Economics Letters 104 (2009) 5–8

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Testing for unit root in nonlinear heterogeneous panels[☆]

Nuri Ucar [a,b,*], Tolga Omay [c]

[a] PGlobal Advisory Services, Koza 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

ARTICLE INFO

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

ABSTRACT

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.

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

$$\Delta y_{i,t} = \alpha_i + \phi_i y_{i,t-1} + \gamma_i y_{i,t-1}\left[1 - exp\left(-\theta_i y_{i,t-d}^2\right)\right] + \varepsilon_{i,t} \tag{1}$$

where $d \geq 1$ 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; Micheal et al., 1997), we set $\phi_i = 0$ for all i (i.e, $y_{i,t}$ has a unit root process in the middle regime) and $d=1$, which gives specific PESTAR(1) model :

$$\Delta y_{i,t} = \alpha_i + \gamma_i y_{i,t-1}\left[1 - exp\left(-\theta_i y_{i,t-1}^2\right)\right] + \varepsilon_{i,t} \tag{2}$$

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 γ_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

$$\Delta y_{i,t} = \alpha_i + \delta_i y_{i,t-1}^3 + \varepsilon_{i,t} \tag{3}$$

where $\delta_i = \theta_i \gamma_i$.

☆ Our thanks go to anonymous referee, Furkan Emirmahmutoglu, Huseyin Guler and Peren Arin for helpful comments.

* Corresponding author. Department of International Trade Management, Cankaya University, Ogretmenler Cad. No:14, 06530 Balgat, Ankara, Turkey. Tel.: +90 3122844500/153; fax: +90 3122865044.
E-mail address: nucar@pglobal.com.tr (N. Ucar).

0165-1765/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.econlet.2009.03.018

We establish the hypotheses for unit root testing based on regression (3) as follows:

$$\begin{aligned} \mathcal{H}_0 &: \delta_i = 0, \text{ for all } i, \quad \text{(linear nonstationarity)} \\ \mathcal{H}_1 &: \delta_i < 0, \text{ for some } i, \quad \text{(nonlinear stationarity)} \end{aligned} \tag{4}$$

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-ratio of δ_i in regression (3) defined by

$$t_{i,NL} = \frac{\Delta \mathbf{y}_i' \mathbf{M}_\tau \mathbf{y}_{i,-1}^3}{\hat{\sigma}_{i,NL}\left(\mathbf{y}_{i,-1}' \mathbf{M}_\tau \mathbf{y}_{i,-1}\right)^{3/2}} \tag{5}$$

where $\hat{\sigma}^2_{i,NL}$ is the consistent estimator such that $\hat{\sigma}^2_{i,NL} = \Delta \mathbf{y}_i' \mathbf{M}_\tau \Delta \mathbf{y}_i / (T-1), \mathbf{M}_\tau = I_T - \tau_T(\tau_T'\tau_T)^{-1}\tau_T'$. Notice here that $\Delta \mathbf{y}_i = (\Delta y_{i,1}, \Delta y_{i,2}, \dots, \Delta y_{i,T})'$, $\mathbf{y}_{i,-1}^3 = \left(y_{i,0}^3, y_{i,1}^3, \dots, y_{i,T-1}^3\right)$ and $\tau_T = (1, 1, \dots, 1)'$

Furthermore, for a fixed T, we have

$$\bar{t}_{NL} = \frac{1}{N}\sum_{i=1}^{N} t_{i,NL} \tag{6}$$

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 σ_i^2 and σ_i^4 if $y_{10} = 0$ for all $i = 1,2,\dots,N$.*

