

Cointegration In Panels With Correlated Units¹

Cointegrazione in panel con unità correlate

Stefano Fachin

Facoltà di Scienze Statistiche
Università di Roma "La Sapienza"
s.fachin@caspur.it

Abstract: L'obiettivo di questo lavoro è proporre una procedura bootstrap per test di cointegrazione su dati panel con unità dipendenti sia nel breve che nel lungo periodo e valutarne le proprietà mediante simulazioni Monte Carlo. I risultati ottenuti mostrano che la procedura, basata sul "Continuous-Path Block Bootstrap", ha errore di prima specie sempre vicino al livello di significatività del test e buona potenza. Costituisce quindi una interessante alternativa alle procedure parametriche, soprattutto nei numerosi casi in cui le ipotesi alla base di esse non sono rispettate.

Keywords: Panel cointegration, Bootstrap, Continuous-Path Block Bootstrap, Group Mean Panel Cointegration tests.

1. Introduction

The analysis of non-stationary panels is currently one of the most actively researched areas in econometrics, so much so that any list of references is both likely to be incomplete and bound to be rapidly obsolete. This wide interest has two reasons, the first being the empirical importance of these methods. Indeed, at the outset their development has been strictly linked with the attempt to rescue the PPP hypothesis, which received little empirical support from non-stationarity analysis on time series. The suspect that this was mainly due to power problems with the sample sizes commonly available directed the attention towards panel data, an obvious way to increase the amount of information available and hence power (*inter alia*, Frankel and Rose, 1996). However, the PPP is just one of the many instances of natural use of cross-sections of time series; other obvious examples are convergence and studies of money demand in the Euro area. The second reason is the many challenging methodological problems which remain to be solved. The most important of these is probably abandoning the highly convenient, but unfortunately totally unbelievable, assumption of independence across units typical of the first generation tests. Ignoring short-run cross-correlation and cross-units long-run dependence can cause severe size distortion (evidence is given respectively by O'Connell, 1998; Maddala and Wu, 1999; Banerjee *et al.*, 2001), so that the power gain delivered by the panel dimension, which is the very reason for its use, is entirely fictitious. Although several panel unit root testing procedures allowing for cross-section dependence of fairly general form have been recently proposed (Bai and Ng, 2002; Chang, 2002; Harvey and Bates, 2003; Moon and Perron, 2003; Phillips and Sul, 2002; Pesaran, 2003), progress under this respect seems to be much slower in the case of cointegration tests. The methods based on full system estimation (Groen and Kleinbergen, 2003; Larsson *et al.*, 2001) are in principle the most natural solution, but, as the number of parameters to be estimated grows with the square

¹Research supported by MIUR and CNR. This paper is based on work carried out while the author was visiting the Cambridge University Statistical Laboratory. I would like to thank G.A. Young for his hospitality, and Peter Pedroni, Hashem Pesaran and Dimitris Politis for suggestions and comments. The usual disclaimers apply.

of the cross-section dimension, their empirical relevance tends to be limited to panels with a large number of time observations and a small number of units. Within the class of single-equation methods very much the same applies to the GLS approach (O’Connell, 1998; Pedroni, 1997). As the need to resort to panel techniques is highest when the time dimension is smallest, these are unfortunately the least interesting cases. Given this background, the aim of this paper is to evaluate the scope for bootstrap tests of cointegration, a natural solution already advocated by Maddala and Wu (1999) but still largely unexplored, and, to the best of our knowledge, only for panel unit root tests (Maddala and Wu, 1999; Chang, 2004; Smith *et al.*, 2004). We will confine our attention to single-equation procedures, and, within this class, to the Group t -statistic (Pedroni, 1997), which could be seen as the cointegration version of the popular panel unit root test by Im *et al.* (2003). While the extension to other single-equation tests is immediate, so that considering just one specific test does not cause much loss of generality, the exclusion of full information methods is justified by their smaller empirical appeal already mentioned above. We will rely on simulations, rather than analytical asymptotic results, as the latter are not only very difficult to obtain under general forms of cross-section dependence, but also of very little help in assessing the performance in panels with small time and cross-section dimensions. The paper is organised as follows: we shall first recall the modelling set-up and outline the proposed testing procedures (section 2), then describe the design and results of the Monte Carlo experiment (section 3). Some conclusions are finally drawn (section 4)

