case. DGP 1: ρ_i=0; $\gamma_{ij} \sim Uni[0, 0.2]$; DGP 2: ρ_i=0; $\gamma_{ij} \sim Uni[-1, 3]$; DGP 3: $\rho_i \sim Uni[0.2, 0.4]$; $\gamma_{ij} \sim Uni[-1, 3]$; DGP 4: $\rho_i \sim Uni[-0.2, -0.4]$; $\gamma_{ij} \sim Uni[-1, 3]$; DGP 5: $\rho_i \sim Uni[0.4, 0.8]$; $\gamma_{ij} \sim Uni[-1, 3]$.

Table 2: Size and power of a 5% CIPS test when the autoregressive order is estimated.

			C. & P.		*BIC*		*Z. & A.*		*True L*	
N	***T***	**DGP**	size	power	size	power	size	power	size	power
20	**50**	**1**	4.6	41.4	6.4	61.7	2.3	18.7	5.1	61.0
		2	3.6	33.6	5.3	54.5	2.7	22.6	5.4	61.8
		3	1.2	11.5	1.2	13.9	3.8	18.1	5.2	39.3
		4	16.5	68.8	33.7	91.9	3.0	22.0	6.0	51.1
		5	1.6	9.0	1.6	9.9	3.2	14.5	3.9	25.6
	100	**1**	4.6	97.7	6.1	99.9	3.3	85.3	5.1	99.9
		2	4.9	97.5	5.4	99.8	4.2	88.7	5.7	99.9
		3	1.5	82.1	1.2	79.6	3.4	80.6	4.9	98.8
		4	16.3	99.0	37.0	100.0	3.7	90.6	5.1	99.7
		5	2.6	64.6	2.2	55.3	3.4	63.4	4.9	91.5
50	**50**	**1**	3.2	60.1	6.1	85.8	3.3	26.8	4.7	84.6
		2	2.0	46.1	4.5	77.1	2.6	28.3	5.3	85.5
		3	0.3	13.1	0.4	15.6	2.9	28.3	4.3	58.2
		4	19.5	85.1	40.2	97.9	3.2	30.0	5.1	73.7
		5	0.4	7.6	0.5	8.5	1.2	14.3	3.3	36.4
	100	**1**	4.1	100.0	6.3	100.0	2.9	98.9	4.8	100.0
		2	3.8	99.9	5.2	100.0	2.6	98.7	5.3	100.0
		3	0.6	97.2	0.4	97.4	2.5	97.6	4.0	100.0
		4	17.8	100.0	49.6	100.0	1.7	99.6	4.5	100.0
		5	1.4	84.6	0.8	76.0	1.7	89.1	4.1	99.9
80	**50**	**1**	2.6	65.9	4.7	90.5	2.8	27.5	4.2	92.2
		2	1.6	52.3	2.9	82.9	1.4	30.7	4.4	92.2
		3	0.1	11.9	0.1	14.8	1.2	23.9	4.3	64.8
		4	19.8	89.3	43.0	98.8	2.5	33.9	4.3	83.1
		5	0.2	6.3	0.2	6.5	1.1	15.5	2.8	37.8
	100	**1**	4.8	100.0	5.1	100.0	1.7	99.7	5.4	100.0
		2	3.1	100.0	4.9	100.0	2.4	99.8	5.1	100.0
		3	0.6	99.6	0.1	99.4	2.0	99.3	4.8	100.0
		4	20.3	100.0	52.5	100.0	1.1	100.0	5.3	100.0
		5	0.7	92.8	0.5	80.5	1.2	93.6	3.7	100.0

Notes: the DGPs are generated according to the expressions (1), (2) and (4) with ϕ_i=1 $\forall i$ in the size case and $\phi_i \sim Uni[0.85, 0.95]$ for the power case. DGP 1: ρ_i=0; $\gamma_i \sim Uni[0, 0.2]$; DGP 2: ρ_i=0; $\gamma_i \sim Uni[-1, 3]$; DGP 3: $\rho_i \sim Uni[0.2, 0.4]$; $\gamma_j \sim Uni[-1, 3]$; DGP 4: $\rho_i \sim Uni[-0.2, -0.4]$; $\gamma_i \sim Uni[-1, 3]$; DGP 5: $\rho_i \sim Uni[0.4, 0.8]$; $\gamma_i \sim Uni[-1, 3]$.

