What would Nelson and Plosser find had they used panel unit root tests?*

Christophe Hurlin †

Revised Version. June 2007

Abstract

In this study, we systemically apply nine recent panel unit root tests to the same fourteen macroeconomic and financial series as those considered in the seminal paper by Nelson and Plosser (1982). The data cover OECD countries from 1950 to 2003. Our results clearly point out the difficulty that applied econometricians would face when they want to get a simple and clear-cut diagnosis with panel unit root tests. We confirm the fact that panel methods must be very carefully used for testing unit roots in macroeconomic or financial panels. More precisely, we find mitigated results under the cross-sectional independence assumption, since the unit root hypothesis is rejected for many macroeconomic variables. When international cross-correlations are taken into account, conclusions depend on the specification of these cross-sectional dependencies. Two groups of tests can be distinguished. The first group tests are based on a dynamic factor structure or an error component model. In this case, the non stationarity of common factors (international business cycles or growth trends) is not rejected, but the results are less clear with respect to idiosyncratic components. The second group tests are based on more general specifications. Their results are globally more favourable to the unit root assumption.

- *Key Words* : Unit Root Tests, Panel Data.
- *J.E.L Classification* : C23, C33

*I am grateful to Benoit Perron, Rafal Kierzenkowski and Regis Breton for helpful comments and suggestions. I am grateful to Faouzi Boujedra for research assistance. All the tests used have been programmed under Matlab. Codes and data are available on the website: http://www.univ-orleans.fr/deg/masters/ESA/CH/churlin_R.htm

†LEO, University of Orléans. Rue de Blois. BP 6739. 45067 Orléans Cedex 2. France. E-mail address: christophe.hurlin@univ-orleans.fr.

1 Introduction

Many studies have been devoted to the finite sample properties of panel unit root tests. In this literature, it is now standard to distinguish first generation tests that are based on the assumption of independent cross section units and second generation tests that allow for cross-section dependence (see Banerjee, 1999; Baltagi and Kao, 2000; Choi, 2004; Hurlin and Mignon, 2004; Breitung and Pesaran, 2005 for a survey). The empirical power and size of first generation unit root tests have been simulated under various assumptions in Maddala and Wu (1999), Levin, Lin and Chu (2002), Choi (1999, 2001), Breitung (2000), Banerjee, Marcellino and Osbat (2005). The relative performance of second generation unit root tests has been studied in particular by Gutierrez (2003, 2005), Gengenbach, Palm, and Urbain (2004) or Baltagi, Bresson and Pirotte (2005). The results of these studies depend very much on the underlying data generating process used in the Monte Carlo simulations. In particular, it appears that the finite sample properties of panel unit root tests depend on (*i*) the homogeneity assumption used under the alternative, (*ii*) the existence of cross-section dependences, (*iii*) the specification of these cross-section dependences (factor structure or weak cross section dependence), (*iv*) the relative sample sizes T and N, (*v*) the existence of long-run cross-unit relationships, etc.. In many configurations, the panel unit root tests have severe size biases in finite samples. In some cases, the empirical size of the tests is substantially higher than the nominal level, so that the null hypothesis of a unit root is rejected very often, even if correct. It is for instance the case when the assumption of no cross-unit correlation or cross-unit cointegrating is violated and first generation unit root tests are used (Banerjee, Marcellino and Osbat, 2005). But, it may be also the case when second generation unit root tests are used in a context of cross-unit dependences for which they are not designed. The use of a factor model (Bai and Ng, 2004) in the case of weak correlation may do not yield valid test procedures. On the contrary, the use of unit root tests that allow for weak dependence may also lead to severe size biases in some cases[1]. Given these results, some authors warn against the

[1] It is for instance the case, when the cross-section dependences are specified as standard spatial error processes (Baltagi, Bresson and Pirotte, 2005) and nonlinear IV tests (Chang, 2002) or bootstrap based tests (Chang, 2004) are used.

use of panel methods for testing for unit roots in some cases. In particular, Banerjee, Marcellino and Osbat (2005) clearly ask the question: should we use panel methods for testing for PPP?

In this perspective, our paper aims to evaluate the advantages and drawbacks of panel unit root tests for macroeconomic and financial series. But, our methodological approach is different from the approaches previously mentioned. Rather than simulating some Monte Carlo experiments and evaluating empirical size and power in many different configurations, we propose here to respond to the question: what results would Nelson and Plosser (1982) obtain if they have used panel unit root tests[2]? For that, we systemically apply panel unit root tests to the same 14 macroeconomic and financial variables (including measures of output spending, money, prices and interest rates) as those studied by Nelson and Plosser. The series are considered for a panel of OECD countries over the period 1950-2003 given data availability. More precisely, we consider nine panel unit root tests among the most used in the literature. Four first generation tests are studied: (*i*) the tests of Levin, Lin and Chu (2002) based on a homogenous alternative assumption, (*ii*) the tests of Im, Pesaran and Shin (2003) that allow for a heterogeneous alternative and the Fisher type tests of (*iii*) Maddala and Wu (1999) and (*iv*) Choi (2001). The common feature of these first generation tests is the restriction that all cross-sections are independent. However, it is well-known that this cross-unit independence assumption is quite restrictive in many empirical applications. So, we also consider some second generation unit root tests that allow cross-unit dependencies. A growing literature is now devoted to these tests with among others, the papers by Bai and Ng (2004), Choi (2002), Phillips and Sul (2003), Moon and Perron (2004), Pesaran (2004) and Chang (2002, 2004). The main issue is to specify the cross-sectional dependencies, since as pointed out by Quah (1994), individual observations in a cross-section have no natural ordering. Consequently, various specifications and a lot of different testing procedures have been proposed. In our study, we consider two groups of tests. The first group tests are based on a dynamic factor model (Bai and Ng, 2004; Moon

[2] Based on standard time series unit root tests, the seminal paper by Nelson and Plosser pointed out that American macroeconomic series feature, quasi systematically, stochastic tendencies and unit root properties.

and Perron, 2004; Pesaran, 2003) or an error-component model (Choi, 2002). The cross-sectional dependency is then due to the presence of one or more common factors or to a random time effect. The tests of the second group are defined by opposition to these specifications based on common factor or time effects. In this group, some specific (O'Connell, 1998; Taylor and Sarno, 1998) or more general (Chang, 2002 and 2004) specifications of the cross-sectional correlations are proposed in the literature. Here, we limit our analysis to the IV nonlinear test proposed by Chang (2002).

Contrary to a standard finite sample exercise based on Monte Carlo simulations, our approach does not allow determining what the most robust test is (since we do not know the true data generating process). We can not give some recommendations about the appropriate circumstances for using each test. However, our study clearly points out the difficulty that applied econometricians would face when they want to get a simple and clear-cut diagnosis with panel unit root tests. They confirm the fact that one must be very careful with the use of panel root tests on macroeconomic time series. In particular, we highlight the influence of (*i*) the heterogeneous specification of the model, (*ii*) the cross-sectional independence assumption and the (*iii*) the specification of these dependences.

The rest of the paper is organized as follows. In section 2, we present the data and the results of three first generation tests. Section 3 presents the results obtained from five representative tests of the second generation. The last section concludes.

2 First Generation Unit Root Tests

What would Nelson and Plosser find had they used first generation panel unit root tests? To answer this question, we consider in this study the same series as those used in the seminal paper of Nelson and Plosser (1982). The only difference is that we consider these series for a panel of OECD countries[3]. The data are annual with starting dates from 1952 to 1971 and ending dates from 2000 to 2003. This sample size

[3]The second slight difference is that we consider GDP (and GDP per capita, real GDP) rather than GNP for data availability.

is relatively short compared to the sample used initially for the United States by Nelson and Plosser (from 1860 or 1909 to 1970 according to the series), but it corresponds to that of most of macroeconomic or financial panels. Since some panel unit root tests require the use of a balanced panel, for each variable we consider the same balanced database for all the tests. The data includes the maximum of OECD countries given the data availability. The lists of countries and data sources are reported in appendix A.

2.1 Levin and Lin unit root tests

One of the most popular first generation unit root test is undoubtedly the test proposed by Levin, Lin and Chu (2002). Let us consider a variable observed on N countries and T periods and a model with individual effects and no time trends. As it is well known, Levin and Lin (LL thereafter) consider a model in which the coefficient of the lagged dependent variable is restricted to be homogenous across all units of the panel:

$$\Delta y_{it} = \alpha_i + \rho \, y_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta y_{i,t-z} + \varepsilon_{it} \tag{1}$$

for $i = 1,..,N$ and $t = 1,..,T$. The errors ε_{it} $i.i.d.$ $\left(0, \sigma^2_{\varepsilon_i}\right)$ are assumed to be independent across the units of the sample. In this model, LL are interested in testing the null hypothesis $H_0 : \rho = 0$ against the alternative hypothesis $H_1 : \rho = \rho_i < 0$ for all $i = 1,..N$, with auxiliary assumptions about the individual effects ($\alpha_i = 0$ for all $i = 1,..N$ under H_0). This restrictive alternative hypothesis implies that the autoregressive parameters are identical across the panel.

In a model with individual effects, the standard t-statistic t_ρ based on the pooled estimator $\widehat{\rho}$ diverges to negative infinity. That is why, LL suggest using the following adjusted t-statistic:

$$t^*_\rho = \frac{t_\rho}{\sigma^*_T} - N \, T \, \widehat{S}_N \, \left(\frac{\widehat{\sigma}_{\widehat{\rho}}}{\widehat{\sigma}^2_{\widehat{\varepsilon}}} \right) \left(\frac{\mu^*_T}{\sigma^*_T} \right) \tag{2}$$

where the mean adjustment μ^*_T and standard deviation adjustment σ^*_T are simulated by authors (Levin, Lin and Chu, 2002, table 2) for various sample sizes T. The adjustment term is also function of the average of individual ratios of long-run to short-run variances, $\widehat{S}_N = (1/N) \sum_{i=1}^{N} \widehat{\sigma}_{y_i} / \widehat{\sigma}_{\varepsilon_i}$, where $\widehat{\sigma}_{y_i}$ denotes a kernel estimator of the long-run variance for the country i. LL suggest using a Bartlett kernel function and a homogeneous truncation lag parameter given by the simple formula $\overline{K} = 3.21T^{1/3}$. They demonstrate that, under the non stationary null hypothesis, the adjusted t-statistic t_ρ^* converges to a standard normal distribution.