2. Testing Panel Cointegration

The object of the study are the long-run properties of two non-stationary random variables X and Y , with information available for N different units (indexed by i) over a sample of T observations (indexed by t); we thus have what we may call a cross-section of time series, or a panel; if N/T is not large the first definition may actually be more fitting than the latter. Especially in the case of small T , the idea of strengthening the results of tests aimed at assessing the existence of a long-run relationship between X and Y by exploiting the cross-section dimension is quite natural. If the power of the available testing procedure is not very high the risk of failing to reject a wrong null hypothesis in a single small sample is substantial; however, by pooling together or averaging the information from several, distinct small samples we should indeed be able to shed some more light on the issue and limit the chance of wrong inferences. As mentioned in the Introduction we shall confine our attention to single-equation methods. Excluding for notational convenience the case of deterministic time trends, the hypothesis of a long-run relation between X_i and Y_i may be expressed as

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}^y, i = 1, \dots, N; t = 1, \dots, T, \quad (1)$$

and tested by one of the many available procedures aimed at evaluating the properties of the residuals of (1). Let $x_{it} = x_{it-1} + u_{it}^x$ and $u_{it} = [u_{it}^x u_{it}^y]'$; then, under the assumption $E(u_{it} u_{js}') = 0 \forall i \neq j$ and $\forall t, s$, Pedroni (1999) showed that both standardised sums of individual statistics and statistics computed on the pooled residuals of the N equations, (respectively *group mean panel cointegration statistics* and *panel cointegration statistics* in Pedroni’s terminology) are asymptotically normal. The null hypothesis is in both cases that of a unit root in the residuals of (1); the alternative hypothesis is that the roots are in all units smaller than one, for the panel cointegration statistics with the further restriction that the value is the same across units. The means and variances, which depend on

the number of regressors and deterministic terms, have been computed by simulation and tabulated by Pedroni, so that inference is standard Gaussian. A first noticeable limitation of the method is that the number of regressors (i.e., the cointegration rank) is assumed to be constant over units. However, the critical point is the assumption of independence. Although including time dummies common to all units (a standard practice in the modelling of panel data, equivalent to de-meaning under homogeneity of the regressors) may eliminate a considerable amount of short-run cross-correlation, it is also important to stress that long-run dependence is also ruled out, with the implication that no regressors may be common to different units. Both cases (cross-units cointegration in the regressors and common regressors) are empirically relevant, e.g. when the units are regions within an economy. When the fundamental assumption of independence across units is satisfied the available simulation evidence suggests that the Group t -statistic may deliver good results, provided some care is exercised. Power is generally high (Gutierrez, 2003), but Banerjee *et al.* (2001) report substantial overrejection if the time dimension is small in absolute terms ($T < 50$), or not high enough relatively to the cross-section dimension (for instance, with $N = 64$ $T = 200$ is not enough). Since we must expect gaussian inference to collapse when there is strong cross-correlation or common regressors it is natural to explore the results delivered by the bootstrap. In principle the bootstrap should present a number of further advantages, allowing use of robust between groups statistic, e.g. the median of the cointegration statistics of the individual units, and heterogeneity in the cointegration rank over units. Obviously, the key point is how to construct the pseudodata: our proposal is to apply separately to the entire cross-sections of time-series, $\{X_1 X_2 \dots X_N\}_{t=1}^T$ and $\{Y_1 Y_2 \dots Y_N\}_{t=1}^T$, the Continuous-Path Block Bootstrap (CBB). Developed by Paparoditis and Politis (2001), this is a block resampling method designed to construct non-stationary pseudodata. The pseudoseries is obtained in essentially two steps: first, a block bootstrap series is constructed integrating within each block the resampled first differences of a series known² to be non-stationary; second, the end points of the blocks are chained so to eliminate jumps between blocks. As the resampling is applied to the entire cross-section the pseudoseries will clearly preserve the cross-correlation structure and the non-stationary individual patterns. On the other hand, the blocks are chosen independently for the X 's and the Y 's, so that the pseudoseries are independent by design. This point is obviously critical, as a bootstrap test requires resampling under the null hypothesis, which here is no cointegration. Denoting by G a between group statistic (which may be Pedroni's Group t -statistic, the median of the individual cointegration ADF's, or any other between groups cointegration statistic³), the bootstrap procedure includes five simple steps:

1. compute the Group statistic $\hat{G}$ for the data set $\{X_1 X_2 \dots X_N, Y_1 Y_2 \dots Y_N\}_{t=1}^T$;
2. construct separately by CBB two sets of N pseudoseries, $\{X_1^* X_2^* \dots X_N^*\}_{t=1}^{T^*}$ and $\{Y_1^* Y_2^* \dots Y_N^*\}_{t=1}^{T^*}$;
3. compute the Group statistics G^* for the pseudo-data set, $\{X_1^* X_2^* \dots X_N^*, Y_1^* Y_2^* \dots Y_N^*\}_{t=1}^{T^*}$;
4. repeat steps (2) and (3) a large number (say, B) of times;

²For a variant designed for non-stationarity testing see Paparoditis and Politis (2003).

