Maastricht University

Christian Gengenbach, Alain Hecq,
Jean-Pierre Urbain

**Are Panel Unit Root Tests Useful
for Real-Time Data?**

RM/11/012

METEOR

Maastricht University School of Business and Economics
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

Are Panel Unit Root Tests Useful for Real-Time Data?

Christian GENGENBACH **Alain HECQ**[1]

Jean-Pierre URBAIN

Maastricht University, Department of Quantitative Economics

31st January 2011

Abstract

With the development of real-time databases, N vintages are available for T observations instead of a single realization of the time series process. Although the use of panel unit root tests with the aim to gain in efficiency seems obvious, empirical and simulation results shown in this paper heavily mitigate the intuitive perspective.

Keywords: Real-time data, panel unit root tests

JEL: C23, C82, E01

1 Introduction

Before being considered definitive, many series produced and published by statistical offices are provisional, as they are partly based on estimates. They are subsequently subject to revisions when new information is available. Figures for the same sampling period may change over time, resulting in different releases, called vintages, of the same phenomenon. The use of different releases to infer the univariate properties of the process is still an open question, in particular given the problem induced by the presence of common trends across vintages (see e.g. Patterson and Heravi, 1991, 2004, Patterson, 2000, 2002). Patterson and Heravi (2004) mention the potential gain from using panel unit tests. However, they do not investigate this issue further. Therefore, the main purpose of this note is to evaluate the usefulness of second generation panel unit root tests that allow for such global stochastic trends.

A word of notation. We observe N vintages available for T observations each. For a generic variable y that we will adopt in this paper, y_{t-1}^t denotes the figure published at time t of the value of y for time $t-1$. The first difference operator $\Delta = (1-L)$ runs over both indexes with $\Delta y_{t-1}^t = (1-L)y_{t-1}^t = y_{t-1}^t - y_{t-2}^{t-1}$. To make things clearer, Table 1 provides what is called the real-time data matrix for $N=6$ vintages. Each column refers to a vintage series, while the point of time of the referred data are in the rows. We will only consider the diagonals of Table 1, since the vertical sequences only differ on their last few points. For these diagonals, we make the distinction between joint processes of the type $(y_{t-2}^{t-1}, y_{t-2}^t)'$ and $(y_{t-1}^t, y_{t-2}^t)'$, $t = 1 \ldots T$. Hecq and Jacobs (2009) call the first approach *OBS (Observations Balanced System)* and the second *VBS (Vintage Balanced System)*.

Several practical and methodological issues emerge in this framework. First, there are several technical problems with the series, such as the redefinition of the series through time (an entire column changes), or the presence of typos (in individual observations). Moreover, it is not obvious in this setting whether the log transform of the vintages must be taken. Neither is it easy to know to what extent the seasonal adjustment of different vintages creates distortions. Second, and more importantly for our paper, the presence of few common stochastic trends (in the extreme case there are $N-1$ cointegrating vectors among N series), may annihilate the advantage of the existing cross-sectional dimension. In addition to that, the variance-covariance matrix of the disturbance terms for N diagonal releases shows a very particular pattern and will heavily impact the results.

2 Panel unit root tests

Although N vintages are available for the same variable, early panel unit root tests, such as those proposed by Levin *et al.* (2002) (LLC hereafter) or Im *et al.* (2003) (IPS hereafter), are in this framework misleading, as they assume cross-sectional independence (e.g. Banerjee et al., 2004). We will therefore focus mainly on so-called second generation panel unit root tests (see Breitung and Pesaran, 2008). In a detailed simulation study, Gengenbach, Palm and Urbain (2010) compare several of the tests. They found the PANIC approach proposed by Bai and Ng (2004) to be the best in case of cross-member cointegration, while tests based on defactored data tend to reject the null hypothesis of a unit root in this case.

