Testing for unit root in nonlinear heterogeneous panels☆

Nuri Ucar a, b, *, Tolga Omay c

a PGlobal Advisory Services, Raza sok. 37/4, GOP, Ankara, Turkey
b Department of International Trade Management, Cankaya University, Ogretmenler Cad. No:14, 06530 Balgat, Ankara, Turkey
c Department of Economics, Cankaya University, Ogretmenler Cad. No:14, 06530 Balgat, Ankara, Turkey

A R T I C L E   I N F O

Article history:
Received 10 June 2007
Received in revised form 3 March 2009
Accepted 12 March 2009
Available online 28 March 2009

Keywords:
Nonlinear
Panel unit root
Sieve bootstrap

JEL classification:
C12
C22
O47

A B S T R A C T

We develop unit root tests for nonlinear heterogeneous panels where the alternative hypothesis is an exponential smooth transition (ESTAR) model, and provide their small sample properties. We apply our tests for investigating the income convergence hypothesis in the OECD sample.

© 2009 Published by Elsevier B.V.

1. Introduction

By using the nonlinear time series framework Kapetanios et al. (2003), hereafter KSS, and the panel unit root testing framework of Im et al. (2003), hereafter IPS, this paper proposes unit root tests for nonlinear heterogeneous panels.

Section 2 of this paper develops the proposed test statistics and represents their critical values. Section 3 provides the small sample performance of our normalized test in comparison with the power of the IPS test. Section 4 presents the application of our aforementioned tests in the empirical investigation of income convergence for the OECD sample.

2. The model and testing framework

Let \( y_{it} \) be panel exponential smooth transition autoregressive process of order one (PESTAR(1)) on the time domain \( t = 1, 2, \ldots, T \) for the cross section units \( i = 1, 2, \ldots, N \). Consider \( y_{it} \) follows the data generating process (DGP) with fixed effect (heterogeneous intercept) parameter \( \alpha_i \):

\[
\Delta y_{it} = \alpha_i + \gamma y_{it-1} \left( 1 - \exp \left( -\frac{\theta y_{it}^2}{\phi_i} \right) \right) + \epsilon_{it}
\]

where \( \phi_i \geq 1 \) is the delay parameter and \( \theta_i > 0 \) implies the speed of mean reversion for all \( i \).

By taking the previous literature into consideration (e.g. Balke and Fomby, 1997; Michael et al., 1997), we set \( \phi_i = 0 \) for all \( i \) (i.e., \( y_{it} \) has a unit root process in the middle regime) and \( d = 1 \), which gives specific PESTAR(1) model:

\[
\Delta y_{it} = \alpha_i + \gamma y_{it-1} \left( 1 - \exp \left( -\theta y_{it}^2 \right) \right) + \epsilon_{it}
\]

Nonlinear panel data unit root test based on regression (2) is simply to test the null hypothesis \( \theta_i = 1 \) for all \( i \) against \( \theta_i > 0 \) for some \( i \) under the alternative. However, direct testing of the \( \theta_i = 0 \) is somewhat problematic because \( \gamma_i \) is not identified under the null. This problem is achieved by applying first-order Taylor series approximation to the PESTAR(1) model around \( \theta_i = 0 \) for all \( i \). Hence, we obtain the auxiliary regression

\[
\Delta y_{it} = \alpha_i + \hat{\delta}_i y_{it-1}^2 + \epsilon_{it}
\]

where \( \hat{\delta}_i = \theta_i \gamma_i \).
We establish the hypotheses for unit root testing based on regression (3) as follows:

\[ H_0 : \delta_i = 0 \quad \text{for all } i \quad \text{(linear nonstationarity)} \]
\[ H_1 : \delta_i < 0 \quad \text{for some } i \quad \text{(nonlinear stationarity)} \]

(4)

