

CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass



<http://www.cass.city.ac.uk/cea/index.html>

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

On the Asymptotic t-test for Large Nonstationary Panel Models

Lorenzo Trapani

CEA@Cass Working Paper Series

WP-CEA-11-2008

On the asymptotic t -test for large nonstationary panel models

Lorenzo Trapani*
Cass Business School

October 7, 2008

Abstract

This paper derives an asymptotic t -test for the long-run average in a heterogeneous nonstationary panel model. The asymptotics of the Least Squares Dummy Variable (LSDV) and of the Pooled-OLS (POLS) estimators for the slope parameter is studied under various circumstances (serial correlation, weak endogeneity, strong cross sectional dependence and mixed non-stationary errors) and a modified estimator of the asymptotic variance is derived. The asymptotic variance is computed up to a simple transformation of the residual and no nuisance parameters need to be estimated. The resulting t -statistics are shown to have a standard normal limiting distribution. Asymptotic tests based on the standardised version of the t -statistic are shown to have good size properties even for n as small as 25.

*Cass Business School, Faculty of Finance, 106 Bunhill Row, London EC1Y 8TZ, Tel.: +44 (0) 207 040 5260; email: L.Trapani@city.ac.uk. I wish to thank Giovanni Urga for carefully reading an earlier version of this work and providing inspiring comments. The usual disclaimer applies.

On the asymptotic t -test for large nonstationary panel models

Abstract. This paper derives an asymptotic t -test for the long-run average in a heterogeneous nonstationary panel model. The asymptotics of the Least Squares Dummy Variable (LSDV) and of the Pooled-OLS (POLS) estimators for the slope parameter is studied under various circumstances (serial correlation, weak endogeneity, strong cross sectional dependence and mixed non-stationary errors) and a modified estimator of the asymptotic variance is derived. The asymptotic variance is computed up to a simple transformation of the residual and no nuisance parameters need to be estimated. The resulting t -statistics are shown to have a standard normal limiting distribution. Asymptotic tests based on the standardised version of the t -statistic are shown to have good size properties even for n as small as 25.

KEYWORDS: Panel data, t -test, Asymptotics, Monte Carlo, Common Factors.

1 Introduction

Consider the heterogeneous panel regression model

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}, \quad (1)$$

where $i = 1, \dots, n$, $t = 1, \dots, T$ and the variables y_{it} and x_{it} are both $I(1)$ for each i . In (1), each unit can be either a cointegration regression (with $u_{it} \sim I(0)$) or a spurious regression (when $u_{it} \sim I(1)$). This situation could e.g. correspond to the case (often found in empirical applications) where u_{it} is truly stationary in accord with some theory, but it is observationally equivalent to an $I(1)$ process due to mis-specification. A comprehensive discussion of some possible circumstances whereby this can arise is in recent contributions by Choi, Hu and Ogaki (2008), Fuertes (2008) and Ng (2008).

In the context of time series analysis, it is well known that spurious regression means lack of any relationship between the variables y_{it} and x_{it} . Conversely, in a large panel framework (where both n and T are large), the seminal contributions by Kao (1999) and Phillips and Moon (1999a) show that spurious regression does not imply absence of a long-run relationship, and the long run average parameter $E(\beta_i) = \beta$ can still be estimated consistently at a rate $O_p(n^{-1/2})$. However, asymptotics and inference differ from the time series case; particularly, Kao (1999) shows that t -tests based on estimates of β diverge as $T \rightarrow \infty$, thereby leading to null rejection frequency asymptotically equal to 1. Thus, care should be used when carrying out significance tests using estimates from a nonstationary panel, and it is crucial to implement a correct standardization of the estimated β in order to conduct hypothesis testing.