Bai and Ng (2004) propose a procedure to test for a unit root in the common factor and the idiosyncratic component separately. For an observed data series y_{it}, they consider the following model,

$$y_{it} = d_{it} + \sum_{l=1}^{k} \lambda_{il} f_{lt} + v_{it},$$

where d_{it} is a deterministic component, f_{lt} $l = 1, \ldots, k$ are the common factors with corresponding factor loadings λ_{il}, and v_{it} is the idiosyncratic component. The vector of common factors $f_t = (f_{1t}, \ldots, f_{kt})'$ is assumed to follow a process $\Delta f_t = \Phi(L)\eta_t$, where $\eta_t \sim i.i.d.(0, \Sigma_f)$ with $\text{rank}(\Phi(1)) = k_1$. So, f_t potentially contains $k_1 \leq k$ stochastic trends and $k - k_1$ stationary components. The idiosyncratic components are modeled as an AR(1) process $v_{it} = \delta_i v_{it-1} + e_{it}$, where e_{it} follows a mean zero, stationary and invertible MA process. In this setup, the goal of PANIC is to determine the number of non-stationary common factors k_1 and to test whether $\delta_i = 1$ for $i = 1, \ldots, N$. Bai and Ng (2004) propose a principal component estimator for the unobserved idiosyncratic

Table 1: Real-time data matrix: notations

$\frac{vintages}{calendar\ time}$	$t-5$	$t-4$	$t-3$	$t-2$	$t-1$	t
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t-7$	y_{t-7}^{t-5}	y_{t-7}^{t-4}	y_{t-7}^{t-3}	y_{t-7}^{t-2}	y_{t-7}^{t-1}	y_{t-7}^{t}
$t-6$	y_{t-6}^{t-5}	y_{t-6}^{t-4}	y_{t-6}^{t-3}	y_{t-6}^{t-2}	y_{t-6}^{t-1}	y_{t-6}^{t}
$t-5$	$-$	y_{t-5}^{t-4}	y_{t-5}^{t-3}	y_{t-5}^{t-2}	y_{t-5}^{t-1}	y_{t-5}^{t}
$t-4$	$-$	$-$	y_{t-4}^{t-3}	y_{t-4}^{t-2}	y_{t-4}^{t-1}	y_{t-4}^{t}
$t-3$	$-$	$-$	$-$	y_{t-3}^{t-2}	y_{t-3}^{t-1}	y_{t-3}^{t}
$t-2$	$-$	$-$	$-$	$-$	y_{t-2}^{t-1}	y_{t-2}^{t}
$t-1$	$-$	$-$	$-$	$-$	$-$	y_{t-1}^{t}
	$-$	$-$	$-$	$-$	$-$	$-$

component, common factors and factor loadings. Further, they suggest to use an ADF statistic to test for a unit root in an individual series associated with a group mean Fisher $p - value$ approach (denoted $P_{\hat{v}_i}$). To test for unit roots in the extracted common factors, Bai and Ng (2004) suggest an ADF test if only a single common factor has been estimated (denoted as $ADF_{\hat{f}}$), and two modified Q tests if $\hat{k} > 1$.

An alternative is to follow the cross-sectional robust bootstrapped panel unit root tests proposed by Palm, Smeekes and Urbain (2010). They consider pooled Levin $et\ al.$ (2002) type tests and group mean Im $et\ al.$ (2003) type tests, respectively based on the OLS or group mean. Palm $et\ al.$ (2010) show that it is asymptotically valid for a number of cross-sectional correlation models, including cross-sectional cointegration.

3   Monte Carlo Simulations

For the sake of presentation we consider the last three diagonals in a VBS framework. Let us further assume that the variables are $I(1)$ with $r = N - 1$ cointegrating relationships of the type $\beta_i = (1 - 1)$ for $i = 1, ..., N - 1$. The VECM representation for $p = 1$ can be written as follows

$$
\begin{pmatrix} \Delta y_{t-1}^t \\ \Delta y_{t-2}^t \\ \Delta y_{t-3}^t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} y_{t-2}^{t-1} \\ y_{t-3}^{t-1} \\ y_{t-4}^{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}.
$$

The parameters in the Monte Carlo simulations have been calibrated on several systems of real-time data. It emerges that the following restrictions $\alpha_{11} = \alpha_{12} = \alpha_{22} = \alpha_{31} = c_2 = c_3 = 0, \alpha_{21} = \alpha_{32} = 1$ can be imposed, leading to the following structure for the levels

$$
\begin{cases} y_{t-1}^t = c_1 + y_{t-2}^{t-1} + \varepsilon_{1t} \\ y_{t-2}^t = y_{t-2}^{t-1} + \varepsilon_{2t} \\ y_{t-3}^t = y_{t-3}^{t-1} + \varepsilon_{3t} \end{cases}
$$

namely a system with a set of $N - 1$ unbiased revisions and with the first release that follows a random walk with drift. The variance-covariance Σ_ε matrix of the disturbances has by construction a very particular shape. First the variance of ε_{1t} is very high, then variances along the diagonal decrease. Next the correlation between older diagonal vintages is quite high while that covariance is very small with new releases. For instance here is

the covariance matrix we use for 3 series

$$\Sigma_\varepsilon = \begin{bmatrix} 4.11 & 0.74 & 0.67 \\ & 2.71 & 1.10 \\ & & 1 \end{bmatrix}.$$

