A fixed-T version of Breitung's panel data unit root test

Karavias, Yiannis; Tzavalis, Elias

DOI:
10.1016/j.econlet.2014.04.029

Citation for published version (Harvard):
Karavias, Y & Tzavalis, E 2014, 'A fixed-T version of Breitung's panel data unit root test', *Economics Letters*, vol. 124, no. 1, pp. 83-87. https://doi.org/10.1016/j.econlet.2014.04.029

A fixed-T Version of Breitung's Panel Data Unit Root Test

Yiannis Karavias[a,*] and Elias Tzavalis[b]

a: School of Economics, University of Nottingham

b: Athens University of Economics & Business

Abstract

We extend Breitung's (2000) panel data unit root test to the case of fixed time (T) dimension while still allowing for heteroscedastic and serially correlated error terms. The analytic local power function of the new test is derived assuming that only the cross section dimension of the panel grows large. It is found that, if the errors are serially correlated, the test has non-trivial power. Monte Carlo experiments show that the suggested test is more powerful when the number of cross section units is moderate or large, regardless of the number of time series observations.

JEL classification: C22, C23

Keywords: Panel unit root; local power function; serial correlation; incidental trends

The authors would like to thank the editor Badi Baltagi, two anonymous referees, Jörg Breitung and Steve Leybourne for useful comments.

*Corresponding author. University of Nottingham, Nottingham NG7 2RD, UK. Tel: +441159515480, Email addresses: ioannis.karavias@nottingham.ac.uk ; etzavalis@aueb.gr

1 Introduction

To improve the power performance of panel data unit root tests in the presence of heterogeneous individual linear trends, known as incidental trends, Breitung (2000) proposed a statistic based on an orthogonal transformation of the individual series of the panel. The test does not require an inconsistency adjustment of the estimator of the autoregressive parameter φ as opposed to other tests in the literature, see e.g., Baltagi (2013) for a survey. Although it was found to be consistent and have superior power in small samples for values of φ not far from unity (e.g., $\varphi = 0.95$), its asymptotic local power in a $T^{-1}N^{-1/2}$ neighbourhood of unity is trivial and equivalent to that of the asymptotically bias corrected tests (see, Moon et al (2007)).

In this paper, we extend Breitung's (2000) test in two directions. First, we allow the time dimension T of the panel to be finite (fixed) while allowing for heterogeneity, heteroscedasticity, and serial correlation in the error terms. Second, we derive the fixed-T asymptotic local power function of the new test. These extensions make the application of the test valid in cases of short-T panels, often met in practice, and under higher than first order serial correlation. The paper provides a number of interesting results. First, it shows that the fixed-T version of the test can further improve its small sample size and power performance in short panels, compared to its large-T version. Second, the new test also has trivial asymptotic local power in a $N^{-1/2}$ neighbourhood of unity when the error terms are independently distributed over time, which explains analytically Breitung's (2000) findings in his Monte Carlo experiment. Third, when the error terms are serially correlated, the estimator of φ becomes inconsistent and thus, the test needs an inconsistency correction. Fourth, there are forms of serial correlation of the error terms for which the test has non-trivial asymptotic local power.

The paper is organized as follows. Section 2 introduces the new test and provides its local power function. Section 3 presents the results of our Monte Carlo exercise, while Section 4 concludes the paper. All proofs are relegated to the Appendix.

2 The test statistic and its asymptotic local power

Consider the following AR(1) panel data model with individual effects and incidental trends:

$$y_i = \varphi y_{i-1} + (1-\varphi)a_i e + \varphi\beta_i e + (1-\varphi)\beta_i \tau + u_i, \quad i = 1, 2, ..., N, \tag{1}$$

where $y_i = (y_{i1}, ..., y_{iT})'$ and $y_{i-1} = (y_{i0}, ..., y_{iT-1})'$ are $T \times 1$ vectors, u_i is the $T \times 1$ vector of error terms u_{it}, and a_i and β_i are respectively the individual effects and the slopes of incidental trends of the model. The $T \times 1$ vectors e and τ have elements $e_t = 1$ and $\tau_t = t$ for $t = 1, ..., T$. Next, define the autoregressive coefficient φ as $\varphi_N = 1 - \frac{c}{\sqrt{N}}$ (see also Madsen (2010)). Then, the null hypothesis of a unit root in φ against its alternative of stationarity (i.e., $\varphi < 1$) can be respectively written as

$$H_0\colon c = 0 \quad \text{and} \quad H_1\colon c > 0,$$

where c is the local to unity parameter. The asymptotic distribution of the extension of Breitung's (2000) test statistic is derived under the following assumption.