In a recent contribution, Fuertes (2008) studied the behaviour of t -statistics for the long-run average parameter β under the null $H_0 : \beta = \beta_0$. Using a comprehensive Monte Carlo exercise, Fuertes analysed the amount of size distortion when using the traditional formula

$$t_\beta = \frac{\hat{\beta} - \beta_0}{se(\hat{\beta})}, \quad (2)$$

where $\hat{\beta}$ is either the Least Squares Dummy Variable (LSDV) or the Pooled-OLS (POLS) estimator of β in

$$y_{it} = \alpha_i + \beta x_{it} + v_{it}, \quad (3)$$

with $v_{it} = u_{it} + (\beta_i - \beta) x_{it}$ and

$$se(\hat{\beta}) = \sqrt{\frac{1}{nT} \frac{\sum_{i=1}^n \sum_{t=1}^T \hat{v}_{it}^2}{\sum_{i=1}^n \sum_{t=1}^T x_{it}^2}}, \quad (4)$$

where \hat{v}_{it} is the regression residual. The null rejection frequency is shown to be very high (around 70% for $T = 300$, with a nominal level at 5%) and increasing with T , which is a consequence of the theoretical result that $t_\beta = O_p(\sqrt{T})$ - see Theorem 1(c) in Kao (1999). This problem is compounded when the panel is a mixture of spurious and cointegrated regressions, since the asymptotic distribution of $T^{-1/2}t_\beta$ depends upon various nuisance parameters including the proportion of spurious regressions in (1). As we show below, this is a consequence of $se(\hat{\beta})$ being an inconsistent estimator for the standard error of $\hat{\beta}$. In order to overcome this issue and make the t -test usable, Fuertes (2008) proposes sieve bootstrap. Although the theory of bootstrap for cross-sectionally dependent panels is still to be fully explored, and bootstrap may be computationally burdensome, the simulations in Fuertes (2008) show that when applying an appropriate bootstrap scheme the size distortion is reduced. However, as $T \rightarrow \infty$, t_β diverges and thus even the bootstrap becomes inapplicable.

This note is an application of recent theoretical advances in the asymptotics for panel models with strong cross-sectional dependence derived by Kao, Trapani and Urga (2008), and it is aimed at complementing the results discussed by Fuertes (2008) by deriving an asymptotic version of the t -test for the long run parameter β . The main contribution of this paper is to propose and study a consistent estimator of the standard error of the LSDV and POLS estimators for β , defined respectively as $\tilde{se}(\hat{\beta}^{POLs})$ and $\tilde{se}(\hat{\beta}^{LSDV})$. Thus, $\hat{\beta} - \beta_0$ can be properly studentized by $\tilde{se}(\cdot)$, thereby obtaining a nuisance free statistic, whose limiting distribution is standard normal. This result is asymptotic and it could be viewed as both an alternative or a complement to bootstrap, in light of its computational convenience and due to the fact that, as $(n, T) \rightarrow \infty$, the test still has a standard normal distribution, and thus it does not become invalid.

The derivation of $\tilde{se}(\cdot)$ and of the resulting asymptotic test is done by studying the asymptotics of $\hat{\beta}$ under the null and on studentising $\hat{\beta} - \beta_0$ by its asymptotic standard deviation, taking an approach similar to the one in Bai and Kao (2006). The modified t -statistic $\tilde{t}_\beta = (\hat{\beta} - \beta_0) / \tilde{se}(\hat{\beta})$ is shown to be $N(0, 1)$ as $(n, T) \rightarrow \infty$ under serial dependence, in presence of strong cross sectional dependence arising from the presence of a factor-loading structure in the error term u_{it} and for arbitrary values of the proportion of spurious regres-

sions in the panel which is allowed to range in the interval $[0, 1]$. Thus, the test is shown to be robust with respect to all the cases considered in Fuertes (2008).

Tests based on \tilde{t}_β are a useful complement of sieve bootstrap algorithms for at least three reasons.