Table 3: Estimated size of a 5% CIPS test when the truncation lag L is fixed

N	***T***	**DGP**	***L*=0**	***L*=1**	***L*=2**	***L*=3**	***L*=4**	***L*=5**	***L*=6**	***L*=7**	***L*=8**
20	**50**	**1**	**4.3**	4.6	2.7	2.3	1.9	1.0	1.1	0.5	0.1
		2	**6.3**	5.0	3.4	2.7	1.7	1.1	1.4	1.4	0.8
		3	0.0	**5.6**	4.2	3.8	2.4	2.0	1.1	1.5	0.8
		4	73.9	**5.1**	3.5	3.0	2.0	1.7	1.4	0.6	0.4
		5	0.0	**4.8**	3.2	3.2	2.5	2.0	1.3	0.5	0.3
	100	**1**	**5.3**	5.6	4.8	4.0	3.3	3.2	2.2	2.1	1.7
		2	**6.2**	5.8	5.2	5.4	4.2	3.4	3.4	2.7	2.5
		3	0.0	**4.9**	4.1	4.6	3.4	3.7	2.6	2.5	2.1
		4	74.3	**5.0**	3.6	3.9	3.7	3.1	2.6	2.8	2.2
		5	0.0	**4.0**	3.8	4.1	3.4	3.6	2.9	2.7	1.9
50	**50**	**1**	**4.5**	3.6	3.4	3.3	2.0	1.2	0.4	0.1	0.2
		2	**5.4**	4.5	2.7	2.6	1.4	0.7	0.1	0.3	0.0
		3	0.1	**4.9**	2.6	2.9	1.0	1.1	0.7	0.3	0.1
		4	85.1	**5.1**	3.2	3.2	1.3	1.0	0.6	0.7	0.3
		5	0.0	**2.4**	1.5	1.2	0.7	0.5	0.3	0.3	0.2
	100	**1**	**5.3**	5.3	3.8	3.9	2.9	2.7	2.6	1.8	1.3
		2	**4.7**	4.6	3.2	3.0	2.6	2.5	1.6	1.9	1.4
		3	0.0	**3.5**	3.3	2.7	2.5	2.3	1.1	1.5	1.1
		4	89.1	**4.3**	2.7	2.2	1.7	1.5	1.0	1.2	0.5
		5	0.0	**3.4**	2.2	2.3	1.7	1.4	1.0	1.5	1.0
80	**50**	**1**	**5.2**	4.9	2.3	2.8	0.6	0.6	0.0	0.2	0.0
		2	**4.3**	3.9	1.7	1.4	0.6	0.4	0.1	0.0	0.1
		3	0.0	**3.9**	1.4	1.2	0.5	0.5	0.2	0.2	0.1
		4	90.7	**5.6**	2.2	2.5	0.8	0.4	0.4	0.4	0.0
		5	0.0	**2.2**	1.0	1.1	0.3	0.5	0.2	0.1	0.1
	100	**1**	**4.6**	3.8	3.4	3.7	1.7	2.0	1.3	1.0	0.5
		2	**6.1**	5.1	3.5	4.0	2.4	1.9	1.8	1.8	1.7
		3	0.0	**3.7**	3.1	2.5	2.0	1.6	1.0	0.9	0.6
		4	94.8	**4.1**	2.1	2.1	1.1	1.3	0.9	1.7	1.0
		5	0.0	**3.5**	2.4	2.1	1.2	1.4	0.9	0.9	0.6

Notes: For the characteristics of the five DGPs see note under table II. In bold type, the estimated sizes of the CIPS test corresponding to the true autoregressive order L.