Assumption A

(i) $\{u_i\}$, $i \in \{1, 2, ..., N\}$, are independent random vectors with means $E(u_i) = 0$ and heterogeneous variance-covariance matrices $\Gamma_i \equiv E(u_i u_i') \equiv [\gamma_{i,ts}]$, where $\gamma_{i,ts} = E(u_{it} u_{is}) = 0$ for $t < s$ and $s = t+p+1, ..., T$. The maximum order of serial correlation in u_i is $p = T-2$. All $4 + \varepsilon$ mixed moments are finite.

(ii) $\Gamma = \lim_N \frac{1}{N} \sum_{i=1}^{N} \Gamma_i$ is a finite, positive definite matrix and $\lim_N (N\Gamma)^{-1} \Gamma_i = 0$, for all i. β_i are independent across i with finite $4 + \varepsilon$ moments and with $\lim_N \frac{\max(E(\beta_i^2))}{(\sum_{i=1}^{N} E(\beta_i^2))} = 0$.

(iii) u_{it} are independent of y_{i0}, a_i and β_i,while y_{i0} and a_i are independent across i and have finite $4 + \epsilon$ moments.

Assumption A allows us to derive the distribution of the fixed-T version of Breitung's (2000) panel unit root test statistic under H_0. Condition (i) determines the order of serial correlation of u_{it} and together with condition (ii) provide the necessary assumptions for the application of the Lindeberg-Feller CLT. Condition (iii) is needed for the derivation of the local power function.

Breitung's (2000) test is based on the forward orthogonal deviations transformation of the individual series of model (1), y_{it}, which removes individual effects and incidental trends. In a first step, the initial observations y_{i0} are subtracted from y_{it}, i.e., $z_{it} = y_{it} - y_{i0}$. Then, define the following $(T-1) \times T$ matrices:

$$A = \begin{pmatrix} 0_{1\times T} \\ GH \end{pmatrix} \text{ and } B = \begin{pmatrix} 0_{1\times(T-2)} & 0 & 0 \\ I_{T-2} & 0_{(T-2)\times 1} & -\frac{1}{T}\tau_{T-2} \end{pmatrix}, \text{ where}$$

$$G = \begin{pmatrix} \sqrt{\frac{T-2}{T-1}} & & & 0 \\ & \sqrt{\frac{T-3}{T-2}} & & \\ & & \ddots & \\ 0 & & & \sqrt{\frac{1}{2}} \end{pmatrix} \text{ and } H = \begin{pmatrix} 1 & -\frac{1}{T-1} & \cdots & \cdots & \cdots & -\frac{1}{T-1} \\ & \ddots & -\frac{1}{T-2} & & & -\frac{1}{T-2} \\ & & \ddots & & & \vdots \\ & & & 1 & -\frac{1}{2} & -\frac{1}{2} \\ \cdots & \cdots & \cdots & 0 & 1 & -1 \end{pmatrix},$$

with dimensions $(T-2)\times(T-2)$ and $(T-2)\times T$ respectively, and vector $\tau_{T-2} = (1, 2, ..., T-2)'$. In the case that $u_{it} \sim IID(0, \sigma^2)$, multiplying Δz_i with matrix A and z_i with matrix B implies the following orthogonal moment conditions under H_0: $c = 0$:

$$E(z_i' B' A \Delta z_i) = 0. \tag{2}$$

These can be tested based on the following least squares estimator of φ:

$$\hat{\varphi}_{FOD} = 1 + \left(\sum_{i=1}^{N} z_i' B' B z_i \right)^{-1} \left(\sum_{i=1}^{N} z_i' B' A \Delta z_i \right),$$

which is equal to Breitung's (2000) estimator plus 1. This estimator is consistent under H_0: $c = 0$, i.e., $p\lim_N \hat{\varphi}_{FOD} = 1$. In the more general case where $\Gamma \neq \sigma^2 I_T$, estimator $\hat{\varphi}_{FOD}$ becomes inconsistent and its asymptotic bias is equal to $p\lim_N(\hat{\varphi}_{FOD} - 1) = \frac{tr((\Lambda+I_T)'B'A\Gamma)}{tr((\Lambda+I_T)'B'B(\Lambda+I_T)\Gamma)}$, where Λ is a $T \times T$ matrix which has unities at its lower than its main diagonals, and zero elsewhere, and I_T is a $T \times T$ identity matrix.[1] Thus, to test moment conditions (2), $\hat{\varphi}_{FOD}$ needs to be corrected for its inconsistency (see, e.g., Harris and Tzavalis (1999)).