The results of the LL tests in a model with individual effects are reported in table 1. In order to assess the sensitivity of the results to the choice of the kernel function and the selection of the bandwidth parameters in the adjustment terms, for each variable we compute three statistics. The first one, denoted t_ρ^*, is based on a Bartlett kernel function and the common lag truncation parameter $\overline{K}$ proposed by LL. The second statistic, denoted t_ρ^{*B}, is also based on a Bartlett kernel but with individual lag truncation parameters selected according the Newey and West's procedure (1994). The last statistic, denoted t_ρ^{*C}, is computed with a Quadratic Spectral kernel and individual lag truncation parameters. Finally, we also consider a model with individual effects and deterministic trends to asses the sensitivity of our results to the specification of the deterministic component. The corresponding adjusted t-statistic based on a Bartlett kernel is denoted t_3^{*C}.

The results would have surprised Nelson and Plosser. The LL tests clearly indicate that stationarity is a common feature of the main macroeconomic variables. Indeed, at a 5% significance level, the tests strongly reject the null of non stationarity for 11 macroeconomic series out of 14, including real GDP, nominal GDP, employment etc. The unit root hypothesis is not rejected only for bond yield, common stock prices and velocity. Besides, except for velocity, these results are robust to the choice of kernel function and bandwidth parameter. These conclusions, except for the unemployment rate and the money stock, are also robust to the specification of the deterministic component, *i.e.* with or without time trends. Various explanations to these surprising results are possible. The first one is based on the mispecification of one or more on the N individual ADF lag lengths in the model (1). Im, Pesaran and Shin (2003) show the importance of correctly choosing these individual lag orders for the LL tests. In our study, individual lag lengths are optimally chosen using the general-to-specific

(GS) procedure of Hall (1994) with a maximum lag length set to 4. However, similar qualitative results (not reported) are obtained when individual lag lengths are chosen by information criteria (AIC or BIC).

Then, others explanations must be evoked. The second one is linked to the restrictive homogeneous assumption used in LL. In particular, this assumption implies that all panel members are forced to have identical orders of integration. The null hypothesis is that all series contain a unit root, while the alternative hypothesis is that no series contains a unit root, that is, all are stationary. Then with as few as one stationary series, the rejection rate rises above the nominal size of the test, and continues to increase with the number of $I(0)$ series in the panel.

2.2 Heterogeneous Panel Unit Root Tests

At this stage, the question is: if Nelson and Plosser would have used heterogeneous panel unit root tests, would they have also concluded to non stationarity of macroeconomic and financial variables? In order to answer this question, we consider two heterogeneous tests based on the cross-sectional independence assumption: the well-known test proposed by Im, Pesaran and Shin (2003) (IPS thereafter) and two equivalent Fisher type tests (Choi, 1999; Maddala and Wu, 1999). It is well known that the main advantage of these test compared to LL one, is to allow for heterogeneity in the value of ρ_i under the alternative hypothesis. The corresponding model with individual effects and no time trend becomes:

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta y_{i,t-z} + \varepsilon_{it} \tag{3}$$

The null hypothesis is defined as $H_0 : \rho_i = 0$ for all $i = 1, ..N$ and the alternative hypothesis is $H_1 : \rho_i < 0$ for $i = 1, ..N_1$ and $\rho_i = 0$ for $i = N_1+1, .., N$, with $0 < N_1 \leq N$. The alternative hypothesis allows unit roots for some (but not all) of the individual. In this context, the IPS test is based on the (augmented) Dickey-Fuller statistics averaged across groups. Let $t_{iT}(p_i, \beta_i)$ with $\beta_i = \left(\beta_{i,1}, .., \beta_{i,p_i}\right)$ denote the t-statistic for testing unit root in the i^{th} country. The IPS statistic is then defined as:

$$t\_bar_{NT} = \frac{1}{N} \sum_{i=1}^{N} t_{iT} \left(p_i, \beta_i \right) \tag{4}$$

Under the assumption of cross-sectional independence, this statistic is shown to sequentially converge to a normal distribution. IPS propose two corresponding standardized t-bar statistics. The first one, denoted $Z_{t\,bar}$, is based on the asymptotic moments of the Dickey Fuller distribution. The second standardized statistic, denoted W_{tbar}, is based on the means and variances of $t_{iT} \left(p_i, 0 \right)$ evaluated by simulations under the null $\rho_i = 0$. Although the tests Z_{tbar} and W_{tbar} are asymptotically equivalent, simulations show that the W_{tbar} statistic, which explicitly takes into account the underlying ADF orders in computing the mean and the variance adjustment factors, performs much better in small samples. In table 2, both statistics are reported. We also report the value of the W_{tbar} statistic in a model with deterministic trend. For each country, the values of the mean and variance used in the standardization of W_{tbar} are taken from the IPS simulations (Im, Pesaran and Shin, 2003, table 3) for the time length T and the corresponding individual lag order p_i. Individual ADF lag orders are optimally chosen according to the same GS method as that used for LL tests. In order to asses the sensitivity of our results to the choice of the lag orders, we also report the value of the standardized t-bar statistic based on Dickey-Fuller statistics ($p_i = 0, \forall i$) and denoted $Z_{t\,bar}^{DF}$.

Using IPS tests, Nelson and Plosser would have obtain mitigate results. If we consider the standardized statistic W_{tbar}, the unit root hypothesis is not rejected for 8 macroeconomic variables out of 14 at a 5% significance level: nominal GDP, real per capita GDP, employment GDP deflator, consumer prices, velocity, bond yield and common stock prices. We find in these subset, the three variables for which the LL tests do not reject the null hypothesis. Except for the nominal GDP, the results are robust to the use the standardized statistic Z_{tbar} based on asymptotic moments instead of W_{tbar}. More surprising, except for the nominal GDP and the unemployment rate, the results are also robust when we consider the statistic $Z_{t\,bar}^{DF}$ based on the average of Dickey Fuller individuals statistics. Finally, the results are globally robust to the specification of deterministic component. With time trends, the null hypothesis is not rejected for four other variables (industrial production, unemployment rate, money stock and real wages) whereas the null is now rejected for the velocity.

Special care need to be exercised when interpreting the results of the six variables for which the null hypothesis is rejected (real GDP, industrial production, unemployment rate, wages, real wages and money stock). Due to the heterogeneous nature of the alternative, rejection of the null hypothesis does not necessarily imply that the non stationarity is rejected for all countries, but only that the null hypothesis is rejected for a sub-group of $N_1 < N$ countries. Therefore, such a result is not incompatible with the fact that, based on pure time series, the ADF tests lead to accept the non stationarity hypothesis for the majority of OECD countries. For instance, let us consider the real GDP over the period 1963-2003, for which the IPS leads to rejection of the non stationarity hypothesis. At a 5% significance level, the pure time series ADF tests conclude to the presence of a unit root in 17 out of 25 GDP processes (see table 8).

These conclusions are confirmed by the Fisher (1932) type tests proposed by Choi (2001) and Maddala and Wu (1999). The null and alternative assumptions are the same as in IPS. But in these tests, the strategy consists in combining the observed significant levels from the unit root individual tests. Let us consider pure time series unit root test statistics (ADF, ERS, Max-ADF etc.). Since these statistics are continuous, the corresponding p-values, denoted p_i, are uniform $[0,1]$ variables. Consequently, under the assumption of cross-sectional independence, the statistic proposed by Maddala and Wu (1999) and defined as:

$$P_{MW} = -2\sum_{i=1}^{N} \log (p_i) \qquad (5)$$

has a chi-square distribution with $2N$ degrees of freedom, when T tends to infinity and N is fixed. As noted by Banerjee (1999), the obvious simplicity of this test and its robustness to statistic choice, lag length and sample size make it extremely attractive. For large N samples, Choi (2001) proposes a similar standardized statistic:

$$Z_{MW} = -\frac{\sum_{i=1}^{N} \log (p_i) + N}{\sqrt{N}} \qquad (6)$$

Under the cross-sectional independence assumption, Z_{MW} converges to a standard normal distribution under the unit root hypothesis. For each macroeconomic variable, we compute both statistics P_{MW} and Z_{MW} based on individual ADF tests. We also consider the same statistics computed in a model with time trend. The results are reported on table 3. They globally confirm our previous conclusions. It is not surprising, since the Fisher tests based on ADF and the IPS tests are directly comparable. The crucial element that distinguishes the two tests is that the Fisher test is based on combining the significance levels of the different tests, and the IPS is based on combining the test statistics. However, these tests are similar in the sense that they combine independent individual tests. If we consider the P_{MW} test at a 5% significant level, we do not reject the unit root for 7 out of 14 variables. The only difference with the IPS results is for the nominal GDP, for which we reject the null here. This is precisely the only variable for which the two IPS standardized statistics, W_{tbar} and Z_{tbar}, do not give the same conclusions. Except for the real per capita GDP, the conclusions are identical with the Choi's standardized statistic. Besides, the results are globally robust to the specification of the deterministic component, except for industrial production, nominal GDP and money.

In summary, with the heterogeneous panel unit root tests based on the cross-sectional independence assumption, the conclusions on the non stationarity of OECD macroeconomic variables are no clear-cut. The unit root hypothesis is strongly rejected for four macroeconomic variables (real GDP, wages, real wages and money stocks), which are generally considered as non stationary for the most of OECD countries. The non stationarity is also rejected for the unemployment rate as in Nelson and Plosser (1982). The non stationarity is robust to the choice of the test and the choice of the standardization only for six variables: employment, GDP deflator, consumer prices, velocity, bond yield and common stock prices. So, we are far from the clear-cut results obtained by Nelson and Plosser for the United States. The issue is then to know if these surprising results are due to the restrictive assumption of cross-sectional independence.

3 Second Generation Unit Root Tests

The second generation unit root tests relax the cross-sectional independence assumption. Then, the issue is to specify these cross-sectional dependencies. The simplest way consists in using a factor structure model. At least three panel unit root tests based on this approach have been proposed: Phillips and Sul (2003), Bai and Ng (2004), Moon and Perron (2004). For all these tests, the idea is to shift data into two unobserved components: one with the characteristic that is cross-sectionally correlated and one with the characteristic that is largely unit specific. Thus, the testing procedure consists in two steps: in a first one, data are de-factored, and in a second step, panel unit root test statistics based on de-factored data and/or common factors are then proposed. The issue is to know if this factor structure allows obtaining clear cut conclusions about stationarity of macroeconomic variables.