First, the test has a standard normal limiting distribution under a very comprehensive set of assumptions. Thus, it can be used with no need for pre-testing. Particularly, unit root tests do not need to be applied to the residuals \hat{u}_{it} in order to detect which units are spurious regressions; moreover, the estimation of the fraction of spurious regressions is not required either (see also Ng, 2008, and Trapani, 2008). This is an important advantage since carrying out unit root test for the residuals in the individual regressions in (1) may lead to wrong inference due to the low power of unit root tests. Note that under some resampling schemes, the bootstrap does require some pretesting to identify which units are cointegrated and which ones are spurious regressions under some schemes (e.g. the pretesting sieve bootstrap, discussed in Fuertes, 2008). From a technical viewpoint, our results are derived for $(n, T) \rightarrow \infty$ jointly (see Phillips and Moon, 1999a, for the definition of joint versus sequential asymptotics), under the restriction $n/T \rightarrow 0$. The literature has so far considered only sequential limits, for $T \rightarrow \infty$ followed by $n \rightarrow \infty$. Whilst sequential limits entail that T must be "much bigger" than n , the panels that can be considered in this framework are those with large T and n of moderate size - see also the discussion in Phillips and Moon (1999a, p. 1059).

Second, the test statistic is asymptotically bounded (unlike t_β) and its limiting distribution is free from nuisance parameters; therefore, the accuracy of the bootstrap algorithm when applied to \tilde{t}_β instead of t_β directly should be enhanced, also allowing for asymptotic refinements. Third, the test is computationally very convenient and it can be implemented in any software that performs panel estimation, such as Eviews, OxMetrics or Stata.

The structure of this note is as follows. Section 2 lays out the assumptions and discusses the asymptotics; numerical evidence from synthetic data is reported in Section 3. Section 4 concludes. The notation employed throughout the paper is fairly standard: standard Brownian motions are denoted as $W(r)$ for $r \in [0, 1]$; integrals involving Brownian motions such as e.g., $\int_0^1 W(s)ds$ are referred to as $\int W$ when there is no ambiguity over limits; $\|\cdot\|$ denotes the Euclidean norm, \xrightarrow{d} convergence in distribution, \xrightarrow{p} convergence in probability and $[\cdot]$ the integer part; the symbol " $\bar{\cdot}$ ", when applied to a time series, denotes demeaning; for a generic Brownian motion B , we define $\bar{B} = B - \int B$ to denote the demeaned version of B .

2 Assumption and main results

Recall the heterogeneous panel model (1)

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it},$$

and the pooled regression (3)

$$y_{it} = \alpha_i + \beta x_{it} + v_{it}.$$

The regressor x_{it} is assumed to have the following DGP

$$x_{it} = x_{it-1} + e_{it}^x,$$

and we let $v_{it} = u_{it} + (\beta_i - \beta) x_{it}$. Strong cross-sectional dependence is assumed to be introduced by considering a factor-loading specification for the error term u_{it} :

$$u_{it} = \left[\varphi_i F_t^{(\lambda)} + e_{it}^{u(\lambda)} \right] d_\lambda + \left[\varphi_i F_t^{(1-\lambda)} + e_{it}^{u(1-\lambda)} \right] (1 - d_\lambda), \quad (5)$$

where we define $d_\lambda = 1$ for $i = 1, \dots, \lfloor n\lambda \rfloor$ and zero otherwise. We assume that $\{F_t^{(\lambda)}, e_{it}^{u(\lambda)}\}$ is non-stationary and $\{F_t^{(1-\lambda)}, e_{it}^{u(1-\lambda)}\}$ is stationary, i.e.

$$\begin{aligned} F_t^{(\lambda)} &= F_{t-1}^{(\lambda)} + \varepsilon_t^F, \\ e_{it}^{u(\lambda)} &= e_{it-1}^{u(\lambda)} + \varepsilon_{it}^e, \end{aligned}$$

and

$$\begin{aligned} F_t^{(1-\lambda)} &= \rho_i^F F_{t-1}^{(1-\lambda)} + \varepsilon_t^F, \\ e_{it}^{u(1-\lambda)} &= \rho_i^e e_{it-1}^{u(1-\lambda)} + \varepsilon_{it}^e, \end{aligned}$$

with $|\rho_i^F| < 1$ and $|\rho_i^e| < 1$; $\{\varepsilon_t^F, \varepsilon_{it}^e\}$ is assumed to be a linear stationary process in both cases.