Theorem 1 *Let conditions (i) and (ii) of Assumption A hold and $N \to \infty$, then under $H_0: c = 0$ we have*

$$UB_T = \sqrt{N} V^{-1/2} \hat{\delta} \left(\hat{\varphi}_{FOD} - 1 - \frac{\hat{b}}{\hat{\delta}} \right) \xrightarrow{d} N(0, 1),$$

where $\frac{\hat{b}}{\hat{\delta}} = \frac{tr(\Phi_p \hat{\Gamma})}{\frac{1}{N}\sum_{i=1}^{N} z_i' B' B z_i}$, $\hat{\Gamma} = \frac{1}{N}\sum_{i=1}^{N} \Delta z_i \Delta z_i'$, $\Phi_p = \Psi_p - \frac{e'\Psi_p e}{e'Me} M$ *with* Ψ_p *a* $T \times T$ *matrix having in its diagonals* $\{-p, .., 0, ...p\}$ *the corresponding elements of matrix* $\Xi = (\Lambda + I_T)' B' A$, *and zero elsewhere,* M *is a* $T \times T$ *selection matrix with elements* $m_{ts} = 0$, *if* $\gamma_{ts} \neq 0$, *and* $m_{ts} = 1$, *if* $\gamma_{ts} = 0$, *and* $V = vec(\Xi' - \Phi_p')' \Theta vec(\Xi' - \Phi_p')$ *where* $\Theta = \frac{1}{N} \sum_{i=1}^{N} Var(vec(\Delta z_i \Delta z_i'))$.[2]

The bias correction of $\hat{\varphi}_{FOD}$, assumed by Theorem 1, relies on selection matrices Ψ_p and Φ_p. Ψ_p selects the non-zero elements of Γ (estimated by $\hat{\Gamma} = \frac{1}{N}\sum_{i=1}^{N} \Delta z_i \Delta z_i'$), i.e., $\gamma_{i,ts} = E(u_{it} u_{is}) \neq 0$, for $t, s < p$, to correct for the bias of the numerator of $\hat{\varphi}_{FOD}$ coming from the serial correlation effects in u_{it}. Since, under H_0, $\hat{\Gamma}$ is not a consistent estimator of Γ due to the nuisance parameter effects of incidental trends, i.e. $p\lim_N \hat{\Gamma} = \Gamma + \beta^2 ee'$ by

[1] This happens because $tr((\Lambda + I_T)'B'A) = 0$ and $tr((\Lambda + I_T)'B'A\Gamma) \neq 0$.

[2] An alternative specification of UB_T for $u_{it} \sim NIID(0, \sigma^2)$ is $UB_{T,2} = \sqrt{N} V_2^{-1/2} (\hat{\varphi}_{FOD} - 1) \xrightarrow{d} N(0, 1)$, where $V_2 = \frac{2tr(A_\Xi^2)}{tr((\Lambda+I_T)'B'B(\Lambda+I_T))^2}$, with $A_\Xi = \frac{1}{2}(\Xi + \Xi')$.

condition (ii) where $\beta^2 = p\lim_N \frac{1}{N}\sum_{i=1}^N E(\beta_i^2)$, the limiting distribution of $\hat{\varphi}_{FOD} - 1$ must be also corrected for these effects, defined β^2. This is done by premultiplying $\hat{\Gamma}$ with Φ_P. The latter adjusts Ψ_p by $tr\left(M\hat{\Gamma}\right)/e'Me$, which is a consistent estimator of β^2. Implementing test statistic UB_T requires a consistent estimator of variance V. Under H_0, this is given as $\hat{V} = vec(\Xi' - \Phi_p')'\hat{\Theta}vec(\Xi' - \Phi_p')$ where $\hat{\Theta} = \frac{1}{N}\sum_{i=1}^N (vec(\Delta z_i \Delta z_i')vec(\Delta z_i \Delta z_i')')$. The main difference between UB_T and Breitung's statistic is the replacement of a T-consistent variance estimator of u_i with a N-consistent one.

To study the asymptotic power of UB_T under H_1: $c > 0$, we will rely on a "slope" parameter defined in local power functions of form $\Phi(z_a + ck)$ as k, where Φ is the standard normal cdf and z_a denotes the *α-level* percentile. Since Φ is strictly monotonic, a larger k means greater power for the same value of c. If $k > 0$, then test statistic UB_T will have non-trivial power. If $k = 0$, it will have trivial power, which is equal to a. Finally, if $k < 0$, it will be biased. In the next theorem, we derive the limiting distribution of UB_T under H_1.