3.1 Bai and Ng unit root tests

In this context, the unit root tests by Bai and Ng (2004) provide a complete procedure to test the degree of integration of series. They decompose a series y_{it} as a sum of three components: a deterministic one, a common component expressed as a factor structure and an error that is largely idiosyncratic. The process y_{it} is non-stationary if one or more of the common factors are non-stationary, or the idiosyncratic error is non-stationary, or both. Instead of testing for the presence of a unit root directly in y_{it}, Bai and Ng propose to test the common factors and the idiosyncratic components separately. Let us consider a model with individual effects and no time trend:

$$y_{it} = \alpha_i + \lambda_i' F_t + e_{it} \tag{7}$$

where F_t is a $r \times 1$ vector of common factors and λ_i is a vector of factor loadings. Among the r common factors, we allow r_0 stationary factors and r_1 stochastic common trends with $r_0 + r_1 = r$. The corresponding model in first differences is:

$$\Delta y_{it} = \lambda_i' f_t + z_{it} \tag{8}$$

where $z_{it} = \Delta e_{it}$ and $f_t = \Delta F_t$ with $E(f_t) = 0$. The common factors in Δy_{it} are estimated by the principal component method. Let us denote $\widehat{f}_t$ these estimates, $\widehat{\lambda}_i$ the corresponding loading factors and $\widehat{z}_{it}$ the estimated residuals. Bai and Ng propose a differencing and re-cumulating estimation procedure which is based on the cumulated variables:

$$\hat{F}_{mt} = \sum_{s=2}^{t} \hat{f}_{ms} \qquad \hat{e}_{it} = \sum_{s=2}^{t} \widehat{z}_{is} \tag{9}$$

for $m = 1, .., r$ and $i = 1, .., N$. Then, they test the unit root hypothesis in the idiosyncratic component e_{it} and in the common factors F_t with the estimated variables $\hat{F}_{m\,t}$ and $\hat{e}_{i\,t}$.

In order to test the non stationarity of idiosyncratic components, Bai and Ng propose to pool individual ADF t-statistics computed with the de-factored estimated components $\hat{e}_{it}$ in a model with no deterministic terms. Let $ADF^{c}_{\widehat{e}}(i)$ be the ADF t-statistic for the idiosyncratic component of the i^{th} country. The asymptotic distribution of $ADF^{c}_{\widehat{e}}(i)$ coincides with the Dickey Fuller distribution for the case of no constant. Therefore, a unit root test can be done for each idiosyncratic component of the panel. The great difference with unit root tests based on the pure time series is that the common factors, as global international trends or international business cycles for instance, have been withdrawn from data. In order to asses the importance of this transformation, both individual ADF test results are compared for the real GDP, on table 8. For 12 countries, the conclusions of both tests are opposite at a 5% significant level: for 8 countries, the ADF tests on the initial series lead to reject the null, whereas the idiosyncratic component is founded to be non-stationary. However, these individual time series tests have the same low power as those based on initial series. That is why, pooled tests (similar to the first generation ones) are also proposed. But in this case, estimated idiosyncratic components $\hat{e}_{i,t}$ are asymptotically independent across units. Bai and Ng consider two Fisher's type statistics, respectively denoted $P^{c}_{\widehat{e}}$ and $Z^{c}_{\widehat{e}}$. The second one corresponds to a standardized Choi's type statistic. The results are reported on table 4. At a 5% significant level, the non stationarity of idiosyncratic components is not rejected only for 6 out of 14 variables: industrial production, employment, consumer prices, real wages, velocity and common stock prices. For the eight others variables, including real and nominal GDP, the null is strongly rejected.

In the Bai and Ng's perspective, the rejection of the non stationarity of the idiosyncratic component does not imply that the series are stationary, since the common factors may be non-stationary. In order to test the non-stationarity of the common factors, Bai and Ng (2004) distinguish two cases. When there is only one common factor among the N variables ($r = 1$), they use a standard ADF test in a model with an intercept. The corresponding ADF t-statistic, denoted $ADF^c_{\widehat{F}}$, has the same limiting distribution as the Dickey Fuller test for the constant only case. If there are more than one common factors ($r > 1$), Bai and Ng test the number of common independent stochastic trends in these common factors, denoted r_1. Naturally, if $r_1 = 0$ it implies that there are N cointegrating vectors for N common factors, and that all factors are $I(0)$.

In order to determine r_1, Bai and Ng propose a sequential procedure based on two statistics. The first statistic of test, denoted MQ_f, assumes that the non-stationary components are finite order vector-autoregressive processes. The second statistic, denoted MQ_c, allows the unit root processes to have more general dynamics. The corresponding results are reported on table 4. For each variable, the number of common factors is estimated according to IC_2 or BIC_3 criteria (see Bai and Ng, 2002) with a maximum number of factor equal to 5. Given these criteria, there is only one common factor in real GDP and in real per capita GDP, which can be analyzed as an international stochastic growth factor. For both variables, this common factor is found to be non stationary. For all other variables, the estimated number of common factors ranges from 2 to 4. Whatever the test used, MQ_c or MQ_f, the number of common stochastic trends is always equal to the number of common factors. So, it seems that for all macroeconomic variables, except for the real GDP, at least two independent non stationary common factors can be identified among OECD countries. The conclusions are globally in favour of non stationarity for all financial and macroeconomic variables. More precisely, we found that if the macroeconomic series are non-stationary, this property seems to be more due to the common factors, as international business cycles or growth trends, than to the idiosyncratic components.

3.2 Moon and Perron unit root tests

Moon and Perron (2004) also use a factor structure to model cross-sectional dependence. Their model is slightly different from that used by Bai and Ng (2004), since they assume that the error terms are generated by r common factors and idiosyncratic shocks.

$$y_{it} = \alpha_i + y_{it}^0 \tag{10}$$

$$y_{it}^0 = \rho_i \, y_{i,t-1}^0 + \mu_{it} \tag{11}$$

$$\mu_{it} = \lambda_i' F_t + e_{it} \tag{12}$$

where F_t is a $r \times 1$ vector of common factors and λ_i is a vector of factor loadings. The idiosyncratic component e_{it} is assumed to be $i.i.d.$ across i and over t. The null hypothesis corresponds to the unit root hypothesis $H_0 : \rho_i = 1, \forall i = 1, .., N$ whereas under the alternative the variable y_{it} is stationary for at least one cross-sectional unit. The testing procedure is the same as in Bai and Ng: in a first step, data are de-factored, and in a second step, panel unit root test statistics based on de-factored data are proposed. The main difference is that the Moon and Perron unit root test is only based on the estimated idiosyncratic components.

Moon and Perron treat the factors as nuisance parameters and suggest pooling de-factored data to construct a unit root test. The intuition is as follows. In order to eliminate the common factors, panel data must be projected onto the space orthogonal to the factor loadings. So, the de-factored data and the de-factored residual no longer have cross-sectional dependencies. Then, it is possible to define standard pooled t-statistics, as in IPS, and to show their asymptotic normality. Let $\widehat{\rho}_{pool}^{+}$ be the modified pooled OLS estimator using the de-factored panel data. Moon and Perron define two modified t-statistics which have a standard normal distribution under the null hypothesis:

$$t_a = \frac{T\sqrt{N}\left(\widehat{\rho}_{pool}^{+} - 1\right)}{\sqrt{2\gamma_e^4 / w_e^4}} \underset{T,N\rightarrow\infty}{\overset{d}{\longrightarrow}} N\left(0,1\right) \tag{13}$$

$$t_b = T\sqrt{N}\left(\widehat{\rho}_{pool}^{+} - 1\right)\sqrt{\frac{1}{NT^2}\text{trace}\left(Z_{-1} Q_\Lambda Z_{-1}'\right)\frac{w_e^2}{\gamma_e^4}} \underset{T,N\rightarrow\infty}{\overset{d}{\longrightarrow}} N\left(0,1\right) \tag{14}$$

where w_e^2 denotes the cross-sectional average of the long-run variances $w_{e_i}^2$ of residuals e_{it} and γ_e^4 denotes the cross-sectional average of $w_{e_i}^4$. Moon and Perron propose feasible statistics t_a^* and t_b^* based on an estimator of the projection matrix and estimators of long-run variances $w_{e_i}^2$. The corresponding results are reported on table 5. For each variable, the number of common factors r is estimated according to the same[4] criteria (IC_2 or BIC_3) used for the Bai and Ng (2004) unit root test (see table 4). In order to asses the sensitivity of our results to the choice of the kernel function used to estimate $w_{e_i}^2$, we compute both statistics t_a^* and t_b^* with a Bartlett and with a Quadratic Spectral kernel. In both cases, bandwidth parameters are optimally chosen according to the Newey and West (1994) procedure. Besides, the results for a model with time trends are reported. The computation of unit root test statistic, denoted $t_a^{\#}$, is then slightly different from this presented above (see Moon and Perron, 2004). In this case, we use the same criteria to estimate the number of common factor as in the model with individual effects only.

In a model with individual effects, the null is strongly rejected for all variables. The only exceptions are the real GDP and wages when the t_b^* statistic is considered. These results confirm the rejection of non-stationarity of idiosyncratic components, when they are defined in a factor structure model (see Fisher' type statistics $Z_{\widehat{e}}^c$ and $P_{\widehat{e}}^c$, table 4). However, this rejection is not robust to the specification of the deterministic component. When time trends are included in the model, the conclusions are more in favour of the unit root hypothesis. The unit root is not rejected for nine variables and particularly for the real GDP and the real per capita GDP.