³Recall however that pivotal statistics should deliver superior results.

5. compute the bootstrap significance level; assuming that the rejection region is the left tail of the distribution, $p^* = \text{prop}(G^* < \hat{G})$.

A critical point of the algorithm is clearly the choice of the block length. We will not enter into the issue, which is still essentially open, and fix the block length at 10% of the sample size, a rather common choice. This entails that the performances of the bootstrap test are likely to be inferior to those that may be obtained from fine-tuning the block length.

3. Monte Carlo Experiment

3.1 Design

The Monte Carlo experiment is based on a panel extension of the classical Engle-Granger bivariate data generating process (DGP), adopted by e.g. Kao (1999) and Gutierrez (2003). The correlation across units is obtained as a consequence of a common factor in the disturbances, as in Pesaran (2003). We then have a cross-section of N pairs of T observations over time of two, possibly cointegrated, non-stationary random variables, X and Y , driven by shocks $(u^j, j = x, y)$ which are the sum of an idiosyncratic component $(\epsilon^j, j = x, y)$ and a stationary common factor $(f_t^j, j = x, y)$; the former generates the long-run path, while the latter produces short-run correlation across units. More precisely:

$$\begin{cases} x_{it} = \frac{a_1 u_{it}^y + u_{it}^x}{1 - a_1} \\ y_{it} = x_{it} + u_{it}^y \end{cases} \quad (2)$$

where possibly $x_i = x_j$ for some i, j , and

$$\begin{cases} u_{it}^x = \gamma_i^x f_t^x + \epsilon_{it}^x \\ u_{it}^y = \gamma_i^y f_t^y + \epsilon_{it}^y \end{cases} \quad (3)$$

The coefficients $\gamma_i^x, j = x, y$, are the factor loadings and determine the strength of the short-run cross-correlation across units. The structure of the idiosyncratic component is:

$$\begin{cases} \epsilon_{it}^y = \phi_i \epsilon_{it-1}^y + e_{it}^y \\ \epsilon_{it}^x = \sum_j \epsilon_{it-1}^j \end{cases} \quad (4)$$

so that when $\phi_i = 1$ there is no cointegration between X_i and Y_i , while the closer ϕ_i is to zero the higher is the speed of adjustment to the long-run equilibrium in unit i . Given that we are mostly interested in a comparison between the performance of the tests based upon the asymptotic distributions and the bootstrap procedures, without much loss of generality in the power simulations we will consider only the case $\phi_i \sim \text{Uniform}(0.2, 0.4)$. Finally,

$$\begin{cases} e_{it}^y \sim N(0, \sigma_i) \\ e_{it}^x \sim N(0, \sigma_i) \end{cases} \quad (5)$$

with σ_i drawn from a Uniform density defined over an interval centred on 1, so to allow for some heterogeneity across units; here without loss of generality we chose $[0.5, 1.5]$. Clearly, with such a set-up we cannot aim at a full factorial design: the amount of information produced will be exceedingly large. Rather, we will define a benchmark case representative of a set-up as close as possible to reality, and then explore variations in a few relevant directions:

- *Benchmark*: $N = 10; T = 30; \gamma_i^x, \gamma_i^y \sim \text{Uniform}(-1, 6)$. Both the time and the cross-section dimensions are small, with the latter dominating the former consistently with the assumptions underlying the asymptotics of the test. The number of time observations is similar to that often available for international macroeconomic data sets with annual data; the number of cross-section units could appear small, but it is plausible if we take for instance the largest world or European economies, or the NUTS 2 areas within one European economy. The cross-correlation is substantial (on the average about 0.65).
- *Independent units*: as Benchmark, except $\gamma_i^x = \gamma_i^y = 0, \forall i$; in order to assess the performance of the bootstrap tests we clearly need to consider a case satisfying the assumptions underlying the asymptotic test.
- *Common regressor*: as Benchmark, except $X_i = X_1, i = 2, \dots, N$; in this case the asymptotic test should collapse.
- *Larger cross-section*: as Benchmark, except $N = 30$; given that the sample size is identical for the time series and cross-section dimension, this case can be expected to be troublesome.
- *Larger T* : as Benchmark, except $T = 100$; this can be considered a rather large sample size in the time dimension, and all methods should deliver good results.