Theorem 2 *Under Assumption A and H_1: $c > 0$, we have*

$$UB_T = \sqrt{N}V^{-1/2}\hat{\delta}\left(\hat{\varphi}_{FOD} - 1 - \frac{\hat{b}}{\hat{\delta}}\right) \xrightarrow{d} N(-ck, 1), \tag{3}$$

as $N \to \infty$, where

$$k = \frac{tr(\Lambda' B' A \Lambda \Gamma) + tr(B' A \Lambda \Gamma) + tr(\Lambda' B' A \Gamma) + tr(F' B' A \Gamma) - tr(\Lambda' \Phi_p \Gamma) - tr(\Phi_p \Lambda \Gamma)}{\sqrt{V}}, \tag{4}$$

where F is defined in the Appendix.

The result of Theorem 2 implies that UB_T can have non-trivial power, as k can be positive. For instance, this can happen when u_{it} follows MA(1) process $u_{it} = \varepsilon_{it} + \theta_i \varepsilon_{it-1}$ with $\theta_i < 0$. It can be attributed to the fact that the individual series of the panel y_{it} will look similar to those generated by model (1) with a common linear trend for all i. In this case the incidental trends problem does not apply (see Moon et al (2007)). However, power becomes trivial if u_{it} are serially uncorrelated. Then, UB_T suffers from the problem of zero local power due to incidental trends, noted by Moon et al (2007) for large-T panel unit root tests.[3] This explains Breitung's (2000) Monte Carlo findings. Note that k also depends on the moments of nuisance parameters β_i, entering its denominator k through variance function V. For instance, if u_{it} and β_i are zero-mean normally distributed random variables,

[3]The limiting distribution of $UB_{T,2}$ under H_1: $c > 0$ becomes $UB_{T,2} = \sqrt{N}V_2^{-1/2}(\hat{\varphi}_{FOD} - 1) \xrightarrow{d} N(-ck_2, 1)$, where $k_2 = 0$, which means that the test has trivial power.

then V is given as $V = 2tr((A_{FOD}\Gamma + \beta^2 A_{FOD}ee')^2)$, where $A_{FOD} = \frac{1}{2}(\Xi + \Xi' - \Phi_p - \Phi_p')$ (see proof of Theorem 1).

3 Simulation Results

The aim of our simulation study is twofold: first, to examine if the small sample size and power performance of the fixed-T test statistic UB_T is satisfactory compared to its large-T version and, second, to investigate if the asymptotic local power function of the test can approximate its actual power. In our analysis, we assume that u_{it} are generated as $u_{it} = \varepsilon_{it}+\theta\varepsilon_{it-1}$, with $\varepsilon_{it} \sim NIID(0,1)$ and $\theta \in \{-0.8, -0.4, 0, 0.4, 0.8\}$. We set $y_{i0} = 0$ and $a_i = 0$, without loss of generality as these parameters do not appear in the local power function. For β_i, we consider $\beta_i = 0$ and $\beta_i \sim NIID(0,1)$, while $\varphi \in \{1, 0.95\}$, $N \in \{20, 50, 100\}$ and $T \in \{7, 10, 15, 20, 50\}$. Rejection frequencies are computed based on 10000 replications at 5% significance level.

	N	20				50				100			
	φ/T	10	20	50	100	10	20	50	100	10	20	50	100
UB_T	1	0.093	0.104	0.110	0.114	0.077	0.087	0.088	0.093	0.070	0.070	0.073	0.080
	0.95	0.061	0.083	0.267	0.840	0.065	0.106	0.582	0.997	0.074	0.158	0.854	1
UB	1	0.082	0.074	0.063	0.061	0.079	0.069	0.066	0.057	0.075	0.066	0.059	0.057
	0.95	0.055	0.069	0.291	0.886	0.059	0.101	0.547	0.998	0.064	0.138	0.823	1

Table 1: Size and size-adjusted power of test statistics UB_T and UB, for $\theta = 0$ and $p = 0$.

Table 1 presents the size and the size-adjusted power of UB_T and Breitung's statistic, denoted UB. This is done for $\theta = 0$ and $\beta_i = 0$, for all i (see also Breitung (2000)). The results of the table indicate that both the size and power of UB_T are satisfactory (see De Blander and Dhaene (2012)). Its power increases with N or T, but faster with T than N. For small N and large T, UB has better size and more power than UB_T. However, as N increases UB_T improves its size and is more powerful than the UB test irrespective of T. This qualifies application of UB_T also in cases where both dimensions N and T of the panel are large.