We consider $M = 1000$ replications; we use $c_1 \sim U[-3, 3]$. To save space we only report the outcome for $T = 15$ and $T = 50$ time observations. We add 50 observations to initialize the processes. For $N = \{2, 5, 10, 20\}$ we compute panel unit root tests on both OBS and VBS systems. We can of course only report OBS results for $T > N$. The number of lags is $l = \{0, 2\}$ in the estimated models.

Although invalid in lightly cross-sectionally dependent setup, Table 2 reports results for the first generation panel unit root tests. Tests with an intercept only as well as those with both an intercept and a linear trend are reported. None of the procedure gives decent results for our DGP even when N, T increase. Sizes are above the nominal 5% for almost all cases. For the IPS statistics, the size seems to increase with the cross-section dimension N. The size of the LLC statistics improves in N, however the tests are still over-sized for all considered panel dimensions.

Table 3 reports results for the Bai and Ng (2004) tests statistics. For the unit root test on the common factor, $ADF_{\hat{f}}$, the rejection frequencies are in general above the 5% nominal size. Furthermore, these size distortions are increasing in both N and T. The test on the idiosyncratic component, $P_{\hat{v}_i}$, which should give a high rejection frequency, also show disappointing results. The power of this statistic ranges from 13.4% to 98.8%.

Finally, Table 4 indicates that Palm *et al.* (2010) procedures are also undersized for small N. Rejection frequencies increase in N leading to some severe size distortions in larger panels. This finding is in line with results reported by Palm *et al.* (2010) for the cross-sectional cointegration case. However, size distortions are more severe using our DGP. Size can be improved by increasing the block length for the bootstrap, which raises the question of optimal block length selection in applications. The results should be interpreted in the light of the variance-covariance matrix that we use to mimic the behavior of vintage dynamics, but which violates assumptions made by Palm *et al.* (2010).

Table 2: Size of first generation unit root tests

| | | T = 15 | | | | T = 50 | | | |
| | | l = 0 | | l = 2 | | l = 0 | | l = 2 | |
		VBS	OBS	VBS	OBS	VBS	OBS	VBS	OBS
N = 2	IPS_c	3.7	3.7	5.2	5.9	2.5	2.3	3.1	3.4
	$IPS_{c,t}$	11.4	11.2	10.8	10.8	12.6	12.4	11.0	11.8
	LLC_c	92.8	88.1	43.0	40.8	99.7	99.9	95.3	96.7
	$LLC_{c,t}$	87.3	81.6	34.9	32.4	99.7	99.9	91.7	94.4
N = 5	IPS_c	15.5	18.7	7.1	12.0	11.4	12.5	5.8	7.7
	$IPS_{c,t}$	43.9	35.7	13.3	17.2	68.5	65.2	26.6	26.0
	LLC_c	84.4	64.3	54.8	33.4	99.0	99.8	98.7	99.3
	$LLC_{c,t}$	78.7	53.6	40.0	24.3	99.0	99.7	98.5	98.1
N = 10	IPS_c	48.9	28.9	14.1	38.2	34.5	38.8	15.7	18.0
	$IPS_{c,t}$	74.3	12.1	19.1	38.1	94.1	90.2	41.6	39.8
	LLC_c	76.0	19.3	48.9	0	99.6	99.4	98.3	98.1
	$LLC_{c,t}$	66.7	8.3	26.4	0.1	99.6	99.2	96.2	94.9
N = 20	IPS_c	93.0	-	27.2	-	84.8	83.1	36.9	42.5
	$IPS_{c,t}$	95.4	-	26.3	-	99.8	97.6	66.1	56.7
	LLC_c	63.2	-	44.5	-	99.9	93.2	95.6	86.9
	$LLC_{c,t}$	49.9	-	21.4	-	99.9	91.7	90.0	76.7