3.3 Choi unit root tests

Unlike in previous tests, Choi (2002) uses an error-component model to specify the cross sectional correlations. In spite of this first difference, his testing procedure is similar to those developed in Bai and Ng (2004) or in Moon and Perron (2004). However, the method used to eliminate non-stochastic trend components and cross-sectional cor-

[4]The corresponding estimated numbers of factors are exactly the same except for employment and velocity. This slight difference is due to the fact that in Bai and Ng (2004) the information criteria are computed from demeaned first differences whereas in Moon and Perron the residuals $\widehat{y}$ are used.

relations is appreciably different from previous ones. Indeed, the second originality of Choi's unit root tests is that cross-sectional correlations and deterministic components are eliminated by GLS-based detrending (Elliott, Rothenberg and Stock, 1996; ERS thereafter) and conventional cross-sectional demeaning for panel data. Let us consider a model defined as:

$$y_{it} = \alpha_i + f_t + v_{it} \tag{15}$$

$$v_{it} = \sum_{j=1}^{q_i} d_{i,j} v_{i,t-j} + \varepsilon_{it} \tag{16}$$

where ε_{it} are $i.i.d.\left(0, \sigma_{\varepsilon,i}^2\right)$ and assumed to be cross-sectional independent. α_i and f_t respectively denote the unobservable individual effect and the unobservable time effect. The null hypothesis corresponds to the presence of a unit root in the remaining random component v_{it}, i.e. $H_0 : \sum_{j=1}^{q_i} d_{i,j} = 1, \forall i = 1, .., N$. The alternative hypothesis is $\sum_{j=1}^{q_i} d_{i,j} < 1$ for some cross-section units. The test is constructed by first demeaning the data by GLS as in ERS. Assuming that the largest root of v_{it} is $1 + c/T$ for all i, two quasi-differenced series are built for $t \geq 2$:

$$\widetilde{y}_{it} = y_{it} - \left(1 + \frac{c}{T}\right) y_{i,t-1} \qquad \widetilde{c}_{it} = 1 - \left(1 + \frac{c}{T}\right) \tag{17}$$

To obtain the GLS estimators $\widehat{\alpha}_i$ of parameters α_i, the variable $\widetilde{y}_{it}$ are regressed on $\widetilde{c}_{it}$. Choi suggests here to follow ERS in setting $c = -7$ for all i. In a second step, the residuals $\widetilde{y}_{it} - \widehat{\alpha}_i$ are cross-sectionally demeaned.

$$z_{it} = (y_{it} - \widehat{\alpha}_i) - \frac{1}{N} \sum_{i=1}^{N} (y_{it} - \widehat{\alpha}_i) \tag{18}$$

The deterministic components α_i and f_t are eliminated from y_{it} by the time series and cross-sectional demeaning. It implies that the transformed variables z_{it} are independent across i for large T and N. Then, it is possible to test unit root with the cross-sectional independent variables z_{it}. Choi uses a standard ADF t-statistic based on the regression:

$$\Delta z_{it} = \rho_i z_{i,t-1} + \sum_{j=1}^{q_i - 1} \beta_{i,j} \Delta z_{i,t-j} + u_{it} \tag{19}$$

This statistic, called Dickey-Fuller-GLS statistic, has the Dickey and Fuller distribution. Based on these individual tests, Choi proposes three Fisher's type statistics.

$$P_m = -\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[\ln\left(p_i\right) + 1\right] \tag{20}$$

$$Z = -\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}\left(p_i\right) \tag{21}$$

$$L^* = \frac{1}{\sqrt{\pi^2 N/3}} \sum_{i=1}^{N} \ln\left(\frac{p_i}{1-p_i}\right) \tag{22}$$

where p_i denotes the asymptotic p-value of the Dickey-Fuller-GLS statistic for the country i and where $\Phi\left(.\right)$ is the cumulative distribution function for a standard normal variable. Under the null hypothesis, all these statistics have a standard normal distribution. The results of the three unit root tests P_m, Z and L^* are reported on table 6. We also report the inverse normal test for the model with time trends. The conclusions are very similar to those drawn by Nelson and Plosser. At a 5% significant level, the non stationarity is not rejected for 11 out of 14 variables, whatever the choice of the statistic. For the velocity, the bond yield and unemployment rates the unit root hypothesis is not rejected only with the Z and L^* statistics. The results are identical in a model with time trends.

3.4 Pesaran unit root tests

Pesaran (2003) proposes a different approach to deal with the problem of cross-sectional dependencies. He considers a one-factor model with heterogeneous loading factors for residuals, as in Phillips and Sul (2003). However, instead of basing the unit root tests on deviations from the estimated common factors, he augments the standard Dickey Fuller or Augmented Dickey Fuller regressions with the cross section average of lagged levels and first-differences of the individual series. If residuals are not serially correlated, the regression used for the i^{th} country is defined as:

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + c_i \overline{y}_{t-1} + d_i \Delta \overline{y}_t + v_{it} \tag{23}$$

where $\overline{y}_{t-1} = (1/N) \sum_{i=1}^{N} y_{i,t-1}$ and $\Delta \overline{y}_t = (1/N) \sum_{i=1}^{N} \Delta y_{it}$. Let us denote $t_i\left(N,T\right)$ the t-statistic of the OLS estimate of ρ_i. The Pesaran's test is based on these individual cross-sectionally augmented ADF statistics, denoted CADF. A truncated version, denoted CADF*, is also considered to avoid undue influence of extreme outcomes that could arise for small T samples (see Pesaran, 2003, for more details). In both cases, the idea is to build a modified version of IPS t-bar test based on the average of individual CADF or CADF* statistics (respectively denoted CIPS and CIPS*, for cross-sectionally augmented IPS).

$$CIPS = \frac{1}{N}\sum_{i=1}^{N} t_i\left(N,T\right) \qquad CIPS^* = \frac{1}{N}\sum_{i=1}^{N} t_i^*\left(N,T\right) \tag{24}$$

where $t_i^*\left(N,T\right)$ denotes the truncated CADF statistic. All the individual CADF (or CADF*) statistics have similar asymptotic null distributions which do not depend on the factor loadings. But they are correlated due to the dependence on the common factor. Pesaran proposes simulated critical values of CIPS and CIPS* for various samples sizes. Finally, this approach readily extends to serially correlated residuals with the introduction of lagged terms $\Delta\overline{y}_{t-j}$ and $\Delta y_{i,t-j}$ for $j = 1,..,p$.

The results of the Pesaran CIPS and CIPS* tests are reported on table 7 for a lag order ranging from one to four. Whatever the choice of the lag length p, at a 5% significant level, the non stationarity hypothesis is not rejected for 6 variables: real GDP, real per-capita GDP, industrial production, employment, consumer prices and wages. The same conclusion is get for the real wages, except for a model with one lag. On the contrary, the non stationarity is robustly rejected for nominal GDP and bond yield. If we consider a 10% significant level, the non stationarity is also rejected at all lags for the unemployment rate and the money stocks. For velocity, common stock prices and GDP deflator, the conclusions depend on the choice of the lag order. Truncated statistics are exactly equivalent to the non truncated ones at all lags, except for real GDP, money stock and velocity. These results are sensibly different from those obtained with the standard IPS tests except for 5 variables (unemployment rate, real per capita GDP, employment, consumer prices and money stock). However, the standard IPS statistics do not present a systematic bias compared with the cross sectionally augmented ones. Indeed, the standard IPS statistic leads to reject the null for the real GDP whereas the augmented one leads to accept the unit root hypothesis for this variable. On the contrary, for the bond yield, the conclusions are reversed.

If we consider individual CADF statistics compared to ADF ones for the real GDP over the period 1963-2003 (see table 8), the conclusions are clearer. The CADF tests do not reject the non stationarity of the real GDP for 24 countries out of 25, whereas it was the case for only 17 countries with the ADF tests. Therefore, when we take into account the common factor in OECD real GDPs, via the introduction of cross sectionally augmented terms, the non stationarity of the real GDP seems to be largely accepted.

3.5 Chang nonlinear IV unit root tests

The second approach to model cross-sectional dependencies consists in imposing few or none restrictions on the covariance matrix of residuals (O'Connell, 1998; Taylor and Sarno, 1998; Chang, 2002 and 2004). Such an approach raises some important technical problems since the usual Wald type unit root tests based on standard estimators have limit distributions that are dependent in a very complicated way upon various nuisance parameters defining correlations across individual units. In this context, Chang (2002) proposes a solution that consists in using a nonlinear instrumental variable (IV thereafter). More precisely, she derives a nonlinear IV estimator of the autoregressive parameter in simple ADF model. She proves that the corresponding t-ratio (denoted Z_i) asymptotically converges to a standard normal distribution. Note that this asymptotic Gaussian result is very unusual and entirely due to the nonlinearity of the IV. Moreover, it can be shown that the asymptotic distributions of individual Z_i statistics are independent across cross-sectional units. So, panel unit root tests based on the cross-sectional average of individual independent statistics can be implemented. Chang proposes an average IV t-ratio statistic, denoted S_N and defined as:

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i \tag{25}$$

In a balanced panel, this statistic has a limit standard normal distribution. The instruments are generated by an Instrument Generating Function (IGF thereafter) which corresponds to a nonlinear function $F\left(y_{i,t-1}\right)$ of the lagged values $y_{i,t-1}$. It must be a regularly integrable function which satisfies $\int_{-\infty}^{\infty} xF\left(x\right)dx \neq 0$. This assumption can be interpreted as the fact that the nonlinear instrument $F(.)$ must be correlated with the regressor $y_{i,t-1}$. Chang provides several examples of regularly integrable IGFs. In our application, we consider three functions in order to assess the sensitivity of the results to the choice of the IGF. The first is $IGF_1(x) = x \exp(-c_i |x|)$ where $c_i \in \mathbb{R}$ is determined by $c_i = 3T^{-1/2}s^{-1}(\Delta y_{it})$ where $s^2(\Delta y_{it})$ is the sample standard error of Δy_{it}. The two others are $IGF_2(x) = \mathbb{I}(|x| < K)$ and $IGF_3(x) = \mathbb{I}(|x| < K) * x$, where K denotes a truncation parameter. The IV estimator constructed from the IGF_2 function is simply the trimmed OLS estimator based on observations in the interval $[-K, K]$.