Table 2, which presents size and power of statistic UB_T for non-zero θ, indicates that positive serial correlation ($\theta > 0$) in error terms u_{it} increases considerably the power of UB_T, even for very small values of T and N. Also, the size performance of UB_T is unaffected when u_{it} are negatively correlated ($\theta < 0$). To see how well the asymptotic theory approximates the local power of UB_T, Table 3 presents power values when $\varphi = 1 - c/\sqrt{N}$, for $c = 1$,

$N \in \{50, 100, 300, 1000\}$, $T = 10$ and two cases of β_i: $\beta_i = 0$ and $\beta_i \sim NIID(0,1)$. The results of Table 3 indicate that the estimates of the power obtained by our Monte Carlo experiment tend to approximate their theoretical values (TV). For $\theta < 0$, the test has non-trivial local power while for $\theta > 0$, it is biased. Finally, the power losses for $\beta_i \sim NIID(0,1)$ are not very large. They become minimal for $\theta = 0$, where β_i does not affect the local power function.

	N	20				50				100			
θ	φ/T	10	20	50	100	10	20	50	100	10	20	50	100
-0.8	1	0.054	0.054	0.057	0.054	0.051	0.057	0.054	0.057	0.053	0.054	0.056	0.051
	0.95	0.061	0.073	0.099	0.114	0.070	0.086	0.128	0.164	0.075	0.105	0.173	0.246
-0.4	1	0.051	0.059	0.066	0.080	0.050	0.055	0.066	0.071	0.054	0.054	0.059	0.061
	0.95	0.062	0.091	0.252	0.711	0.074	0.115	0.435	0.695	0.077	0.138	0.656	0.998
0.4	1	0.079	0.096	0.113	0.111	0.070	0.084	0.082	0.089	0.061	0.069	0.082	0.079
	0.95	0.092	0.161	0.489	0.950	0.093	0.181	0.728	0.999	0.093	0.216	0.924	1.00
0.8	1	0.074	0.097	0.111	0.122	0.068	0.078	0.090	0.090	0.064	0.073	0.080	0.078
	0.95	0.095	0.168	0.496	0.958	0.090	0.185	0.747	0.999	0.100	0.219	0.927	1.00

Table 2: Size and power of the fixed-T panel root test statistic UB_T when $\theta \neq 0$ and $p = 1$.

$\beta_i = 0$, $i = 1, ..., N$						$\beta_i \sim N(0,1)$, $i = 1, ..., N$					
$\theta\backslash N$	50	100	300	1000	TV	$\theta\backslash N$	50	100	300	1000	TV
-0.8	0.125	0.123	0.113	0.096	0.067	-0.8	0.091	0.086	0.084	0.076	0.059
-0.4	0.142	0.132	0.109	0.099	0.059	-0.4	0.089	0.086	0.075	0.068	0.054
0	0.222	0.182	0.115	0.086	0.050	0	0.203	0.154	0.105	0.081	0.050
0.4	0.286	0.213	0.132	0.088	0.045	0.4	0.173	0.138	0.102	0.077	0.047
0.8	0.308	0.233	0.147	0.096	0.044	0.8	0.191	0.154	0.111	0.079	0.046

Table 3: Local power values of statistic UB_T for $T = 10$, when $u_{it} = \varepsilon_{it} + \theta\varepsilon_{it-1}$ and $p = 1$.

4 Conclusions

This paper extends Breitung's (2000) panel unit root test to the case of fixed-T time dimension and derives its local power. It shows that the new test can further improve its small sample size and power performance in short panels, compared to its large-T version. In addition to this, allowing for serial correlation in error terms leads to a test which can have non-trivial local power in the presence of incidental trends. Monte Carlo analysis confirms the asymptotic results provided by the paper.

References

[1] Baltagi, B.H., 2013. Econometric analysis of panel data, 5th Edition. Wiley and Sons, NY.

[2] Breitung J., 2000. The local power of some unit root tests for panel data. In Badi H. Baltagi, Thomas B. Fomby, R. Carter Hill (ed.) Nonstationary Panels, Panel Cointegration, and Dynamic Panels (Advances in Econometrics, Volume 15), Emerald Group Publishing Limited, pp.161-177.

[3] De Blander, R., Dhaene, G., 2012. Unit root tests for panel data with AR(1) errors and small T. The Econometrics Journal, 15, 101-124.

[4] Harris R.D.F. and E. Tzavalis 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics, 91, 201-226.

[5] Madsen E., 2010. Unit root inference in panel data models where the time-series dimension is fixed: a comparison of different tests. Econometrics Journal 13, 63-94.

[6] Moon H.R., Perron B. & Phillips P.C.B., 2007. Incidental trends and the power of panel unit root tests. Journal of Econometrics, 141(2), 416-459.

[7] Schott J.R. 1996. Matrix Analysis for Statistics, Wiley-Interscience.