In all these cases we assumed cointegration to hold or not in all units at the same time; hence, the data reflect the formulation of the null and alternative hypotheses. However, it may well be the case that cointegration holds only in some of the units. We then defined a last case:

- *Variable cointegrating rank*: as Benchmark, with cointegration in 80% of the units; $N = 5, 10$.

In all simulations both the number of Monte Carlo replications and bootstrap redrawings has been fixed at 500. A few pilot studies made clear that a higher number of either would not have added much in terms of precision, while increasing significantly the cost and time scale of the experiment. In all cases we fitted a panel regression including common time dummies, computed the individual cointegration ADF statistics selecting the lag length by t -tests (Ng and Perron, 1995), and finally computed the asymptotic and bootstrap p -values for the Group t -statistic (respectively, Group t_A and Group t_B in tables 1-6) as well as the bootstrap p -values for the median of the ADF's (t_{Median} in the tables). Note that given the number of regressors and specification of the deterministic kernel the mean and variance of the Group t -statistic can be considered known constants, and are thus irrelevant for the bootstrap calculations (this implies that using the bootstrap the number of regressors, or cointegration rank, can vary across units). In both cases the bootstrap p -values have been also computed applying the Fast Double Bootstrap (Davidson and MacKinnon, 2000), a procedure designed to correct possible bias in the bootstrap estimates of the p -values.

32 Results

The results are reported in a separate table for each case, starting with the Benchmark case in Table 1. As expected, use of the standard normal p -values causes heavy overrejection, while the bootstrap procedures always have Type I errors close to nominal and

power acceptable when tests are performed with nominal sizes greater than 5% (and note that a 10% bootstrap test has a Type I error actually smaller than an asymptotic 1% test). On the other hand, if the units are actually independent the bootstrap test performs as well as the asymptotic test (Table 2). Thus, the bootstrap tests seem to be clearly superior to the asymptotic procedure, with no significant differences between the statistics based upon the mean (Group t_B) and the median (t_{Median}) of the individual ADF's. If the regressor is the same in all units (an extreme case of long-run dependence), the asymptotic test produces as expected disastrous results, while those of the bootstrap tests are in some cases actually better than in the benchmark case. This is especially true if we consider the FDB p -values (Table 3). The large T (Table 4) and large N (Table 5) results are also very interesting. While a hundred time observations are not enough for the asymptotic procedure to achieve Type I errors close to nominal size in presence of short-run dependence across ten units, all the bootstrap tests are very close to this objective while at the same delivering a power performance essentially equal to that of the asymptotic tests. When the number of units and time observations is the same, namely 30, we find, consistently with Banerjee *et. al*, that the asymptotic tests overrejects heavily, while the bootstrap test has again Type I error close to the nominal size and very high power. Finally, it is interesting to consider the experiments with variable cointegration rank (Table 6). Recall that the null hypothesis is cointegration in *no* unit, while the alternative hypothesis is cointegration in *all* units. When the cross-section dimension is very small ($N = 5$) the chance of falsely claiming cointegration to hold in all units when this is actually true only in 4 out of 5, is only marginally higher than the Type I Error in the Benchmark case. However, when we have ten units and the same proportion of cointegrated units the probability of this wrong inference is much higher even if we use the bootstrap. The question is: is it really wrong? Indeed, the view that cointegration holding in 80% of the units may be enough to conclude that the two variables under study are related may not be wholly unreasonable.

Table 1
Rejection Rates - Benchmark Case
 $T = 30, N = 10$, correlated units

	Group t_A	Group t_B		t_{Median}	
		<i>simple</i>	<i>FDB</i>	<i>simple</i>	<i>FDB</i>
α		A. Size: $\phi_i = 1 \forall i$			
0.01	0.15	0.01	0.01	0.01	0.01
0.05	0.24	0.05	0.04	0.05	0.04
0.10	0.31	0.11	0.09	0.11	0.08
0.20	0.39	0.23	0.17	0.19	0.17
α		B. Power: $\phi_i \sim Uniform(0.2, 0.4)$			
0.01	0.85	0.30	0.52	0.37	0.28
0.05	0.90	0.62	0.51	0.67	0.57
0.10	0.92	0.75	0.66	0.75	0.68
0.20	0.96	0.89	0.89	0.85	0.79

Table 2
Rejection Rates - Independent Units
 $T = 30, N = 10$, uncorrelated units