Individual nonlinear IV t-ratio statistics for the real GDP over the period (1963-2003) are reported on table 8. All Z_i statistics have been computed in a model with individual effects and with IGF_1. The results are clearly in favor of the unit root. At a 5% significant level, the null unit root hypothesis is not rejected for 23 out of 25 countries. Then, this approach leads to clarify the conclusions of the panel unit root tests. Recall that, the pure time series ADF tests remain inconclusive and reject the null for eight countries. The results are even stronger with the panel tests reported in table 9. The S_N statistics based on the instrument generating functions IGF_2 and IGF_3 provide strong evidence in favor of the unit root. The null is not rejected for all the considered variables and the corresponding p-values are always very close to one. The results (not reported) are identical in a model with time trends. Chang (2002) founded the same type of conclusive results in her study of the PPP: her test always provides robust results against the null hypothesis. However, it is important to note that Im and Pesaran (2003) found very large size distortions with this test. Using a common factor model with a sizeable degree of cross section correlations, they show that the test suffers from severe size distortions, even when N is small relative to T.

4 Conclusion

The non-stationarity of the macroeconomic or financial variables remains an open question, especially since this concept has been deeply renewed in the context of nonlinear approaches or models with structural breaks. This debate is largely beyond scope of this paper. However, based on linear time series models without structural breaks, the results of Nelson and Plosser (1982) are generally considered as a reference for the main OECD aggregates. The issue is to know if an applied econometrician would obtain the same kind of general results with panel unit root tests. Our results show that the conclusions based on panel unit root tests are not clear-cut. We confirm the fact that panel methods must be very carefully used for testing unit roots in macroeconomic or financial panels.

What would Nelson and Plosser find had they used panel unit root tests? The table 10 summarizes the response. As we can observe, there is no global regularity, but our study highlights the importance of the specification of cross-sectional dependencies and heterogeneity. Our results raised three main points. Firstly, the unit root hypothesis is largely rejected when homogenous specifications (LL, 2002) are used to test the non-stationarity hypothesis. Secondly, the results based on heterogeneous specifications are more in favour of the non stationary hypothesis. However, under the cross-sectional independence assumption (IPS, 2003; Maddala and Wu, 1999; Choi, 2001), results are mitigated: the null is rejected for some macroeconomic variables generally considered as non-stationary such as the real GDP. Thirdly, when international cross-correlations are taken into account, conclusions depend on the specification of these cross-sectional dependencies. Two groups of tests can be distinguished. The first group tests are based on a dynamic factor model (Bai and Ng, 2004; Moon and Perron, 2004; Pesaran, 2003) or an error-component model (Choi, 2002). In this case, the non stationarity of common factors (international business cycles or growth trends) is genrally not rejected, but the results are less clear with respect to idiosyncratic components. The second group of tests is defined by opposition to these specifications based on common factor or time effects. In this case, it seems that the results are globally and clearly more in favor of the unit root assumption for most of main macroeconomic and financial indicators.

A Data appendix

As in Nelson and Plosser (1982), all series except the bond yields are transformed to natural logs. The data sources for the 14 series are:

Real GDP ($T = 41, N = 25$). Source: Economic Outlook, OECD. Code: GDPVD (gross domestic product, volume, at 2000 PPP, US$). Base 100 in 2000. The sample is balanced with 25 countries observed over the period 1963-2003. Excluded countries are Hungary, Korea, Czech Republic, Poland and the Slovak Republic.

Nominal GDP ($T = 41, N = 25$). Source: Economic Outlook, OECD. Code: GDPV (gross domestic product, volume, market prices). Base 100 in 2000. The sample is balanced with 25 countries observed over the period 1963-2003. Excluded countries are Hungary, Korea, Czech Republic, Poland and the Slovak Republic.

Real per capita GDP ($T = 36, N = 25$) Source: World Development Indicators, World Bank. Code: NY.GDP.PCAP.KD (gross domestic product per capita, constant 1995 US$). Base 100 in 1995. The sample is balanced with 25 countries observed over the period 1965-2000. Excluded countries are Turkey, Germany, Czech Republic, Poland and the Slovak Republic.

Industrial Production ($T = 43, N = 24$) : Source: International Financial Statistics, IMF, Washington. Code: line 61. Base 100 in 1995. The sample is balanced with 24 countries observed over the period 1960-2002. Excluded countries are Turkey, New Zealand, Czech Republic, Hungary, Poland and the Slovak Republic.

Employment ($T = 39, N = 23$). Source: Economic Outlook, OECD. Code: ET (total employment). The sample is balanced with 23 countries observed over the period 1965-2003. Excluded countries are: Luxembourg, Mexico, the Netherlands, the Czech Republic, Hungary, Poland and the Slovak Republic.

Unemployment rate ($T = 39, N = 23$). Source: Economic Outlook, OECD. Code: UN (unemployment rate). The sample is balanced with 23 countries observed over the period 1965-2003. Excluded countries are: Luxembourg, Mexico, the Netherlands, the Czech Republic, Hungary, Poland and the Slovak Republic.

GDP Deflator ($T = 41, N = 24$): Source: World Development Indicators, World Bank. Code: NY.GDP.DEFL.ZS. Base 100 in 1995. The sample is balanced with 24 countries observed over the period 1960-2003. Excluded countries are: Canada, Germany, Turkey, Czech Republic, Poland and the Slovak Republic

Consumer prices ($T = 52$, $N = 22$). Source: International Financial Statistics, IMF, Washington. Code: line 64. Base 100 in 2000. The sample is balanced with 22 countries observed over the period 1952-2003. Excluded countries are: Germany, Turkey, Mexico, Korea, Czech Republic, Hungary, Poland and the Slovak Republic.

Wages ($T = 33$, $N = 20$). Source: Economic Outlook, OECD. Code: WR (wage rate of the business sector). Base 100 in 2000. The sample is balanced with 20 countries observed over the period 1971-2003. Excluded countries are: Switzerland, Czech Republic, Hungary, Korea, Luxembourg , Mexico, Norway, Poland, Turkey and the Slovak Republic.

Real Wages ($T = 33$, $N = 20$). Source: Economic Outlook, OECD. Code: WSRE (real compensation rate of the business sector). Base 100 in 2000. The sample is balanced with 20 countries observed over the period 1971-2003. Excluded countries are: Switzerland, Czech Republic, Hungary, Korea, Luxembourg , Mexico, Norway, Poland, Turkey and the Slovak Republic.

Money Stock ($T = 30, N = 19$) Source: Economic Outlook, OECD. Code: MONEYS (money supply, broad definition, M2 or M3). Base 100 in 1995. The sample is balanced with 20 countries observed over the period 1969-1998. Excluded countries are: Luxembourg, Italy, France, Denmark, Turkey, Mexico, Korea, Czech Republic, Hungary, Poland and the Slovak Republic.

Velocity. ($T = 30, N = 18$) Source: Economic Outlook, OECD. Code: VLCTY (velocity of money). The sample is balanced with 18 countries observed over the period 1969-1998. Excluded countries are: Germany, Luxembourg, Italy, France, Denmark, Turkey, Mexico, Korea, Czech Republic, Hungary, Poland and the Slovak Republic.

Bond Yield: ($T = 47, N = 13$). Source: International Financial Statistics, IMF, Washington. Code: line 61. The sample is balanced with 13 countries observed over the period 1952-2002. Excluded countries are: Portugal, Sweden, Ireland, Austria, Finland, Greece, Iceland, Japan, Luxembourg, Spain, Turkey, Mexico, Korea, Czech Republic, Hungary, Poland and the Slovak Republic.

Common stock prices: ($T = 36, N = 11$). Source: Main Economic Indicators, OECD. Code: share prices. Base 100 in 2000. The sample is balanced with 11 countries observed over the period 1968-2003. Excluded countries are: Belgium, Czech Republic, Denmark, Finland, Greece, Hungary, Iceland, Italy, Korea, Mexico, Netherlands, Norway, Poland, Portugal, Spain, Turkey, Luxembourg, United Kingdom and the Slovak Republic.

B References

Bai J. and Ng S. (2002) Determining the Number of Factors in Approximate Factor Models, *Econometrica*, **70(1)**, 191-221.

Bai J. and Ng S. (2004) A PANIC Attack on Unit Roots and Cointegration, *Econometrica*, **72(4)**, 1127-1178.

Baltagi, B.H. and Kao, C. (2000) Nonstationary Panels, Cointegration in Panels and Dynamic Panels: a Survey, in *Advances in Econometrics*, 15, Elsevier Science, 7-51.

Baltagi, B.H., Bresson G., and Pirotte A. (2005) Panel Unit Root Tests and Spatial Dependence, *Working Paper, Department of Economics*, Texas A&M University.

Banerjee, A. (1999) Panel Data Unit Root and Cointegration: an Overview, *Oxford Bulletin of Economics and Statistics*, Special Issue, 607-629.

Banerjee, A., Marcellino M. and Osbat C. (2005) Testing for PPP: Should we use Panel Methods?, *Empirical Economics*, **30**, 77-91.

Breitung, J. and Pesaran, H. (2005) Unit Roots and Cointegration in Panels, *Discussion Paper 42 Economic Studies*, Deutsche Bundesbank.

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

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

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

Choi I. (2002) Combination Unit Root Tests for Cross-Sectionally Correlated Panels, *Mimeo, Hong Kong University of Science and Technology.*

Choi, I. (2004) Nonstationary Panels, in *Palgrave Handbooks of Econometrics.*

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

Fisher, R.A. (1932), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.

Gengenbach, C. Palm, F.C and Urbain, J.-P. (2004) Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling, *Working Paper, University of Maastricht.*

Gutierrez, L. (2003) On the Power of Panel Cointegration Tests: A Monte Carlo comparison, *Economics Letters*, **80**, 105-111.

Gutierrez, L. (2005) Panel Unit Roots Tests for Cross-Sectionally Correlated Panels: A Monte Carlo Comparison, *Oxford Bulletin of Economics and Statistics*, **68(4)**, 519-540

Hall, A., (1994) Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection, *Journal of Business and Economic Statistics*, **12**, 461-70.

Hurlin, C. and Mignon, V. (2004) Second Generation Panel Unit Root Tests, *Mimeo, University of Paris X.*

Im, K.S. and Pesaran, M.H. (2003) On the Panel Unit Root Tests Using Nonlinear Instrumental Variables, *Mimeo, University of Southern California.*

Im, K.S., Pesaran, M.H., and Shin, Y. (2003) Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, **115(1)**, 53-74.

Johansen, S. (1988) Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, **12**, 231-254.

Levin, A., Lin, C.F., and Chu., C.S.J. (2002) Unit Root Test in Panel Data: Asymptotic and Finite Sample Properties, *Journal of Econometrics*, **108**, 1-24.