5 Appendix

Theorem 1: Under H_0: $c = 0$, we have $z_i = z_{i-1} + \beta_i e + u_i$ and $z_{i-1} = \Lambda e \beta_i + \Lambda u_i$. Then, the numerator of $\hat{\varphi}_{FOD} - 1$ becomes $\frac{1}{N}\sum_{i=1}^{N} z_i' B' A \Delta z_i = \frac{1}{N}\sum_{i=1}^{N}(z_{i-1}' + \beta_i e' + u_i) B' A(\beta_i e + u_i) = \frac{1}{N}\sum_{i=1}^{N}(u_i'(\Lambda' + I_T) + \beta_i \tau') B' A(\beta_i e + u_i) = \frac{1}{N}\sum_{i=1}^{N} u_i'(\Lambda' + I_T) B' A u_i$, since $(\Lambda + I_T)e = \tau$ and $\tau' B' = 0_{1\times T}$, $B' A e = 0_{T\times 1}$ by construction. By Chebyshev's Weak Law of Large Numbers:

$$\frac{1}{N}\sum_{i=1}^{N} u_i'(\Lambda' + I_T) B' A u_i = \frac{1}{N}\sum_{i=1}^{N} u_i' \Xi u_i \xrightarrow{p} tr(\Xi\Gamma). \tag{5}$$

To see how $tr(\Xi\Gamma)$ can be estimated by $\hat{b}$, write

$$tr(\Phi_p\hat{\Gamma}) = tr(\Psi_p\hat{\Gamma}) - tr(\Psi_p ee')\frac{tr(M\hat{\Gamma})}{e'Me}.$$

This has

$$\begin{aligned} p\lim_N tr(\Phi_p\hat{\Gamma}) &= tr\left[\Psi_p\left(\Gamma + \beta^2 ee'\right)\right] - tr(\Psi_p ee')\frac{tr(M\left(\Gamma+\beta^2 ee'\right))}{e'Me} \\ &= tr\left(\Psi_p\Gamma\right) + tr\left(\Psi_p ee'\right)\beta^2 - tr(\Psi_p ee')\frac{tr(M\Gamma)+tr(Mee'\beta^2)}{e'Me} \\ &= tr\left(\Psi_p\Gamma\right) + tr\left(\Psi_p ee'\right)\beta^2 - tr(\Psi_p ee')\beta^2 \\ &= tr\left(\Psi_p\Gamma\right) = tr(\Xi\Gamma), \end{aligned}$$

because $tr(M\Gamma) = 0$. Similarly, it can be shown that the denominator of $\hat{\varphi}_{FOD} - 1$ has the following limit:

$$\frac{1}{N}\sum_{i=1}^{N} z_i'B'Bz_i = \frac{1}{N}\sum_{i=1}^{N} u_i'(\Lambda' + I_T)B'B(\Lambda + I_T)u_i \xrightarrow{p} tr((\Lambda' + I_T)B'B(\Lambda + I_T)\Gamma). \quad (6)$$

The relationships (5) and (6) imply that the inconsistency of $\hat{\varphi}_{FOD}$ is given as $p\lim_N(\hat{\varphi}_{FOD} - 1) = \frac{tr(\Xi\Gamma)}{tr((\Lambda'+I_T)B'B(\Lambda+I_T)\Gamma)}$. Thus $\hat{\varphi}_{FOD}$ becomes unbiased, if $tr(\Xi\Gamma) = 0$, i.e. $\Gamma = \sigma^2 I_T$. Combining the above, the limiting distribution of UB_T can be derived as follows:

$$\begin{aligned} \sqrt{N}\hat{\delta}\left(\hat{\varphi}_{FOD} - 1 - \hat{b}/\hat{\delta}\right) &= \\ \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} u_i'(\Lambda' + I_T)B'Au_i - \frac{1}{N}\sum_{i=1}^{N}\Delta z_i'\Phi_p\Delta z_i\right) &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Delta z_i'(\Xi - \Phi_p)\Delta z_i \end{aligned}$$

since $\Delta z_i'\Xi\Delta z_i = u_i'\Xi u_i$, where $E(\Delta z_i'(\Xi - \Phi_p)\Delta z_i) = 0$ by construction of Φ_p and $Var(\Delta z_i'(\Xi - \Phi_p)\Delta z_i) = vec(\Xi - \Phi_p)'Var(vec(\Delta z_i \Delta z_i'))vec(\Xi - \Phi_p)$. The result follows by applying the Lindeberg-Feller CLT. If u_i and β_i are zero-mean normally distributed random variables, then Δz_i is normal with $Var(\Delta z_i'(\Xi - \Phi_p)\Delta z_i) = 2tr(\left(A_{FOD}\left(\Gamma_i + E(\beta_i^2)ee'\right)\right)^2)$.