Table 3: Bai and Ng second generation unit root tests: Monte Carlo simulation

| | | T = 15 | | | | T = 50 | | | |
| | | l = 0 | | l = 2 | | l = 0 | | l = 2 | |
		VBS	OBS	VBS	OBS	VBS	OBS	VBS	OBS
N = 2	$ADF_{\hat{f}}$	5.8	5.5	12.8	13.2	2.8	1.7	3.4	2.4
	$P_{\hat{v}_i}$	51.6	49.9	39.0	39.0	60.2	60.5	28.4	30.2
N = 5	$ADF_{\hat{f}}$	45.5	16.5	35.1	28.7	42.0	6.0	15.5	4.8
	$P_{\hat{v}_i}$	26.9	45.7	26.4	48.7	27.4	45.9	17.8	28.2
N = 10	$ADF_{\hat{f}}$	85.9	40.2	56.7	77.5	89.8	20.0	34.4	12.4
	$P_{\hat{v}_i}$	34.3	48.8	33.5	81.7	20.9	80.5	13.4	56.6
N = 20	$ADF_{\hat{f}}$	-	-	-	-	100	57.6	59.3	37.7
	$P_{\hat{v}_i}$	-	-	-	-	61.7	98.8	35.0	87.3

Table 4: Smeekes, Palm and Urbain bootstrap unit root tests: Monte Carlo simulation

		$T = 15; l = 0$		$T = 50; l = 0$	
		VBS	OBS	VBS	OBS
N = 2	LLC_{bs}	0.9	0.4	0.9	0.5
	IPS_{bs}	2.0	1.9	0.3	0.2
N = 5	LLC_{bs}	9.5	6.0	5.3	2.2
	IPS_{bs}	12.3	10.2	2.7	1.3
N = 10	LLC_{bs}	32.6	6.3	18.7	10.7
	IPS_{bs}	40.0	6.5	12.7	8.1
N = 20	LLC_{bs}	-	-	66.3	51.3
	IPS_{bs}	-	-	48.1	40.5

4 Empirical analysis

We consider seasonally adjusted vintages of the quarterly German gross domestic product at constant price. These are from the GERDA, a real time database that contains historical vintages of some 280 economic indicators.[1] We use 17 diagonals with 40 observations each for the VBS and a minimum of 23 balanced observations in the last OBS series. There are quarterly vintages and there is a new base introduced in the release 1999Q2. Between 1999Q2 and 1999Q1 in April, the Bundesbank also published a vintage for the base 1995=100. We use it to convert the series before in 1995=100. We also convert pre-1999Q2 vintage releases to Euros.

Table 5 shows the values of the LLC and IPS test statistics as well as the associated $p - values$. We can see that if we include a time trend, the unit root null hypothesis is not rejected. The null is rejected in the regression with an intercept only for the VBS framework.

We perform the PANIC approach with the number of factors varying from 1 to 4 to investigate whether the results are robust to a misspecification of k, as well as selecting k using the different information criteria[2]. Denote the panels obtained using 17 vintages of quarterly German GDP as OBS_{gdp} and VBS_{gdp}, respectively. OBS_{gdp} has $N = 17$ and $T = 23$ observations while VBS_{gdp} has $N = 17$ and $T = 39$. We consider models with and without a linear deterministic trend. For OBS_{gdp}, IC_2 select one common factor if we include a deterministic linear trend in the model and two factors if we omit the trend. Including a linear trend, we can reject the unit root null hypothesis for the idiosyncratic data component if we include more than 2 common factors in the

---

[1] http://www.bundesbank.de/statistik/statistik_realtime.en.php
[2] We consider the PC_i and IC_i, $i = 1, 2, 3$, as well as the BIC_3 criteria of Bai and Ng (2002). The later one is not asymptotically consistent but performs well in small samples.

Table 5: First generation unit root tests for the quarterly GDP

	levels		Logs	
	VBS	OBS	VBS	OBS
	$T = 646$	$T = 510$	$T = 646$	$T = 510$
LLC_c	-1.81 (0.03)	-0.41 (0.33)	-2.48 (0.006)	-0.66 (0.25)
IPS_c	3.71 (0.99)	4.95 (1.00)	3.09 (0.99)	4.70 (1)
$LLC_{c,t}$	1.56 (0.94)	1.89 (0.97)	1.34 (0.91)	1.87 (0.96)
$IPS_{c,t}$	0.71 (0.76)	0.94 (0.82)	0.33 (0.63)	0.83 (0.79)

model. When we omit the deterministic trend, we can reject the unit root null for all k. For the VBS_{gdp} panel, the PC criteria select the maximum number of possible factors, $\hat{k} = 4$, while the IC criteria select the minimum number, $\hat{k} = 1$. In the model without deterministic trend it is possible to reject the unit root null hypothesis for the idiosyncratic component, but we cannot reject the null if we include a deterministic trend. For a single common factor, the unit root null is rejected for the deterministic trend case but not without a trend.