Proof. See Mathematical appendix. □

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

$$\tilde{t}_{i,NL} = \begin{cases} t_{i,NL} & \text{if} \quad -C_1 < t_{i,NL} < C_2 \\ -C_1 & \text{if} \quad t_{i,NL} \le C_1 \\ C_2 & \text{if} \quad t_{i,NL} \ge C_2 \end{cases} \tag{7}$$

where C_1 and C_2 are positive constants which are calculated from $C_1 = -E(t_{i,NL}) - \Phi^{-1}(\epsilon/2)\sqrt{\mathrm{Var}(t_{i,NL})}$ and $C_2 = E(t_{i,NL}) + \Phi^{-1}(1-\epsilon/2)\sqrt{\mathrm{Var}(t_{i,NL})}$. Using simulated values of $E(t_{i,NL}) = -1.677$ and $\mathrm{Var}(t_{i,NL}) = 0.721$ 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 $\bar{t}_{NL}$ 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 $\bar{t}_{NL}$ 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 $\bar{t}_{i,NL}$, as defined in Eq. (6) have the limiting standard normal distribution as $N \to \infty$ such that*

$$\bar{Z}_{NL} = \frac{\sqrt{N}\left(\bar{t}_{NL} - E\left(t_{i,NL}\right)\right)}{\sqrt{\mathrm{Var}\left(t_{i,NL}\right)}} \xrightarrow{d} N(0,1) \tag{8}$$

Proof. It follows directly from Lindberg–Levy CLT. □

We produce critical values of $\bar{Z}_{NL}$ 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]

3. Finite sample performances

We use Monte Carlo simulations to illustrate the finite sample properties of our normalized $\bar{Z}_{NL}$ test in comparison with the IPS linear normalized unit root test, $\bar{Z}_{tbar}$.

The data are generated by running the nonlinear panel regression (2) where innovations are assumed to be drawn from a heterogeneous distribution such that $\varepsilon_{i,t} \sim N(0, \sigma_i^2)$ and $\sigma_i^2 \sim U(0.5, 2)$. We compute the empirical sizes of both linear and nonlinear normalized average test statistics under the null of $y_{i,t} = y_{i,t-1} + \varepsilon_{i,t}$ with

Table 1
Moments of $t_{i,NL}$ statistic.

T	$E(t_{i,NL})$	$\mathrm{Var}(t_{i,NL})$
5	−1.866	2.695
10	−1.620	0.823
15	−1.602	0.760
20	−1.602	0.740
25	−1.604	0.737
30	−1.605	0.735
40	−1.616	0.735
50	−1.626	0.727
100	−1.652	0.727
500	−1.675	0.725
1000	−1.677	0.721
100,000	−1.677	0.716

Table 2
Exact critical values of $\bar{t}_{NL}$ statistic.

Only intercept							
$T \downarrow \backslash N \rightarrow$	5	10	15	20	25	50	100
1%							
5	−4.35	−3.61	−3.27	−3.04	−2.89	−2.61	−2.34
	(−3.26)	(−2.80)	(−2.58)	(−2.48)	(−2.40)	(−2.21)	(−2.08)
10	−2.65	−2.33	−2.19	−2.12	−2.06	−1.93	−1.83
30	−2.44	−2.21	−2.10	−2.04	−2.00	−1.90	−1.80
50	−2.45	−2.22	−2.11	−2.05	−2.00	−1.92	−1.82
70	−2.47	−2.23	−2.12	−2.07	−2.02	−1.91	−1.83
100	−2.47	−2.24	−2.14	−2.07	−2.03	−1.91	−1.84
5%							
5	−3.05	−2.74	−2.59	−2.50	−2.43	−2.30	−2.15
	(−2.76)	(−2.47)	(−2.33)	(−2.26)	(−2.21)	(−2.08)	(−2.00)
10	−2.31	−2.10	−2.02	−1.96	−1.92	−1.82	−1.76
30	−2.20	−2.04	−1.96	−1.91	−1.88	−1.80	−1.74
50	−2.22	−2.05	−1.97	−1.93	−1.90	−1.83	−1.76
70	−2.23	−2.06	−1.99	−1.94	−1.91	−1.83	−1.78
100	−2.24	−2.07	−2.00	−1.95	−1.92	−1.83	−1.78
10%							
5	−2.64	−2.45	−2.35	−2.30	−2.25	−2.18	−2.07
	(−2.51)	(−2.30)	(−2.20)	(−2.15)	(−2.11)	(−2.02)	(−1.95)
10	−2.14	−1.98	−1.92	−1.88	−1.85	−1.77	−1.74
30	−2.08	−1.94	−1.88	−1.85	−1.82	−1.76	−1.72
50	−2.09	−1.96	−1.90	−1.87	−1.84	−1.79	−1.73
70	−2.11	−1.97	−1.91	−1.88	−1.85	−1.79	−1.75
100	−2.12	−1.98	−1.92	−1.89	−1.86	−1.79	−1.76

Note:critical values of the truncated statistic are represented in parenthesis if they differ from $\bar{t}_{NL}$.