Theorem 2: To prove the theorem, we will employ the following relationships:

$$\begin{aligned} z_i &= \varphi_N z_{i-1} + X\zeta_i + u_i, \quad i = 1, 2, ..., N & (7) \\ z_{i-1} &= \Omega X\zeta_i + \Omega u_i + (w - e)y_{i0}, & (8) \\ \text{and} \quad \Delta z_i &= (\varphi_N - 1)z_{i-1} + X\zeta_i + u_i, & (9) \end{aligned}$$

where $\zeta_i = \begin{pmatrix} (1-\varphi_N)(a_i - y_{i0}) + \varphi\beta_i \\ (1-\varphi_N)\beta_i \end{pmatrix}$, $X = (e, \tau)$, $w = (1, \varphi_N, \varphi_N^2, ..., \varphi_N^{T-1})'$ and

$$\Omega = \begin{pmatrix} 0 & . & . & . & . & . & 0 \\ 1 & 0 & & & & & . \\ \varphi_N & 1 & . & & & & . \\ \varphi_N^2 & \varphi_N & . & . & & & . \\ . & & . & . & . & & . \\ . & & & . & 1 & 0 & . \\ \varphi_N^{T-2} & \varphi_N^{T-3} & . & . & \varphi_N & 1 & 0 \end{pmatrix}$$

. Note that, for $\varphi_N = 1$, we have $\Omega \equiv \Lambda$. The first order Taylor expansions of Ω and w yield

$$\Omega = \Lambda + F(\varphi_N - 1) + o(1) \quad \text{and} \quad w = e + f(\varphi_N - 1) + o(1), \tag{10}$$

respectively, where $F = \frac{d\Omega}{d\varphi_N} \mid_{\varphi_N=1}$ and $f = \frac{dw}{d\varphi_N} \mid_{\varphi_N=1}$(see also Madsen (2010)). ζ_i can be written more compactly as

$$\zeta_i = \frac{c}{\sqrt{N}}\mu_i + \beta_i e_2, \tag{11}$$

where $\frac{c}{\sqrt{N}} = (1-\varphi_N)$, $\mu_i = (a_i - y_{i0} - \beta_i, \beta_i)'$ and $e_2 = (1,0)'$. The following equalities also hold:

$$\begin{aligned} tr(\Xi) &= 0 \ \text{ and } tr(\Lambda' B' A) = -tr(B'A), \\ e'\Xi &= 0_{1\times T} \ \text{ and } \ \Xi e = 0_{T\times 1}, \\ B'AXe_2 &= 0_{T\times 1}, \\ \beta^2 e_2' X' \Lambda' B' A \Lambda X e_2 &= e_2' X' \Lambda' B' A X \tilde{e}, \\ \beta^2 e_2' X' B' A \Lambda X e_2 &= e_2' X' B' A X \tilde{e}, \\ \beta^2 e_2' X' \Phi_p \Lambda X e_2 &= e_2' X' \Phi_p X \tilde{e}, \\ \beta^2 e_2' X' \Lambda' \Phi_p X e_2 &= \tilde{e}' X' \Phi_p X e_2, \end{aligned} \tag{12}$$

where $\tilde{e} = E(\beta_i \mu_i)$. Consider the following formula of test statistic UB_T:

$$\begin{aligned} &\sqrt{N}\hat{\delta}\left(\hat{\varphi}_{FOD} - \varphi_N - \frac{\hat{b}}{\hat{\delta}}\right) = \\ = \ &\sqrt{N}\hat{\delta}\left(1 + \frac{\frac{1}{N}\sum_{i=1}^N z_i' B' A \Delta z_i}{\frac{1}{N}\sum_{i=1}^N z_i' B' B z_i} - \varphi_N - \frac{\frac{1}{N}\sum_{i=1}^N \Delta z_i' \Phi_p \Delta z_i}{\frac{1}{N}\sum_{i=1}^N z_i' B' B z_i}\right) \end{aligned} \tag{13}$$

$$= \frac{c}{N}\sum_{i=1}^{N} z_i' B' B z_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_i' B' A \Delta z_i - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Delta z_i' \Phi_p \Delta z_i = (I) + (II) + (III).$$