We propose panel unit root tests computed through taking the average of individual KSS statistics. The KSS statistic for the ith individual is simply \( t_i \)-ratio of \( \delta_i \) in regression (3) defined by

\[ t_{i,NL} = \frac{\Delta y_i M_i}{\delta_{i,NL} (y_{i-1} M_i y_{i-1})^{0.5}} \]

(5)

where \( \Delta^2 t_{i,NL} \) is the consistent estimator such that \( \Delta^2 t_{i,NL} = \Delta y_i M_i \Delta y_i / (T - 1) \). \( M_i = I - T_j (T_j T_j) ^{-1} T_j \). Notice here that \( \Delta y_i = (\Delta y_{i,1}, \Delta y_{i,2}, ..., \Delta y_{i,T}) y_{i-1}^3 = (y_{i,0}^3, y_{i,1}^3, ..., y_{i,T-1}^3) \) and \( T_j = (1, 1, ..., 1) \). Furthermore, for a fixed \( T \), we have

\[ t_{i,NL} = \frac{1}{N} \sum_{i=1}^{N} t_{i,NL} \]

(6)

which is invariant average statistic when \( t_{i,NL} \) holds the property for each \( i \) in the following Lemma.

**Lemma 1.** The test statistic \( t_{i,NL} \) is invariant with respect to initial observations \( y_{i,0} \), heterogeneous moments \( \sigma_i^2 \) and \( \sigma_i^4 \) if \( y_{i,0} = 0 \) for all \( i = 1, 2, ..., N \).

**Proof.** See Mathematical appendix. \( \square \)

In addition, finitely bounded first and second moments of \( t_{i,NL} \) have to exist. These moments are produced via stochastic simulations and presented in Table 1.

However, stochastic simulations, although compatible with the existing moments, are not sufficient to confirm the existence of these moments. One possible solution is to use truncation of \( t_{i,NL} \) statistics, say \( \tilde{t}_{i,NL} \), as proposed in Pesaran (2007), defined as follows

\[ \tilde{t}_{i,NL} = \begin{cases} t_{i,NL} & \text{if } C_1 < t_{i,NL} < C_2 \\ C_1 & \text{if } t_{i,NL} < C_1 \\ C_2 & \text{if } t_{i,NL} > C_2 \end{cases} \]

(7)

where \( C_1 \) and \( C_2 \) are positive constants which are calculated from \( C_1 = -E(t_{i,NL}) - \Phi^{-1}(\epsilon/2) / \sqrt{Var(t_{i,NL})} \) and \( C_2 = E(t_{i,NL}) + \Phi^{-1}(1 - \epsilon/2) / \sqrt{Var(t_{i,NL})} \). Using simulated values of \( E(t_{i,NL}) = -1.677 \) and \( \sqrt{Var(t_{i,NL})} = 0.721 \) from Table 1 for the model with intercept term and setting \( \epsilon = 10^{-6} \), we have \( C_1 = 5.8308 \) and \( C_2 = 2.4766 \). The critical values for \( t_{i,NL} \) and its truncated version are generated by Monte Carlo simulations with 50,000 replications and displayed in Table 2.\(^1\)

Furthermore, when the invariance property (given in Lemma 1) and the existence of moments (by truncating \( t_{i,NL} \) distribution) are satisfied, the usual normalization of \( \tilde{t}_{i,NL} \) statistic yields as in the following Lemma.

**Lemma 2.** Individual statistics \( t_{i,NL} \) are iid random variables with finite means and variances; then, average statistics \( \tilde{t}_{i,NL} \), as defined in Eq. (6) have the limiting standard normal distribution as \( N \to \infty \) such that

\[ Z_{NL} = \frac{\sqrt{N} (\tilde{t}_{i,NL} - E(t_{i,NL}))}{\sqrt{Var(t_{i,NL})}} \]

(8)

**Proof.** It follows directly from Lindberg-Levy CLT. \( \square \)

We produce critical values of \( Z_{NL} \) statistic as well as its truncated version because they may be different from the fractiles of the standard normal distribution, particularly for small \( N \) observations, to which they converge as \( N \) goes to infinity (Table 3).\(^2\)