	Group t_A	Group t_B		t_{Median}	
		<i>simple</i>	<i>FDB</i>	<i>simple</i>	<i>FDB</i>
α	A. Size: $\phi_i = 1 \forall i$				
0.01	0.00	0.00	0.01	0.01	0.01
0.05	0.01	0.03	0.04	0.04	0.05
0.10	0.03	0.06	0.08	0.06	0.09
0.20	0.04	0.14	0.17	0.17	0.20
α	B. Power: $\phi_i \sim Uniform(0.2, 0.4)$				
0.01	0.97	0.92	0.89	0.90	0.90
0.05	0.99	0.98	0.97	0.98	0.98
0.10	0.99	0.98	0.99	0.99	0.99
0.20	1.00	0.99	1.00	0.99	0.99

Table 3
Rejection Rates - Common Regressor Case
 $T = 30, N = 10$, correlated units, $X_i = X_1 \forall i$

	Group t_A	Group t_B		t_{Median}	
		<i>simple</i>	<i>FDB</i>	<i>simple</i>	<i>FDB</i>
α	A. Size: $\phi_i = 1 \forall i$				
0.01	0.31	0.02	0.03	0.02	0.03
0.05	0.43	0.11	0.10	0.10	0.10
0.10	0.50	0.18	0.16	0.18	0.16
0.20	0.59	0.31	0.27	0.28	0.25
α	B. Power: $\phi_i \sim Uniform(0.2, 0.4)$				
0.01	0.98	0.74	0.58	0.71	0.58
0.05	0.99	0.91	0.85	0.90	0.84
0.10	0.99	0.94	0.90	0.94	0.91
0.20	0.99	0.97	0.95	0.96	0.95

Table 4
Rejection Rates - Large T
 $T = 100, N = 10$, correlated units

	Group t_A	Group t_B		t_{Median}	
		<i>simple</i>	<i>FDB</i>	<i>simple</i>	<i>FDB</i>
α	A. Size: $\phi_i = 1 \forall i$				
0.01	0.07	0.01	0.01	0.00	0.00
0.05	0.14	0.03	0.03	0.02	0.02
0.10	0.18	0.07	0.07	0.06	0.06
0.20	0.26	0.22	0.19	0.19	0.20
α	B. Power: $\phi_i \sim Uniform(0.2, 0.4)$				
0.01	1.00	0.98	0.87	0.90	0.86
0.05	1.00	1.00	0.87	0.99	0.98
0.10	1.00	1.00	1.00	0.99	1.00
0.20	1.00	1.00	1.00	1.00	1.00

Table 5
Rejection Rates - Large N
 $T = 30, N = 30$, correlated units

	Group t_A	Group t_B		t_{Median}	
		<i>simple</i>	<i>FDB</i>	<i>simple</i>	<i>FDB</i>
α	A. Size: $\phi_i = 1 \forall i$				
0.01	0.52	0.00	0.00	0.00	0.00
0.05	0.62	0.06	0.03	0.05	0.03
0.10	0.70	0.14	0.09	0.14	0.10
0.20	0.76	0.30	0.21	0.30	0.20
α	B. Power: $\phi_i \sim Uniform(0.2, 0.4)$				
0.01	0.99	0.65	0.52	0.67	0.54
0.05	1.00	0.90	0.80	0.87	0.79
0.10	1.00	0.95	0.90	0.93	0.88
0.20	1.00	0.97	0.95	0.97	0.93

Table 6
Rejection Rates - Variable cointegrating rank
 $T = 30$, correlated units, cointegration in $0.8N$ units

	Group t_A	Group t_B		t_{Median}	
		<i>simple</i>	<i>FDB</i>	<i>simple</i>	<i>FDB</i>
α	$N = 5$				
0.01	0.19	0.02	0.02	0.03	0.02
0.05	0.33	0.13	0.08	0.12	0.09
0.10	0.40	0.22	0.17	0.21	0.15
0.20	0.51	0.41	0.33	0.35	0.31
α	$N = 10$				
0.01	0.64	0.19	0.14	0.27	0.20
0.05	0.75	0.42	0.33	0.48	0.40
0.10	0.80	0.57	0.48	0.61	0.52
0.20	0.85	0.72	0.64	0.72	0.66