Maddala, G.S. 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*, special issue, 631-652.

Moon, H. R. and Perron, B. (2004) Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics*, **122**, 81-126

Nelson, C. and Plosser, C. (1982) Trends and Random Walks in Macroeconomics Time Series: Some Evidence and Implications, *Journal of Monetary Economics*, **10**, 139-162.

Newey, W. and West, K. (1994) Automatic Lag Selection and Covariance Matrix Estimation, *Review of Economic Studies*, **61**, 631-653.

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

Osterholm P., (2004) Killing four unit root birds in the US economy with three panel unit root test stones, *Applied Economics Letters*, **11(4)**, 213 - 216.

Pesaran, H.M., (2003) A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, *Mimeo, University of Southern California.*

Phillips, P.C.B. and Sul, D. (2003) Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Econometrics Journal*, **6(1)**, 217-259.

Quah, D. (1994) Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Data, *Economics Letters*, **44**, 9-19.

Stock J.H and Watson M.W. (1988) Testing for Common Trends, *Journal of the American Statistical Association*, **83**, 1097-1107.

Summers, R., and Heston, A., (1991) The Penn World Tables (Mark 5): An Expanded set of International Comparisons, 1950-88, *Quarterly Journal of Economics*, **106(2)**, 327-368.

Table 1: Levin, Lin and Chu (2002) Unit Root Tests

	t_ρ^*	$\widehat{\rho}$	t_ρ^{*B}	t_ρ^{*C}	t_3^{*C}
Real GDP	-13.05^*	-0.023	-13.07^*	-13.05^*	-6.839^*
	(0.00)	$(0.39e^{-5})$	(0.00)	(0.00)	(0.00)
Nominal GDP	-12.66^*	-0.009	-12.68^*	-12.68^*	-3.037^*
	(0.00)	$(0.82e^{-6})$	(0.00)	(0.00)	(0.00)
Real per capita GDP	-6.739^*	-0.021	-6.977^*	-6.944^*	-4.753^*
	(0.00)	$(0.89e^{-5})$	(0.00)	(0.00)	(0.00)
Industrial production	-10.65^*	-0.029	-10.37^*	-10.15^*	-2.402^*
	(0.00)	$(8.26e^{-6})$	(0.00)	(0.00)	(0.00)
Employment	-4.442^*	-0.020	-4.714^*	-4.616^*	-1.943^*
	(0.00)	$(1.43e^{-5})$	(0.00)	(0.00)	(0.02)
Unemployment rate	-4.567^*	-0.063	-5.068^*	-4.848^*	1.357
	(0.00)	$(6.89e^{-5})$	(0.00)	(0.00)	(0.91)
GDP deflator	-9.311^*	-0.007	-9.333^*	-9.333^*	-7.019^*
	(0.00)	$(7.51e^{-7})$	(0.00)	(0.00)	(0.00)
Consumer prices	-6.214^*	-0.003	-6.226^*	-6.224^*	-10.34^*
	(0.00)	$(5.41e^{-7})$	(0.00)	(0.00)	(0.00)
Wages	-26.34^*	-0.044	-26.36^*	-26.36^*	-10.83^*
	(0.00)	$(0.36e^{-5})$	(0.00)	(0.00)	(0.00)
Real wages	-17.00^*	-0.084	-16.95^*	-16.94^*	-8.940^*
	(0.00)	$(0.26e^{-4})$	(0.00)	(0.00)	(0.00)
Money stock	-13.72^*	-0.023	-13.74^*	-13.75^*	1.138^*
	(0.00)	$(0.36e^{-5})$	(0.00)	(0.00)	(0.00)
Velocity	-1.465	-0.056	-1.772^*	-1.829^*	-1.900^*
	(0.07)	$(1.47e^{-4})$	(0.03)	(0.03)	(0.03)
Bond yield	-1.240	-0.096	-1.234	-1.069	4.387
	(0.10)	$(2.51e^{-4})$	(0.10)	(0.14)	(1.00)
Common stock prices	0.320	-0.018	0.024	0.050	0.828
	(0.62)	$(9.01e^{-5})$	(0.50)	(0.51)	(0.79)

Notes: t_ρ^* denotes the adjusted t-statistic computed with a Bartlett kernel function and a common lag truncation parameter given by $\overline{K} = 3.21T^{1/3}$ (Levin and Lin, 2002). Corresponding p-values are in parentheses. $\widehat{\rho}$ is the pooled least squares estimator. Corresponding standard errors are in parentheses. t_ρ^{*B} denotes the adjusted t-statistic computed with a Bartlett kernel function and individual bandwidth parameters (Newey and West, 1994). t_ρ^{*C} denotes the adjusted t-statistic computed with a Quadratic Spectral kernel function and individual bandwidth parameters. Finally, t_ρ^* denotes the adjusted t-statistic computed with a Bartlett kernel function and a common lag truncation parameter, for the model 3 with deterministic trends. Corresponding p-values are in parentheses. * Indicates significant at the 5% level.

Table 2: Im, Pesaran and Shin (2003) Unit Root Tests

	t_bar_{NT}	W	$Z_{t\,bar}$	$t\_bar^{DF}_{NT}$	$Z^{DF}_{t\,bar}$	$W^{(3)}_{t\,bar}$
Real GDP	−2.367	−4.799* (0.00)	−4.812* (0.00)	−2.746	−6.969* (0.00)	−2.691* (0.00)
Nominal GDP	−2.172	−3.689 (0.11)	−3.703* (0.00)	−3.991	−14.06* (0.00)	7.399 (1.00)
Real per capita GDP	−1.420	0.545 (0.70)	0.594 (0.72)	−1.692	−0.937 (0.17)	−2.519* (0.00)
Industrial production	−2.449	−5.173* (0.00)	−5.170* (0.00)	−2.557	−5.774* (0.00)	−0.409 (0.34)
Employment	−1.016	2.663 (0.99)	2.770 (0.99)	−0.542	5.356 (1.00)	0.252 (0.59)
Unemployment rate	−1.919	−2.252* (0.01)	−2.165* (0.01)	−1.718	−1.068 (0.14)	−0.676 (0.24)
GDP deflator	−1.591	−0.411 (0.340)	−0.383 (0.35)	−1.507	0.086 (0.53)	4.946 (1.00)
Consumer prices	−1.113	2.110 (0.982)	2.224 (0.98)	−0.306	6.564 (1.00)	1.640 (0.94)
Wages	−5.229	−18.19* (0.00)	−18.65* (0.00)	−8.824	−36.74* (0.00)	−2.562* (0.00)
Real wages	−3.152	−8.160* (0.00)	−8.191* (0.00)	−3.390	−9.389* (0.00)	−5.224 (0.87)
Money stock	−2.814	−6.259* (0.00)	−6.323* (0.00)	−3.990	−12.09* (0.00)	6.975 (1.00)
Velocity	−1.300	1.011 (0.84)	1.072 (0.85)	−12.30	1.410 (0.92)	−2.018* (0.02)
Bond yield	−1.728	−0.874 (0.19)	−0.834 (0.20)	−1.479	0.198 (0.57)	3.286 (0.99)
Common stock prices	−0.633	3.210 (0.99)	3.330 (0.99)	−0.683	3.144 (0.99)	−0.800 (0.21)

Notes: $t\_bar^{DF}_{NT}$ (respectively t_bar_{NT}) denotes the mean of Dickey Fuller (respectively Augmented Dickey Fuller) individual statistics. $Z^{DF}_{t\,bar}$ is the standardized $t\_bar^{DF}_{NT}$ statistic and associated p-values are in parentheses. $Z_{t\,bar}$ is the standardized t_bar_{NT} statistic based on the moments of the Dickey Fuller distribution. $W_{t\,bar}$ denotes the standardized t_bar_{NT} statistic based on simulated approximated moments (Im, Pesaran and Shin, 2003, table 3). $W^{(3)}_{t\,bar}$denotes the standardized statistic for the model with deterministic trends. The corresponding p-values are in parentheses. * Indicates significant at the 5% level.

Table 3: Maddala and Wu (1999) and Choi (2001) Unit Root Tests

	P_{MW}	Z_{MW}	$P_{MW}^{(3)}$	$Z_{MW}^{(3)}$
Real GDP	130.6* (0.00)	8.062* (0.00)	83.18* (0.00)	3.318* (0.00)
Nominal GDP	82.97* (0.00)	3.297* (0.00)	9.755 (1.00)	−4.024 (1.00)
Real per capita GDP	68.01 (0.06)	1.80* (0.04)	85.29* (0.00)	3.529* (0.00)
Industrial production	131.4* (0.00)	8.512* (0.00)	61.63 (0.08)	1.391 (0.08)
Employment	29.74 (0.96)	−1.689 (0.95)	56.59 (0.13)	1.104 (0.13)
Unemployment rate	68.34* (0.01)	2.329* (0.00)	71.99* (0.00)	2.710* (0.00)
GDP deflator	44.49 (0.61)	−0.358 (0.63)	17.52 (1.00)	−3.110 (0.99)
Consumer prices	18.02 (0.99)	−2.769 (0.99)	25.47 (0.97)	−1.803 (0.96)
Wages	272.6* (0.00)	26.01* (0.00)	91.43* (0.00)	5.750* (0.00)
Real wages	155.7* (0.00)	12.94* (0.00)	98.86* (0.00)	6.581* (0.00)
Money stock	111.3* (0.00)	8.418* (0.00)	16.77 (0.99)	−2.434 (0.99)
Velocity	39.40 (0.32)	0.400 (0.34)	60.12* (0.00)	2.843* (0.00)
Bond yields	25.94 (0.46)	−0.007 (0.50)	9.017 (0.99)	−2.355 (0.99)
Common stock prices	7.633 (0.99)	−2.165 (0.98)	26.84 (0.21)	0.731 (0.23)

Notes: P_{MW} denotes the Fisher's test statistic defined as $P_{MW} = -2\sum_{i=1}^{N}\log(p_i)$, where p_i are the p-values from ADF unit root tests for each cross-section $i = 1,..,N$. Under H_0, P_{MW} has a χ^2 distribution with $2N$ of freedom when T tends to infinity and N is fixed. Z_{MW} is the Choi (2001) standardized statistic used for large N samples: under H_0, Z_{MW} has a $N(0,1)$ distribution when T and N tend to infinity. $P_{MW}^{(3)}$ and $Z_{MW}^{(3)}$ denote the corresponding statistics for the model with time trends. * Indicates significant at the 5% level.