[1] The critical values for both intercept and time trend are available from the authors.
[2] The critical values for both intercept and time trend are available upon request.

Table 3
Exact critical values of $\bar{Z}_{\mathrm{NL}}$ statistic.

Only intercept							
$T\downarrow \backslash N\rightarrow$	5	10	15	20	25	50	100
1%							
5	−3.48	−3.38	−3.26	−3.22	−3.16	−3.03	−2.89
	(−2.73)	(−2.64)	(−2.56)	(−2.60)	(−2.56)	(−2.47)	(−2.47)
10	−2.56	−2.49	−2.48	−2.45	−2.46	−2.40	−2.43
30	−2.52	−2.52	−2.27	−2.31	−2.29	−2.33	−2.35
50	−2.19	−2.23	−2.23	−2.25	−2.26	−2.28	−2.34
70	−2.17	−2.20	−2.21	−2.24	−2.23	−2.21	−2.31
100	−2.15	−2.22	−2.20	−2.21	−2.22	−2.27	−2.30
5%							
5	−1.62	−1.70	−1.71	−1.76	−1.76	−1.77	−1.76
	(−1.82)	(−1.77)	(−1.74)	(−1.77)	(−1.73)	(−1.72)	(−1.72)
10	−1.70	−1.68	−1.71	−1.69	−1.68	−1.69	−1.70
30	−1.57	−1.61	−1.62	−1.62	−1.65	−1.66	−1.68
50	−1.56	−1.58	−1.61	−1.63	−1.62	−1.63	−1.68
70	−1.55	−1.58	−1.59	−1.61	−1.62	−1.62	−1.64
100	−1.54	−1.59	−1.60	−1.59	−1.61	−1.62	−1.64
10%							
5	−1.26	−1.14	−1.17	−1.21	−1.21	−1.26	−1.27
	(−1.36)	(−1.34)	(−1.32)	(−1.35)	(−1.33)	(−1.33)	(−1.33)
10	−1.28	−1.28	−1.31	−1.30	−1.29	−1.30	−1.31
30	−1.23	−1.27	−1.29	−1.28	−1.30	−1.31	−1.33
50	−1.24	−1.26	−1.28	−1.30	−1.29	−1.29	−1.33
70	−1.22	−1.25	−1.26	−1.28	−1.28	−1.28	−1.29
100	−1.23	−1.25	−1.26	−1.25	−1.26	−1.27	−1.28

Note:critical values of the truncated statistic are represented in parenthesis if they differ from $\bar{Z}_{\mathrm{NL}}$.

$\varepsilon_{i,t} = \rho_i\varepsilon_{i,t-1} + u_{i,t}$ where $u_{i,t}\sim N(0,1)$ and $\rho_i\in\{0, 0.5\}$. Heterogeneous intercepts are removed by demeaning of $y_{i,t}$ and initial values $y_{i,0}$ are set as zero for each i. All the experiments are examined for the 5% confidence level (Table 4).

Because both $\bar{Z}_{\mathrm{NL}}$ and Z_{tbar} statistics have the same null hypothesis, their empirical size performances are not far away from each other. Empirical powers are estimated under the stationary alternative by setting the parameters such that $\theta_i\in\{0.01, 1\}$ and $\gamma_i\in\{-0.1, -0.75, -1\}$ for all i (Table 5).