### Table 1: Moments of \( t_{i,NL} \) statistic.

<table>
<thead>
<tr>
<th>( T )</th>
<th>( E(t_{i,NL}) )</th>
<th>( Var(t_{i,NL}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-1.666</td>
<td>2.695</td>
</tr>
<tr>
<td>10</td>
<td>-1.620</td>
<td>0.823</td>
</tr>
<tr>
<td>15</td>
<td>-1.602</td>
<td>0.760</td>
</tr>
<tr>
<td>20</td>
<td>-1.602</td>
<td>0.740</td>
</tr>
<tr>
<td>25</td>
<td>-1.604</td>
<td>0.735</td>
</tr>
<tr>
<td>30</td>
<td>-1.605</td>
<td>0.735</td>
</tr>
<tr>
<td>40</td>
<td>-1.616</td>
<td>0.735</td>
</tr>
<tr>
<td>50</td>
<td>-1.626</td>
<td>0.727</td>
</tr>
<tr>
<td>100</td>
<td>-1.652</td>
<td>0.727</td>
</tr>
<tr>
<td>500</td>
<td>-1.675</td>
<td>0.725</td>
</tr>
<tr>
<td>1000</td>
<td>-1.677</td>
<td>0.721</td>
</tr>
<tr>
<td>100,000</td>
<td>-1.677</td>
<td>0.716</td>
</tr>
</tbody>
</table>

\(^1\) The critical values for both intercept and trend are available from the authors.

\(^2\) The critical values for both intercept and trend are available upon request.
Table 3

Exact critical values of $Z_{\text{NL}}$ statistic.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.48, 3.38</td>
<td>3.26</td>
</tr>
<tr>
<td></td>
<td>(-2.73, -2.64)</td>
<td>(-2.56, -2.60)</td>
</tr>
<tr>
<td>10</td>
<td>-2.56, -2.49</td>
<td>-2.48</td>
</tr>
<tr>
<td>20</td>
<td>-2.52, -2.52</td>
<td>-2.71</td>
</tr>
<tr>
<td>50</td>
<td>-2.19, -2.21</td>
<td>-2.32</td>
</tr>
<tr>
<td>70</td>
<td>-2.17, -2.20</td>
<td>-2.21</td>
</tr>
<tr>
<td>100</td>
<td>-2.15, -2.22</td>
<td>-2.20</td>
</tr>
</tbody>
</table>

Table 4

The size comparisons of alternative tests.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0.051</td>
<td>0.052</td>
</tr>
<tr>
<td>10</td>
<td>0.053</td>
<td>0.056</td>
</tr>
<tr>
<td>20</td>
<td>0.050</td>
<td>0.047</td>
</tr>
<tr>
<td>50</td>
<td>0.054</td>
<td>0.053</td>
</tr>
<tr>
<td>70</td>
<td>0.050</td>
<td>0.055</td>
</tr>
<tr>
<td>100</td>
<td>0.052</td>
<td>0.047</td>
</tr>
</tbody>
</table>

For the case $\gamma = -0.1$, both of the tests suffer from low power, but their power increases monotonically with $N$ and $T$.

4. Empirical application: testing income convergence hypothesis

In this section we reexamine the income convergence hypothesis in the OECD by using our $T_{\text{NL}}$ and augmented $Z_{\text{NL}}$, say $Z_{\text{NLU}}$, statistics and compare our results with the IPS tests $t_{\text{bar}}$ and $W_{\text{bar}}$. The tests proposed in the previous parts were based on the assumption of independence over cross-section units. However, this may not be possible for some observed data. To overcome the cross section dependency problem, we have implemented sieve bootstrap approach for the application on the Summers–Heston real GDP data set for the period 1953–2004. Bootstrap algorithm is below.3

(i) We consider the following OLS regression for each country which allows for different lag orders $p_i$:

$$
\Delta y_{it} = d_i + \delta y_{it-1} + \sum_{j=1}^{p_i} \beta_{j,i} \Delta y_{it-j} + \epsilon_{it}
$$

where deterministic component $d_i$ is considered for $\alpha_i$ and $\alpha_i + \beta_i$. Notice here that lag orders are selected via Schwartz criterion by starting $p_i = 6$ and applying top down strategy.