Table 4: Performance of Campbell and Perron and BIC lag length selection procedures (*N=1*)

			Probability of selecting									
	T	**DGP**	$\hat{L}$**=0**	$\hat{L}$**=1**	$\hat{L}$**=2**	$\hat{L}$**=3**	$\hat{L}$**=4**	$\hat{L}$**=5**	$\hat{L}$**=6**	$\hat{L}$**=7**	$\hat{L}$**=8**	***Mean*** $\hat{L}$
Campbell & Perron	**50**	**1**	**48.9**	5.4	6.0	4.4	5.4	7.2	5.1	8.7	8.9	*2.5*
		2	**43.0**	5.4	5.8	5.9	5.9	8.0	6.2	10.1	9.7	*2.8*
		3	32.8	**17.9**	7.7	5.2	5.5	6.4	8.6	7.6	8.3	*2.7*
		4	38.7	**13.1**	4.4	5.7	5.5	7.5	6.9	9.8	8.4	*2.8*
		5	22.5	**26.5**	7.4	6.5	6.5	9.0	6.9	7.8	6.9	*2.8*
	100	**1**	**43.9**	4.8	4.1	8.0	7.3	7.6	9.8	8.4	6.1	*2.7*
		2	**41.8**	6.0	5.1	6.1	6.3	6.8	8.1	10.9	8.9	*2.9*
		3	24.7	**21.2**	5.4	6.1	7.2	7.7	7.8	9.2	10.7	*3.1*
		4	30.5	**19.8**	5.6	6.2	6.9	6.5	7.9	8.2	8.4	*2.8*
		5	10.8	**33.3**	9.0	6.7	7.5	7.0	7.6	7.5	10.6	*3.2*
BIC	**50**	**1**	**64.7**	11.6	6.2	3.1	2.8	2.7	2.6	2.7	3.6	*1.2*
		2	**68.1**	11.2	5.0	3.3	2.9	2.4	1.2	2.5	3.4	*1.1*
		3	52.4	**23.3**	7.8	4.3	2.9	2.6	2.6	2.0	2.1	*1.2*
		4	55.2	**20.8**	6.4	4.1	3.2	2.8	1.8	2.1	3.6	*1.3*
		5	29.6	**36.9**	10.7	6.2	4.3	3.8	2.4	2.8	3.3	*1.7*
	100	**1**	**80.9**	9.9	4.9	1.8	0.7	0.7	1.0	0.0	0.1	*0.4*
		2	**83.8**	8.3	4.0	1.6	1.1	0.3	0.2	0.5	0.2	*0.3*
		3	51.5	**36.9**	6.7	1.5	1.4	1.2	0.1	0.5	0.2	*0.7*
		4	58.2	**30.4**	5.3	3.0	0.9	0.9	0.7	0.2	0.4	*0.7*
		5	23.6	**54.3**	12.7	5.4	1.3	1.7	0.4	0.3	0.3	*1.2*

Notes: For the characteristics of the five DGPs see note under table II. Moreover, in all cases, ϕ_i=1. The results with $\phi_i \sim Uni[0.85, 0.95]$ are very similar and then not reported. In bold type, the probability to select the correct truncation lag.

References

Ben-David, D. (1994) "Convergence Clubs and diverging economies", CEPR Discussion Paper Series 922, London.

Campbell, J. and P. Perron (1991) "Pitfalls and Opportunities: what Macroeconomists should know about Unit Roots", in Blanchard, O. and S. Fishers (Eds.), NBER Macroeconomics Annual, MIT Press, Cambridge, MA.

Galor, O. (1996) "Convergence? Inferences from Theoretical models", *The Economic Journal*, **106**, 1056-1069.

Ng, S. and P. Perron (1995) "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag", *Journal of the American Statistical Association*, **90**, 268-281.

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.

O'Connell, P. (1998) "Overvaluation of Purchasing Power Parity", *Journal of International Economics*, **44**, 1-19.

Pesaran, M.H. (2006) "A simple Panel Unit Root test in the presence of Cross-section Dependence", Cambridge Working Papers in Economics, 0346.

Zivot, E. and D.W.K. Andrews (1992) "Further evidence on the great cash, the oil-price shock and the unit root hypothesis", *Journal of Business and Economic Statistics*, **10**, 251-270.