For the case $\theta_i = 0.01$ and $\gamma_i\in\{-0.75, -1\}$, $\bar{Z}_{\mathrm{NL}}$ statistic has better power performance than Z_{tbar} statistic for small samples. On the other hand, for the case $\theta_i = 1$ and $\gamma_i = -0.1$, the power of Z_{tbar} outperforms our $\bar{Z}_{\mathrm{NL}}$ statistic because the DGP turns out to be a linear model when θ_i increases sufficiently. Furthermore, when $\theta_i = 0.01$ and $\gamma_i = -0.1$, both of the tests suffer from low power, but their power rises 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 $\bar{t}_{\mathrm{NL}}$ and augmented $\bar{Z}_{\mathrm{NL}}$, say $\bar{Z}_{\mathrm{ANL}}$, statistics

Table 4
The size comparisons of alternative tests.

$T\downarrow \backslash N\rightarrow$	5		10		25		50		100	
					$\rho=0$					
	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}
10	0.051	0.052	0.050	0.053	0.056	0.056	0.052	0.049	0.051	0.052
25	0.053	0.056	0.050	0.049	0.045	0.050	0.052	0.051	0.052	0.050
50	0.050	0.047	0.051	0.051	0.050	0.050	0.051	0.053	0.055	0.045
100	0.054	0.053	0.055	0.052	0.052	0.053	0.053	0.057	0.051	0.052
					$\rho=0.5$					
10	0.050	0.055	0.049	0.047	0.050	0.052	0.049	0.047	0.057	0.059
25	0.052	0.047	0.056	0.060	0.051	0.054	0.053	0.050	0.049	0.054
50	0.054	0.045	0.055	0.058	0.049	0.045	0.053	0.058	0.057	0.060
100	0.044	0.042	0.050	0.041	0.055	0.045	0.042	0.044	0.048	0.058

Table 5
The power comparisons of alternative tests.

$T\downarrow \backslash N$–5			10		25		50		100	
					$\theta=0.01$	$\gamma=-1.00$				
	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}	$\bar{Z}_{\mathrm{NL}}$	Z_{tbar}
10	0.098	0.068	0.105	0.090	0.166	0.154	0.228	0.223	0.381	0.379
25	0.211	0.162	0.346	0.259	0.677	0.594	0.918	0.848	0.996	0.981
50	0.602	0.456	0.887	0.763	0.999	0.977	1.000	1.000	1.000	1.000
100	0.991	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
					$\theta=0.01$	$\gamma=-0.75$				
10	0.102	0.086	0.114	0.098	0.125	0.122	0.191	0.190	0.279	0.279
25	0.182	0.156	0.270	0.216	0.549	0.459	0.798	0.701	0.980	0.949
50	0.465	0.335	0.770	0.618	0.984	0.945	1.000	1.000	1.000	1.000
100	0.956	0.925	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000
					$\theta=0.01$	$\gamma=-0.1$				
10	0.052	0.054	0.066	0.054	0.064	0.059	0.065	0.065	0.080	0.082
25	0.074	0.072	0.092	0.091	0.122	0.130	0.187	0.198	0.286	0.303
50	0.098	0.095	0.159	0.140	0.309	0.296	0.498	0.503	0.774	0.778
100	0.213	0.200	0.352	0.336	0.719	0.694	0.938	0.936	0.998	0.998
					$\theta=1.0$	$\gamma=-0.1$				
10	0.079	0.085	0.082	0.091	0.150	0.169	0.217	0.243	0.314	0.390
25	0.146	0.151	0.206	0.235	0.412	0.496	0.714	0.794	0.933	0.973
50	0.316	0.378	0.588	0.711	0.908	0.979	0.999	1.000	1.000	1.000
100	0.811	0.954	0.979	1.000	1.000	1.000	1.000	1.000	1.000	1.000

and compare our results with the IPS tests t_{nbar} and W_{tbar}.[3] 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.[4] Bootstrap algorithm is below.[5]

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

$$\Delta y_{i,t} = d_i + \delta_i y^3_{i,t-1} + \sum_{j=1}^{p_i} \beta_{i,j}\Delta y_{i,t-j} + \varepsilon_{i,t} \tag{9}$$

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

$$\hat{\varepsilon}_{i,t} = \Delta y_{i,t} - \hat{d}_i - \sum_{j=1}^{p_i} \hat{\beta}_{i,j}\Delta y_{i,t-j} \tag{10}$$