4. Conclusions

The conclusions to be drawn from our simulation study are quite clear-cut. The necessary premise is that ideally panel cointegration estimation and testing should be carried out through a full system approach, as in Groen and Kleinbergen (2003) and Larsson *et al.* (2001). However, this requires sample sizes rarely encountered in practice; hence, single-equation methods have a considerable empirical appeal, though in the case of multivariate studies all the relevant caveats⁴ clearly apply. If the units in the panel are independent (both in the short- and in the long-run) and the time dimension is not small, both in absolute terms and relatively to the cross-section dimension, asymptotic group mean

⁴Well-summarised by Pedroni (1999) as follows: "It is well known that [the cointegrating] vector is generally not unique, and the question of how many cointegrating relationships exist among a certain set of variables is also an important question in many cases.[...] we are implicitly assuming [...] that the researcher has in mind a particular normalization among the variables which is deemed sensible and is simply interested in knowing whether or not the variables are cointegrated".

panel cointegration tests, such as the Group t -statistic examined in this paper, deliver good results. If however, these assumptions do not hold (thus if either there is substantial short-run correlation or long-run dependence, or the number of units is close to that of time observations), the asymptotic test overrejects heavily, in some case disastrously. On the contrary, the bootstrap test put forth in this paper has Type I error close to nominal size a good power properties in all conditions, and can be used with completely heterogenous cointegrating vectors as well as robust statistics such as the median. Hence, it provides a robust, flexible and simple solution to the problem of cointegration testing in panel data.

References

- Bai J. and Ng S. (2002) A panic attack on unit roots and cointegration, *Department of Economics, Boston College*, Unpublished Manuscript.
- Banerjee A., Marcellino M. and Osbat C. (2001) Some cautions on the use of panel methods for integrated series of macro-economic data, Unpublished manuscript, European University Institute.
- Chang Y. (2002) Nonlinear iv unit root tests in panels with cross-sectional dependence, *Journal of Econometrics*, 110, 261–292.
- Chang Y. (2004) Bootstrap unit root tests in panels with cross-sectional dependency, *Journal of Econometrics*, forthcoming.
- Davidson R. and MacKinnon J. (2000) Improving the reliability of bootstrap tests, Discussion Paper 995, Queen's University Institute for Economic Research.
- Groen J. and Kleinbergen F. (2003) Likelihood-based cointegration analysis in panels of vector error correction models, *Journal of Business and Economic Statistics*, 21, 295–318.
- Gutierrez L. (2003) On the power of panel cointegration tests, Unpublished manuscript, University of Sassari.
- Harvey A. and Bates D. (2003) Multivariate unit root tests and testing for convergence, Unpublished manuscript, University of Cambridge.
- Im K., Pesaran H. and Shin Y. (2003) Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, 115, 53–74.
- Kao C. (1999) Spurious regressions and residual-based tests for cointegration in panel data, *Journal of Econometrics*, 90, 1–44.
- Larsson R., Lyhagen J. and Löthgreen M. (2001) Likelihood cointegration tests in heterogeneous panels, *Econometrics Journal*, 4, 109–142.
- Maddala G. and Wu S. (1999) A comparative study of unit root tests with panel data and a new simple test, *Oxford Bulletin of Economics and Statistics*, 61, 631–652.
- Moon H. and Perron B. (2003) Testing for a unit root in panels with dynamic factors, *Journal of Econometrics*, forthcoming.
- Ng S. and Perron P. (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.
- O'Connell P. (1998) The overvaluation of purchasing power parity, *Journal of International Economics*, 44, 1–19.
- Paparoditis E. and Politis D. (2001) The continuous-path block bootstrap, in: *Asymptotics in Statistics and Probability. Papers in honor of George Roussas*, Puri M., ed., VSP Publications, Zeist, The Netherlands, 305–320.

- Paparoditis E. and Politis D. (2003) Residual-based block root testing, *Econometrica*, 71, 813–855.
- Pedroni P. (1997) Cross sectional dependence in cointegration tests of purchasing power parity in panels, Unpublished manuscript, Indiana University.
- Pedroni P. (1999) Critical values for cointegration tests in heterogeneous panels with multiple regressors, *Oxford Bulletin of Economics and Statistics*, 61, 653–670.
- Pesaran M. (2003) A simple panel unit root test in the presence of cross section dependence, Unpublished manuscript, University of Cambridge.
- Phillips P. and Sul D. (2002) Dynamic panel estimation and homogeneity testing under cross section dependence, Discussion Paper 1362, Cowles Foundation, Yale University.
- Smith V., Leybourne S., Kim T.H. and Newbold P. (2004) More powerful panel data unit root tests with an application to mean reversion in real exchange rates, *Journal of Applied Econometrics*, forthcoming.