(ii) Following Basawa et al. (1991), the unit root null is imposed to generate bootstrap samples of residuals. Thus, we estimate the errors as:

$$
\hat{\epsilon}_{it} = \Delta y_{it} - \hat{d}_i - \sum_{j=1}^{p_i} \hat{\beta}_{j,i} \Delta y_{it-j}
$$

(iii) Stine (1987) suggests that residuals have to be centered with $\hat{e}_{i} = \hat{e}_{i} - (T-p+2)^{-1} \sum_{t=p+2}^{T} \hat{e}_{t}$

where $\hat{e}_{i} \sim \tilde{e}_{i}$ and $p = \max(p_i)$. Furthermore, we develop the $N \times T$ $[\tilde{e}_{it}]$ matrix from these residuals. We select randomly a full column with replacement from this matrix at a time to preserve the cross covariance structure of the errors. We denote

3 Separate programming code is developed in Matlab 7.5.
Note: p-values are in the parenthesis.

the bootstrap residuals as \( \hat{\epsilon}_t \), where \( t = 1, \ldots, T \) and \( T = 2T \) in our application.

(iv) We produce bootstrap \( \Delta y^*_t \) recursively from

\[
\Delta y^*_t = \hat{d}_t + \sum_{j=1}^{p} \hat{\beta}_{ij} \Delta y^*_t - j + \hat{\epsilon}^*_t
\]

where \( \hat{d}_t \) and \( \hat{\beta}_{ij} \) are the estimations from step (ii) and \( \Delta y^*_t = \hat{y}_t - \hat{\epsilon}_t = 0 \) for \( p_t = 1, \ldots, 6, \).

(v) We generate nonstationary bootstrap samples from the partial sums:

\[
y^*_t = \sum_{j=1}^{T} \Delta y^*_t
\]

The bootstrap statistics \( \bar{\epsilon}^*_{NL} \) and \( \bar{\zeta}^*_{ANL} \) are computed for each bootstrap replication by running the regression

\[
\Delta y^*_t = \hat{d}_t + \gamma_t \Delta y^*_{t-1} + \sum_{j=1}^{p} \hat{\beta}_{ij} \Delta y^*_t - j + \hat{\epsilon}^*_t
\]

where noting that the last \( T \) observations of \( y^*_t \) and \( \Delta y^*_t \) are used in this regression. The bootstrap empirical distribution of \( \bar{\epsilon}^*_{NL} \) and \( \bar{\zeta}^*_{ANL} \) Statistics, generated by employing 2000 replications, are used to have their p-values. The same procedure is also applied for the IPS statistics \( \bar{\epsilon}^*_{bar} \) and \( \bar{\zeta}^*_{bar} \). The results for 5% significance level are reported in Table 6.

Our tests provide evidence for income convergence when only an intercept is included as well as when both an intercept and a time trend are considered in the regression. On the other hand, the result for \( \bar{\epsilon}^*_{bar} \) and \( \bar{\zeta}^*_{bar} \) obtained from the linear version of regression (14) with intercept and trend fails to reject the null hypothesis of no stochastic income convergence whereas the null is significantly rejected when the same regression equation contains only intercept term.