(iii) Stine (1987) suggests that residuals have to be centered with

$$\tilde{\varepsilon}_t = \hat{\varepsilon}_t - (T-p-2)^{-1} \sum_{t=p+2}^{T} \hat{\varepsilon}_t \tag{11}$$

where $\hat{\varepsilon}_t = (\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t}, \ldots, \hat{\varepsilon}_{N,t})'$ and $p = \max_i (p_i)$. Furthermore, we develop the $N\times T[\tilde{\varepsilon}_{i,t}]$ 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] The OECD sample includes 25 countries, namely, Australia, Austria, Belgium, Canada, Denmark, Finland, France, Greece, Iceland, Ireland, Italy, Japan, Korea, Luxemburg, Mexico, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the UK and the US.

[4] The data were downloaded from http://pwt.econ.upenn.edu/php_site/pwt_index.php.

[5] Separate programming code is developed in Matlab 7.5.

Table 6
Empirical results for alternative tests.

	$\bar{t}^*_{\mathrm{NL}}$	$\bar{Z}^*_{\mathrm{ANL}}$	t^*_{nbar}	W^*_{nbar}
Only intercept	−1.612	0.105	−1.408	0.749
	(0.0035)	(0.0035)	(0.0085)	(0.0085)
Intercept and trend	−2.3744	−1.723	−1.0244	8.0263
	(0.00)	(0.00)	(0.959)	(0.959)

Note: p-values are in the paranthesis.

the bootstrap residuals as $\tilde{\varepsilon}^*_{i,t}$ where $t=1,2,\ldots,T^*$ and $T^*=2T$ in our application.

(iv) We produce bootstrap $\Delta y^*_{i,t}$ recursively from

$$\Delta y^*_{i,t} = \hat{d}_i + \sum_{j=1}^{p_i} \hat{\beta}_{i,j}\Delta y^*_{i,t-j} + \tilde{\varepsilon}^*_{i,t} \tag{12}$$

where $\hat{d}_i$ and $\hat{\beta}_{i,j}$ are the estimations from step (ii) and $\Delta y^*_{i,t-p_i}=0$ for $p_i=1,2,\ldots,6$.

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

$$y^*_{i,t} = \sum_{j=1}^{t} \Delta y^*_{i,j} \tag{13}$$

The bootstrap statistics $\bar{t}^*_{\mathrm{NL}}$ and $\bar{Z}^*_{\mathrm{ANL}}$ are computed for each bootstrap replication by running the regression

$$\Delta y^*_{i,t} = d_i + \gamma_i\left(y^*_{i,t-1}\right)^3 + \sum_{j=1}^{p_i} \theta_{i,j}\Delta y^*_{i,t-j} + v_{i,t} \tag{14}$$

where noting that the last T observations of $y^*_{i,t}$ and $\Delta y^*_{i,t}$ are used in this regression. The bootstrap empirical distribution of $\bar{t}^*_{\mathrm{NL}}$and $\bar{Z}^*_{\mathrm{ANL}}$statistics, generated by employing 2000 replications, are used to have their p-values. The same procedure is also applied for the IPS statistics t^*_{nbar} and W^*_{tbar}.[6] 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 t^*_{nbar} and W^*_{tbar} 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 $\delta_i=0$, $\mathbf{y}_{i,-1}$ can be written as follows:

$$\mathbf{y}_{i,-1} = y_{i0}\tau_T + \mathbf{u}_{i,-1} \tag{A.1}$$

where y_{i0} is a fixed initial value and $\mathbf{u}_{i,-1} = (u_{i0}, u_{i1}, \ldots, u_{i,T-1})'$ with $u_{i,t} = \sum_{j=1}^{t} \varepsilon_{ij}$