Using the LLC and IPS bootstrap tests we cannot reject the null hypothesis that the GDP is I(1). This is the case for all considered specifications, using OBS and VBS panels, logarithms or level data, and models with and without deterministic trend. Bootstrap $p-values$ range from 0.149 to 1.

5 Conclusion

It is intuitively an interesting idea to look at real-time data with the aim to improve the size and the power of unit root tests by using several vintages in a panel, as e.g. suggested by Patterson and Heravi (2004). However there are many new issues that this type of observations imply. Even assuming that the data a free of typos and have been correctly rebased, the large number of cross-member cointegrating relationships makes the first generation of panel unit root tests invalid.

Because of the potential existence of $N-1$ cross-member cointegrating relationships we are left with a unique common stochastic trend of dimension T. While the panel unit root test of Bai and Ng (2004) is theoretically equipped to handle a situation where the non-stationarity is due to a single common factor, the power of the unit root test on the common factor in that situation does not improve over univariate tests on a single vintage. Additionally, we obtain a panel unit root test for the idiosyncratic data component, which we view as a proxy for the measurement error. This statistic enables us to gain insight in the consistency of the revision process. However, in simulations the power of this statistics is quite sensitive to the actual specifications of the sample.

It seems that working with OBS diagonals, namely for the same calendar period, gives more plausible results even though the T dimension is reduced in balanced panels.

The bootstrap panel unit root tests of Palm *et al.* (2010) are also theoretically able to handle cross-member cointegration. However, in the extreme case of a single common stochastic trend driving the non-stationarity of the data they suffer from size distortions and reduced power. Our simulation results are complementary to the ones obtained by Palm *et al.* (2010) in that regard.

While it seems intuitive to use panel techniques to improve unit root tests for real-time data, it appears that existing methods are not yet equipped to handle the strong dependence between individual vintages, induced by the peculiar DGP underlying such data sets.

References

[1] BAI, J. AND S. NG (2004), A PANIC Attack on Unit Roots and Cointegration, *Econometrica* 72, 1127-1177.

[2] BAI, J. AND S. NG (2002), Determining the number of Factors in Approximate Factor Models, *Econometrica* 70, 191-221.

[3] BANERJEE, A., M. MARCELLINO AND C. OSBAT (2004), Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data, *Econometrics Journal* 7, 322-340.

[4] GENGENBACH, C., F.C. PALM AND J.-P. URBAIN (2010), Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling, *Econometric Reviews*, 29, 111-145.

[5] HECQ, A. AND J.P.A.M. JACOBS (2009), On the VAR-VECM Representation of Real Time Data, mimeo, Maastricht University.

[6] IM, K.S., M.H. PESARAN AND Y. SHIN (2003), Testing for unit roots in heterogenous panels, *Journal of Econometrics* 115, 53-74.

[7] LEVIN, A., C.F. LIN AND C.S.J. CHU (2002), Unit root tests in panel data: asymptotic and finite-sample properties, *Journal of Econometrics* 108, 1-24.

[8] PATTERSON, K.D. (2000), Which Vintage of Data to use When There are Multiple Vintages of Data? Cointegration, Weak Exogeneity and Common factors, *Economics Letters*, 69, 115-121..

[9] PATTERSON, K.D. (2002), The Data Measurement Process for UK GNP: Stochastic Trends, Long Memory, and Unit Roots, *Journal of Forecasting*, 21, 245-264.

[10] PATTERSON, K.D. AND S.M. HERAVI (1991), Are Different Vintages of Data on the Components of GDP Co-Integrated?, *Economics Letters*, 35, 409-413.

[11] PATTERSON, K.D. AND S.M. HERAVI (2004), Revisions to Official Data on U.S. GDP: A Multivariate Assessment of Different Vintages, *Journal of Official Statistics*, 20, 573-602.