Mathematical appendix

Proof of Lemma 1. We first note that under \( \delta_t = 0 \), \( y_{t-1} \) can be written as follows:

\[
y_{t-1} = y_0 t_T + u_{t-1}
\]

where \( y_0 \) is a fixed initial value and \( u_{t-1} = (u_0, u_1, \ldots, u_{T-1}) \) with

\[
u_{t-1} = \sum_{j=1}^{T} e_{ij}
\]

Secondly, let’s define the following \( T \times T \) matrix,

\[
A_T = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 \\
1 & 1 & 1 & \cdots & 0 \\
1 & 1 & 1 & \cdots & 0
\end{bmatrix}
\]

and using this matrix, we have \( u_{t-1} = A_T e_t \) where \( e_t = (e_{t1}, e_{t2}, \ldots, e_{tT}) \).

The cubic form of Eq. (A1) can be expressed:

\[
y_{t-1}^3 = y_0^3 t_T + 3 y_0^2 t_T \circ (A_T e_t) + 3 y_0 t_T \circ (A_T e_t)^2 + (A_T e_t)^3
\]

where \( A_T e_t \) is \( T \times 1 \) vector and \( \circ \) denotes for Hadamard product. Furthermore, multiplying both sides with \( M \), yields

\[
M_t y_{t-1}^3 = 3 y_0^2 t_T \circ (M_T A_T e_t) + 3 y_0 t_T \circ [(A_T e_t)^2 - \sigma^2 t] + M_T (A_T e_t)^3
\]

where \( M_T (A_T e_t)^2 = (A_T e_t)^2 - \sigma^2 t \) and \( t = (1, 2, \ldots, T) \). Notice here that taking expectation of the process (A2) on the time domain is

\[
\mathbb{E} \{ y_{t-1}^3 \} = y_0^3 t_T + 3 y_0 \sigma^2 t_T \circ t + \mathbb{E} (A_T e_t)^3 = 0
\]

due to symmetric normal distribution assumption for \( e_t \).

Now setting \( y_{t0} = 0 \) implies

\[
M_t y_{t-1}^3 = M_T (A_T e_t)^3
\]

Therefore, imposing \( y_{t0} = 0 \) removes the nuisance parameter \( \sigma^2 t \) in (A3). Both \( \Delta y_t = e_t \) and (A4) are plugged into the test-statistic (5), we have

\[
t^0_{NL} = \frac{\sqrt{T - \mathbb{E}[y_t^3]}}{\mathbb{E}[y_t^3]^{1/2} \mathbb{E}[(e_t A_T)^3 M_T (A_T e_t)^3]^{1/2}}
\]

which is invariant with respect to \( \sigma^2 t \).

Now let’s define, \( v_t = \frac{d_t}{\sqrt{T}} \sim N(0, I_T) \) and putting this into Eq. (A5) gives directly

\[
t^0_{NL} = \frac{\sqrt{T - \mathbb{E}[y_t^3]}}{\mathbb{E}[v_t^3 M_t (A_T v_t)^3]^{1/2} \mathbb{E}[(v_t A_T)^3 M_t (A_T v_t)^3]^{1/2}}
\]

so that \( \sigma^2 t \) is cancelled out in our test-statistic. □ QED

References


Table 6

<table>
<thead>
<tr>
<th>Test</th>
<th>( t^0_{NL} )</th>
<th>( Z^0_{ANL} )</th>
<th>( c^0_{bar} )</th>
<th>( W^0_{bar} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only intercept</td>
<td>-1.612</td>
<td>0.105</td>
<td>-1.408</td>
<td>0.749</td>
</tr>
<tr>
<td>(0.0035)</td>
<td>(0.0035)</td>
<td>(0.0085)</td>
<td>(0.0085)</td>
<td></td>
</tr>
<tr>
<td>Intercept and trend</td>
<td>-2.3744</td>
<td>-1.723</td>
<td>-1.0244</td>
<td>8.0263</td>
</tr>
<tr>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.959)</td>
<td>(0.959)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) In computing \( Z^0_{ANL} \) and \( W^0_{bar} \) moments are regenerated by taking \( T = 52 \) observations and various lag orders \( p_t \) into consideration.