Secondly, let's define the following $T\times T$ matrix,

$$\Lambda_T = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \ldots & 0 \end{bmatrix}$$

and using this matrix, we have $\mathbf{u}_{i,-1} = \Lambda_T\varepsilon_i$ where $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i1}, \ldots, \varepsilon_{i,T})'$. The cubic form of Eq. (A.1) can be expressed:

$$\mathbf{y}^3_{i,-1} = y^3_{i0}\tau_T + 3y^2_{i0}\tau_T\odot(\Lambda_T\varepsilon_i) + 3y_{i0}\tau_T\odot(\Lambda_T\varepsilon_i)^2 + (\Lambda_T\varepsilon_i)^3 \tag{A.2}$$

where $\Lambda_T\varepsilon_i$ is $T\times 1$ vector and $\odot$ denotes for Hadamard product. Furthermore, multiplying both sides with M_τ yields

$$M_\tau\mathbf{y}^3_{i,-1} = 3y^2_{i0}\tau_T\odot(M_\tau\Lambda_T\varepsilon_i) + 3y_{i0}\tau_T\odot\left[(\Lambda_T\varepsilon_i)^2 - \sigma_i^2 t\right] + M_\tau(\Lambda_T\varepsilon_i)^3 \tag{A.3}$$

where $M_\tau(\Lambda_T\varepsilon_i)^2 = (\Lambda_T\varepsilon_i)^2 - \sigma_i^2 t$ and $t=(1,2,\ldots,T)'$. Notice here that taking expectation of the process (A.2) on the time domain is $E\left(\mathbf{y}^3_{i,-1}\right) = y^3_{i0}\tau_T + 3y_{i0}\sigma_i^2\tau_T\odot t$ and $E(\Lambda_T\varepsilon_i)^3 = 0$ due to symmetric normal distribuion assumption for ε_i.

Now setting $y_{i0}=0$ implies

$$M_\tau\mathbf{y}^3_{i,-1} = M_\tau(\Lambda_T\varepsilon_i)^3 \tag{A.4}$$

Therefore, imposing $y_{i0}=0$ removes the nuisance parameter σ_i^2 in (A.3). Both $\Delta\mathbf{y}_i=\varepsilon_i$ and (A.4) are plugged into the test-statistic (5), we have

$$t_{i,\mathrm{NL}} = \frac{\sqrt{T-1}\varepsilon_i' M_\tau(\Lambda_T\varepsilon_i)^3}{(\varepsilon_i' M_\tau\varepsilon_i)^{1/2}\left[(\varepsilon_i'\Lambda_T')^3 M_\tau(\Lambda_T\varepsilon_i)^3\right]^{1/2}} \tag{A.5}$$

which is invariant with respect to σ_i^2.

Now let's define, $\mathbf{v}_i = \frac{\varepsilon_i}{\sigma_i} \sim N(0, I_T)$ and putting this into Eq. (A.5) gives directly

$$t_{i,\mathrm{NL}} = \frac{\sqrt{T-1}\mathbf{v}_i' M_\tau(\Lambda_T\mathbf{v}_i)^3}{(\mathbf{v}_i' M_\tau\mathbf{v}_i)^{1/2}\left[(\mathbf{v}_i'\Lambda_T')^3 M_\tau(\Lambda_T\mathbf{v}_i)^3\right]^{1/2}} \tag{A.6}$$

so that σ_i^4 is cancelled out in our test-statistic. □ **QED**

References

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration. International Economic Review 38, 627–645.

Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H., Taylor, R.L., 1991. Bootstrapping unstable first order autoregressive processes. Annals of Statistics 19, 1098–1101.

Im, K.S., Pesaran, H., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53–74.

Kapetanios, G., Shin, Y., Snell, A., 2003. Testing for a unit root in the nonlinear STAR framework. Journal of Econometrics 112, 359–379.

Michael, P., Nobey, R.A., Peel, D.A., 1997. Transactions costs and nonlinear adjustment in real exchange rates: an empirical investigation. Journal of Political Economy 105, 862–879.

Pesaran, M.H., 2007. A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22, 265–312.

Stine, R.A., 1987. Estimating properties of autoregressive forecasts. Journal of the American Statistical association 82, 1072–1078.

[6] In computing $\bar{Z}^*_{\mathrm{ANL}}$and W^*_{tbar}moments are regenerated by taking $T=52$ observations and various lag orders p_i into consideration.