Table 4: Bai and Ng (2004) Unit Root Tests

	Criterion	$\widehat{r}$	Idiosyncratic Shocks $Z_{\widehat{e}}^{c}$	$P_{\widehat{e}}^{c}$	Common Factors $\widehat{F}$ $ADF_{\widehat{F}}^{c}$	Trends $\widehat{r}_1$ MQ_c	MQ_f
Real GDP	IC_2	1	3.461* (0.00)	84.61* (0.00)	−1.988 (0.29)	—	—
Nominal GDP	BIC_3	4	8.844* (0.00)	138.4* (0.00)	—	4	4
Real per capita GDP	IC_2	1	1.790* (0.04)	67.90* (0.05)	−1.212 (0.65)	—	—
Industrial production	BIC_3	3	1.072 (0.14)	58.50 (0.14)	—	3	3
Employment	IC_2	2	0.339 (0.36)	49.25 (0.34)	—	3	3
Unemployment rate	BIC_3	2	3.526* (0.00)	79.82* (0.00)	—	2	2
GDP deflator	BIC_3	4	7.803* (0.00)	124.4* (0.00)	—	4	4
Consumer prices	BIC_3	2	0.842 (0.19)	49.72 (0.19)	—	2	2
Wages	BIC_3	3	6.946* (0.00)	102.1* (0.00)	—	3	3
Real wages	BIC_3	3	−0.058 (0.52)	39.47 (0.49)	—	3	3
Money stock	BIC_3	4	5.705* (0.00)	87.73* (0.00)	—	4	4
Velocity	IC_2	2	−1.400 (0.91)	24.11 (0.93)	—	2	2
Bond yield	BIC_3	3	4.161* (0.00)	56.00* (0.00)	—	3	3
Common stock prices	BIC_3	4	−0.474 (0.68)	18.85 (0.65)	—	4	4

Notes: $\widehat{r}$ is the estimated number of common factors, based on IC_2 or BIC_3 criteria functions. For the idiosyncratic components $\widehat{e}_{it}$, only pooled unit root test statistics are reported. $P_{\widehat{e}}^{c}$ is a Fisher's type statistic based on p-values of the individual ADF tests. Under $H_0, P_{\widehat{e}}^{c}$ has a $\chi^2(2N)$ distribution when T tends to infinity and N is fixed. $Z_{\widehat{e}}^{c}$ is a standardized Choi's type statistic for large N samples: under H_0, $Z_{\widehat{e}}^{c}$ has a $N(0,1)$ distribution. p-values are in parentheses. For the idiosyncratic components $\widehat{F}_t$, two cases must be distinguished: if $\widehat{r}=1$, only standard ADF t-statistic, denoted $ADF_{\widehat{F}}^{c}$, is reported with its p-value. If $\widehat{r}>1$, the estimated number $\widehat{r}_1$ of independent stochastic trends in the common factors is reported (columns 6-7). The first estimated value $\widehat{r}_1$ is derived from the filtered test MQ_f and the second one is derived from the corrected test MQ_c. The level of these tests is 5%. * Indicates significant at the 5% level.

Table 5: Moon and Perron (2004) Unit Root Tests

	$\widehat{r}$	t_a^*	t_b^*	$\widehat{\rho}_{pool}^*$	t_a^{*B}	t_b^{*B}	$t_a^{\#}$
Real GDP	1	−13.18*	−6.228*	0.877	−13.46*	−6.369	−0.428
		(0.00)	(0.00)		(0.00)	(0.09)	(0.33)
Nominal GDP	4	−20.41*	−7.528*	0.812	−22.26*	−7.833*	0.082
		(0.00)	(0.00)		(0.00)	(0.00)	(0.53)
Real per capita GDP	1	−12.26*	−5.910*	0.870	−12.28*	−5.927*	−0.310
		(0.00)	(0.00)		(0.00)	(0.00)	(0.37)
Industrial production	3	−8.140*	−4.184*	0.883	−8.244*	−4.344*	−3.876*
		(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Employment	3	−11.09*	−4.975*	0.874	−11.13*	−5.021*	−1.988*
		(0.00)	(0.00)		(0.00)	(0.00)	(0.02)
Unemployment rate	2	−15.63*	−6.192*	0.822	−15.97*	−6.460*	−1.652*
		(0.00)	(0.00)		(0.00)	(0.00)	(0.05)
GDP deflator	4	−25.65*	−8.130*	0.749	−27.82*	−8.556*	−0.974
		(0.00)	(0.00)		(0.00)	(0.00)	(0.16)
Consumer prices	2	−22.71*	−8.367*	0.834	−23.10*	−8.480*	−0.261
		(0.00)	(0.00)		(0.00)	(0.00)	(0.39)
Wages	3	−15.38*	−6.080	0.782	−16.07*	−6.171*	−0.490
		(0.00)	(0.59)		(0.00)	(0.00)	(0.31)
Real wages	3	−10.11*	−6.363*	0.880	−10.07*	−6.352*	−3.945*
		(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Money stock	4	−11.52*	−4.813*	0.789	−12.04*	−4.872*	−0.469
		(0.00)	(0.00)		(0.00)	(0.00)	(0.31)
Velocity	3	−12.53*	−6.952*	0.834	−12.42*	−7.015*	−3.580*
		(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Bond yield	3	−14.28*	−6.059*	0.862	−14.47*	−6.100*	−0.674
		(0.00)	(0.00)		(0.00)	(0.00)	(0.24)
Common stock prices	4	−8.046*	−3.624*	0.868	−8.473*	−3.764*	−4.013*
		(0.00)	(0.00)		(0.00)	(0.00)	(0.00)

Notes: $\widehat{r}$ is the estimated number of common factors (based on IC_2 or BIC_3 criteria function). t_a^* and t_b^* are the unit root test statistics based on de-factored panel data (Moon and Perron, 2004). Corresponding p-values are in parentheses. $\widehat{\rho}_{pool}^*$ is the corrected pooled estimates of the autoregressive parameter. t_a^{*B} and t_b^{*B} are computed with a Bartlett kernel function in spite of a Quadratic Spectral kernel function. In both cases, bandwidth parameters are computed according to the Newey and West (1994) procedure. $t_a^{\#}$ denotes the unit root test statistic for the model with time trends. * Indicates significant at the 5% level.

Table 6: Choi (2002) Unit Root Test

	P_m	Z	L^*	$Z^{(3)}$
Real GDP	−3.137	3.553	3.245	2.651
	(0.99)	(0.99)	(0.99)	(0.99)
Nominal GDP	−0.013	−0.536	−0.610	5.264
	(0.50)	(0.29)	(0.27)	(1.00)
Real per capita GDP	−3.508	4.323	4.137	−0.678
	(0.99)	(1.00)	(1.00)	(0.24)
Industrial production	0.153	−0.697	−0.696	2.839
	(0.43)	(0.24)	(0.24)	(0.99)
Employment	−1.232	1.321	1.185	−0.633
	(0.89)	(0.90)	(0.88)	(0.26)
Unemployment rate	5.989*	−4.828	−5.100	−0.480
	(0.00)	(0.68)	(0.16)	(0.31)
GDP deflator	0.059	−1.143	−1.043	1.724
	(0.47)	(0.12)	(0.14)	(0.95)
Consumer prices	−0.452	−0.909	−0.814	−0.685
	(0.67)	(0.18)	(0.20)	(0.24)
Wages	−1.651	3.381	3.784	7.113
	(0.95)	(0.99)	(0.99)	(1.00)
Real wages	1.370	0.595	0.479	0.877
	(0.08)	(0.72)	(0.68)	(0.80)
Money stock	−1.712	2.601	2.913	4.275
	(0.95)	(0.99)	(0.99)	(1.00)
Velocity	3.106*	−0.975	−1.594	−0.364
	(0.00)	(0.16)	(0.05)	(0.35)
Bond yield	3.237*	−3.389	−3.260	3.011
	($0.6e^{-3}$)	(0.35)	(0.55)	(0.99)
Common stock prices	−1.667	1.417	1.281	−0.863
	(0.95)	(0.92)	(0.90)	(0.19)

Notes: the P_m test is a modified Fisher's inverse chi-square test (Choi, 2001). The Z test is an inverse normal test. The L^* test is a modified logit test. All statistics have a standard normal distribution under H_0 when T and N tend to infinity (Choi, 2002). The null hypothesis of non stationarity is rejected when P_m is greater than the upper tail of the standard normal distribution. For other tests, the null is rejected when the realizations is inferior to the lower tail of the standard normal distribution. $Z^{(3)}$ denotes the inverse normal test based on a model with time trends. Similar results are obtained with the two other statistics. Corresponding p-values are in parentheses. * Indicates significant at the 5% level.