The limiting distribution of the above statistic is derived by taking limits of (I), (II) and (III), for $N \to \infty$. To derive the limit of (I), we will employ (7). Then, (I) can be written as $\frac{c}{N}\sum_{i=1}^{N} z_i' B' B z_i = \frac{c}{N}\sum_{i=1}^{N} \varphi_N^2 z_{i-1}' B' B z_{i-1} + \varphi_N z_{i-1}' B' B X \zeta_i + \varphi_N z_{i-1}' B' B u_i + \varphi_N \zeta_i' X' B' B z_{i-1} + \zeta_i' X' B' B X \zeta_i + \zeta_i' X' B' B u_i + \varphi_N u_i' B' B z_{i-1} + u_i' B B' X \zeta_i + u_i' B' B u_i$. Using (8) and (10) and (11), the first term of the last relationship can be written as $\frac{c}{N}\sum_{i=1}^{N} \varphi_N^2 z_{i-1}' B' B z_{i-1} = \frac{c}{N}\sum_{i=1}^{N} z_{i-1}' B' B z_{i-1} + o_p(1) = \frac{c}{N}\sum_{i=1}^{N} (\beta_i e_2' X' \Lambda' + u_i' \Lambda') B' B (\Lambda X e_2 \beta_i + \Lambda u_i) + o_p(1)$. Since the sum is multiplied by $\frac{1}{N}$, any summand coming from the expansion of it which is also multiplied by $\frac{1}{N}$, or $\frac{1}{\sqrt{N}}$, will be asymptotically negligible, $o_p(1)$. By CWLLN and standard results on quadratic forms (see Schott (1996)), we can show that $\frac{c}{N}\sum_{i=1}^{N} (\beta_i e_2' X' \Lambda' + u_i' \Lambda') B' B (\Lambda X e_2 \beta_i + \Lambda u_i) \xrightarrow{p} c\left[\beta^2 e_2' X' \Lambda' B' B \Lambda X e_2 + tr(\Lambda' B' B \Lambda \Gamma)\right]$. Following analogous arguments to the above, it can be shown that

$$(I) : \frac{c}{N}\sum_{i=1}^{N} z_i' B' B z_i \xrightarrow{p} c \begin{bmatrix} tr(\Lambda' B' B \Lambda \Gamma) + tr(\Lambda' B' B \Gamma) + tr(B' B \Lambda \Gamma) + tr(B' B \Gamma) \\ +\beta^2 e_2' X' \Lambda' B' B \Lambda X e_2 + \beta^2 e_2' X' \Lambda' B' B X e_2 \\ +\beta^2 e_2' X' B' B \Lambda X e_2 + \beta^2 e_2' X' B' B X e_2 \end{bmatrix} \tag{14}$$

Similarly, we can show

$$(II) : \frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_i' B' A \Delta z_i \xrightarrow{p} N(c\mu_1, V_{(II)}) \tag{15}$$

where $\mu_1 = c \begin{bmatrix} -tr(\Lambda' B' A \Lambda \Gamma) - tr(\Lambda' B' A \Gamma) - tr(B' A \Lambda \Gamma) - tr(F' B' A \Gamma) \\ -\beta^2 e_2' X' \Lambda' B' A \Lambda X e_2 + e_2' X' \Lambda' B' A X \tilde{e} + \\ -\beta^2 e_2' X' B' A \Lambda X e_2 + e_2' X' B' A X \tilde{e} \end{bmatrix} + tr(\Lambda' B' A \Gamma) + tr(B' A \Gamma)$ and

$$(III) : -\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Delta z_i' \Phi_p \Delta z_i \xrightarrow{p} N(c\mu_2, V_{(III)}) \tag{16}$$

where $\mu_2 = c \begin{bmatrix} tr(\Lambda' \Phi_p \Gamma) + tr(\Phi_p \Lambda \Gamma) \\ +\beta^2 e_2' X' \Lambda' \Phi_p X e_2 + \beta^2 e_2' X' \Phi_p \Lambda X e_2 \\ -e_2' X' \Phi_p X \tilde{e} - \tilde{e}' X' \Phi_p X e_2 \end{bmatrix} - tr(\Phi_p \Gamma)$. Summing up the results in (14), (15) and (16) and using the results of equations (12), we can prove the result of

Theorem 2 because

$$\sqrt{N}\hat{\delta}\left(\hat{\varphi}_{FOD} - \varphi_N - \frac{\hat{b}}{\hat{\delta}}\right) \xrightarrow{d} N(-c(-(I) - (II) - (III)), V)$$

$$\sqrt{N}\hat{\delta}\left(\hat{\varphi}_{FOD} - 1 - \frac{\hat{b}}{\hat{\delta}}\right) \xrightarrow{d} N(-c(-\mu_1 - \mu_2), V).$$

Note that the variance functions of the limiting distributions of quantities (I) and (II): $V_{(II)}$ and $V_{(III)}$, as well as their covariance do not need to be calculated, given that they are equal to variance V of the test statistic UB_T, under H_0: $c = 0$. This happens because these functions are independent of c (see also Breitung (2000)).