Table 7: Pesaran (2003) Unit Root Tests

	—	$CIPS$				$CIPS^*$
Lag length p	p^*	1	2	3	4	1
Real GDP	1	−2.070 (0.12)	−1.793 (0.46)	−1.708 (0.48)	−1.609 (0.71)	−2.309 (0.15)
Nominal GDP	1	−2.261* (0.03)	−2.286* (0.02)	−2.573* (0.01)	−2.571* (0.01)	−2.261* (0.03)
Real per capita GDP	2	−1.848 (0.37)	−1.823 (0.41)	−1.911 (0.29)	−1.773 (0.48)	−1.848 (0.37)
Industrial production	1	−2.105 (0.10)	−1.804 (0.44)	−1.781 (0.48)	−1.793 (0.46)	−2.105 (0.10)
Employment	1	−1.748 (0.53)	−1.383 (0.92)	−1.613 (0.70)	−1.493 (0.83)	−1.748 (0.53)
Unemployment rate	2	−2.700* (0.01)	−2.276* (0.02)	−2.150 (0.07)	−2.299* (0.02)	−2.700* (0.01)
GDP deflator	2	−2.293* (0.02)	−2.024 (0.16)	−2.270* (0.02)	−2.194* (0.05)	−2.293* (0.02)
Consumer prices	2	−2.163 (0.07)	−1.867 (0.36)	−1.858 (0.38)	−1.751 (0.54)	−2.163 (0.07)
Wages	2	−2.173 (0.06)	−2.148 (0.07)	−1.918 (0.28)	−2.061 (0.13)	−2.173 (0.06)
Real wages	1	−2.325* (0.01)	−1.832 (0.40)	−1.384 (0.92)	−1.053 (0.99)	−2.325* (0.01)
Money stock	1	−3.588* (0.01)	−2.226* (0.04)	−2.468* (0.01)	−2.180 (0.06)	−2.656* (0.01)
Velocity	1	−2.980* (0.01)	−1.847 (0.37)	−2.380* (0.01)	−2.412* (0.01)	−2.663* (0.01)
Bond yield	1	−2.774* (0.00)	−2.851* (0.00)	−3.254* (0.00)	−2.922* (0.00)	−2.774* (0.00)
Common stock prices	1	−2.855* (0.01)	−2.568* (0.01)	−2.251 (0.08)	−2.001 (0.25)	−2.855* (0.01)

Notes: $CIPS$ is the mean of individual Cross sectionally augmented ADF statistics (CADF). $CIPS^*$ denotes the mean of truncated individual CADF statistics. The truncated statistics are reported only for one lag since they are always equal to not truncated ones for higher lag lengths. Corresponding p-values are in parentheses. p^* denotes the nearest integer of the mean of the individual lag lengths in ADF tests. * Indicates significant at the 5% level.

Table 8: Individual Unit Root Tests for Real GDP (1963-2003)

Country	ADF	$\text{ADF}^c_{\hat{e}}$	CADF	IV ADF	Country	ADF	$\text{ADF}^c_{\hat{e}}$	CADF	IV ADF
Australia	−1.69 (0.42)	−0.439 (0.51)	−0.59 (0.89)	3.41 (0.99)	Luxemb.	0.77 (0.99)	−0.676 (0.41)	−2.91 (0.10)	2.75 (0.99)
Austria	−3.89* (0.00)	−1.379 (0.15)	−1.50 (0.62)	2.97 (0.99)	Mexico	−3.71* (0.00)	−1.270 (0.18)	−2.12 (0.35)	3.29 (0.99)
Belgium	−3.94* (0.00)	−0.207 (0.60)	0.18 (0.98)	3.29 (0.99)	Netherla.	−1.70 (0.41)	−0.711 (0.40)	−2.11 (0.35)	−1.40 (0.07)
Canada	−2.92 (0.05)	−2.689* (0.00)	−2.70 (0.18)	3.24 (0.99)	New Zeal.	−0.48 (0.88)	−1.531 (0.11)	−3.11 (0.07)	3.21 (0.99)
Denmark	−3.78* (0.00)	−0.370 (0.54)	−3.25 (0.05)	4.30 (1.00)	Norway	−1.85 (0.34)	−0.457 (0.50)	−1.94 (0.42)	−1.33 (0.09)
Finland	−1.27 (0.62)	−2.025* (0.04)	−2.59 (0.18)	0.72 (0.76)	Portugal	−2.80 (0.06)	−2.306* (0.02)	−1.80 (0.49)	−1.69* (0.04)
France	−2.66 (0.08)	−0.395 (0.53)	−0.89 (0.82)	−1.69 (0.04)	Spain	−1.76 (0.38)	−1.596 (0.10)	−2.35 (0.26)	−1.25 (0.10)
Germany	−1.96 (0.302)	−2.219* (0.02)	−2.46 (0.22)	3.18 (0.99)	Sweden	−0.94 (0.76)	−1.092 (0.24)	−3.25 (0.05)	0.02 (0.51)
Greece	−4.43* (0.00)	−1.221 (0.20)	−2.23 (0.30)	4.77 (1.00)	Switzer.	−3.89* (0.00)	−0.158 (0.62)	−7.11* (0.01)	3.18 (0.99)
Iceland	−1.10 (0.70)	−2.222* (0.02)	−2.17 (0.32)	0.07 (0.53)	Turkey	−1.86 (0.34)	−0.407 (0.53)	−1.47 (0.63)	−1.37 (0.08)
Ireland	0.88 (0.99)	−1.358 (0.15)	−0.86 (0.83)	−1.11 (0.13)	U.K.	−0.45 (0.89)	−1.050 (0.25)	−1.52 (0.61)	3.23 (0.99)
Italy	−4.58* (0.00)	−0.792 (0.36)	−0.93 (0.82)	3.30 (0.99)	U.S.	−1.34 (0.60)	−0.841 (0.34)	−0.58 (0.90)	2.97 (0.99)
Japan	−7.74* (0.00)	−0.842 (0.34)	−1.38 (0.67)	3.71 (0.99)	—	—		—	—

Notes: Individual ADF, CADF (Pesaran, 2003) and IV non linear ADF (Chang, 2002) statistics are reported for each of the 25 countries of the panel for the real GDP (1963-2003). $\text{ADF}^c_{\hat{e}}$ denotes the ADF t-statistic for the idiosyncratic component issued from the Bai and Ng's decomposition with one common factor. Corresponding p-values are reported in parentheses. For the ADF tests, the lag length is optimally chosen using the general-to-specific (GS) procedure of Hall (1994) with a maximum lag length set to 4. Starting with some maximum number of lagged differences and if the last lagged difference is significant at 5%, choose that lag length, if not, reduce the order by one until the last included lag is significant or none are included. * Indicates significant at the 5% level.

Table 9: Chang (2002) Non Linear IV Unit Root Tests

S_N statistics	IGF_1	IGF_2	IGF_3
Real GDP	8.365 (1.00)	17.59 (1.00)	16.48 (1.00)
Nominal GDP	-6.541^* (0.00)	3.473 (0.99)	4.126 (1.00)
Real per capita GDP	12.88 (1.00)	16.86 (1.00)	14.52 (1.00)
Industrial production	14.49 (1.00)	17.98 (1.00)	14.86 (1.00)
Employment	5.297 (1.00)	9.842 (1.00)	7.671 (1.00)
Unemployment rate	−0.499 (0.30)	4.958 (1.00)	0.409 (0.65)
GDP deflator	-5.179^* (0.00)	4.090 (1.00)	3.594 (0.99)
Consumer prices	-4.077^* (0.00)	3.333 (0.99)	5.049 (1.00)
Wages	-1.719^* (0.04)	8.623 (1.00)	6.167 (1.00)
Real wages	10.84 (1.00)	11.84 (1.00)	8.631 (1.00)
Money stock	1.911 (0.97)	10.50 (1.00)	11.08 (1.00)
Velocity	2.696 (0.99)	4.713 (1.00)	3.312 (0.99)
Bond yield	−0.247 (0.40)	0.918 (0.82)	0.430 (0.66)
Common stock prices	4.937 (1.00)	5.894 (1.00)	5.381 (1.00)

Notes: The S_N statistic corresponds to the average of individual non-linear IV t-ratio statistics (Chang, 2002). It has a $N(0,1)$ distribution under H_0. Three Instrument Generating Functions (IGF) are considered: $IGF_1(x) = x\exp\left(-c_i\left|x\right|\right)$, where $c_i \in \mathbb{R}$ is determined by $c_i = 3\,T_i^{-1/2}s^{-1}\left(\Delta y_{it}\right)$ where $s^2\left(\Delta y_{it}\right)$ is the sample standard error of Δy_{it}. $IGF_2(x) = \mathbb{I}(|x| < K)$ and $IGF_3(x) = \mathbb{I}(|x| < K) * x$, where K is the second quantile of $y_{i,t}$. * Indicates significant at the 5% level.

Table 10: Summary of Conclusions

Tests	LLC	IPS	MW	CH	BN$_c$	BN$_i$	MP	CH2	P	IV
Real GDP	I(0)	I(0)	I(0)	I(0)	I(1)	I(0)	I(0)	I(1)	I(1)	I(1)
Nominal GDP	I(0)	I(1)	I(0)	I(0)	I(1)	I(0)	I(0)	I(1)	I(0)	I(1)
Real per capita GDP	I(0)	I(1)	I(1)	I(0)	I(1)	I(0)	I(0)	I(1)	I(1)	I(1)
Industrial production	I(0)	I(0)	I(0)	I(0)	I(1)	I(1)	I(0)	I(1)	I(1)	I(1)
Employment	I(0)	I(1)	I(1)	I(1)	I(1)	I(1)	I(0)	I(1)	I(1)	I(1)
Unemployment rate	I(0)	I(0)	I(0)	I(0)	I(1)	I(0)	I(0)	I(0)	I(0)	I(1)
GDP deflator	I(0)	I(1)	I(1)	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)	I(1)
Consumer prices	I(0)	I(1)	I(1)	I(1)	I(1)	I(1)	I(0)	I(1)	I(1)	I(1)
Wages	I(0)	I(0)	I(0)	I(0)	I(1)	I(0)	I(1)	I(1)	I(1)	I(1)
Real wages	I(0)	I(0)	I(0)	I(0)	I(1)	I(1)	I(0)	I(1)	I(1)	I(1)
Money stock	I(0)	I(0)	I(0)	I(0)	I(1)	I(0)	I(0)	I(1)	I(0)	I(1)
Velocity	I(0)	I(1)	I(1)	I(1)	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)
Bond yield	I(1)	I(1)	I(1)	I(1)	I(1)	I(0)	I(0)	I(0)	I(0)	I(1)
Common stock prices	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(0)	I(1)	I(0)	I(1)

Notes: All results are obtained in a model with fixed effects at the 5% level. LLC denotes the Levin, Lin and Chu (2002) test (statistic t^*_ρ), IPS denotes Im, Pesaran and Shin (2003) test (statistic W), MW denotes the Maddala and Wu (1999) test (statistic P_{MW}), CH denotes the Choi (2001) test (statistic Z_{MW}), BN_c denotes the Bai and Ng (2004) test for common factors (ADF^c_F or MQ statistics), BN_i denotes the Bai and Ng (2004) test for idiosyncratic shocks (statistic P^c_e), MP denotes the Moon and Perron (2004) test (statistic t^*_a), CH2 denotes the Choi (2002) test (statistic P_m), P denotes the Pesaran (2003) test (statistic $CIPS$ with $p = 2$) and IV denotes the Chang (2002) test (statistic IGF_2).