Consider the following assumptions.

Assumption 1: [time series properties of u_{it}] (i) for all i it holds that $|\varphi_i|^{2+\delta} < \infty$ for some $\delta > 0$; (ii) letting $\omega_{it} = [e_{it}^x, \varepsilon_t^F, \varepsilon_{it}^e]'$, it holds that ω_{it} is a linear stationary process with $E \|\omega_{it}\|^{2+\delta} < \infty$ for some $\delta > 0$ and an invariance principle holds for the partial sums

of ω_{it}^c such that for all $r \in [0, 1]$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Tr \rfloor} \omega_{it} \xrightarrow{d} B_{c,i}(r),$$

where $B_i(r)$ is a vector Brownian motion with covariance matrix

$$\Omega_i = \begin{bmatrix} \sigma_x^2 & \cdot & \cdot \\ 0 & \sigma_F^2 & \cdot \\ 0 & 0 & \sigma_u^2 \end{bmatrix}.$$

Assumption 2: [*heterogeneous coefficients*] (i) for all i , the coefficients α_i and β_i are *i.i.d.* with $E(\alpha_i) = \alpha$, $Var(\alpha_i) = \sigma_\alpha^2 < \infty$ and $E(\beta_i) = \beta$, $Var(\beta_i) = \sigma_\beta^2 < \infty$ and $E|\beta_i|^{2+\delta} < \infty$ for some $\delta > 0$; (ii) $\{\alpha_i\}$, $\{\beta_i\}$ and $\{x_{it}, u_{it}\}$ are three mutually independent groups.

The setup considered here allows for a mixed panel, where $\lfloor n\lambda \rfloor$ units are spurious regressions and the rest of the units are cointegration relationships. Note that allowing for $\lambda \in [0, 1]$ means that the boundary cases whereby all units are cointegration/spurious relationships can be accommodated within this framework. Thus, we entertain the cases that (a) all units are cointegrated, (b) all units are spurious regressions and (c) the panel is a mixture of cointegrating and spurious regressions (mixed panel). Hence, the tests discussed below are robust to the boundary cases of cointegration or spurious regression across all units as well as the case of a panel with mixed $I(0)$ and $I(1)$ error terms. Note that model (3) represents a spurious regression for all units i since the error term is always $I(1)$ because it contains $(\beta_i - \beta) \bar{x}_{it} \sim I(1)$ for all i - see also Phillips and Moon (1999a, p. 1080) - and also possibly because $u_{it} \sim I(1)$ for some i .

Assumption 1 considers a broadly general specification for the time series properties of panel y_{it} . Time dependence is assumed through a linear process - this requirement is needed in order to apply the method of proof proposed in Phillips and Solo (1992), but more general forms of time dependence could be allowed, as long as an invariance principle holds for the partial sums of ω_{it} . Cross sectional dependence is allowed for by using a factor-loading specification for the error term (equation (5)) - we refer e.g. to Bai (2003) and Pesaran (2006) for the stationary case and Bai (2004) for the nonstationary case. No assumption is needed on the loadings φ_i or on the factors $F_t^{(\cdot)}$ apart from the validity of the invariance principle and the condition $|\varphi_i|^{2+\delta} < \infty$. The latter is needed in order for the variance of the random variable $\varphi_i F_t^{(\cdot)}$ to be finite for all t . In this respect, it could alternatively be

assumed that the loadings φ_i be random and independent of the other random variables in the model, as long as some Lindeberg condition holds. Note that in this context neither φ_i nor $F_t^{(\cdot)}$ need to be identified or estimated, as they will be treated as nuisance parameters. Correlation between ε_t^F and ε_{it}^e is assumed to be zero; while this is commonly assumed in panel factor analysis (see e.g. Assumption D in Bai, 2004), this requirement is not strictly needed here and it is assumed only for the purpose of a simple notation. Note that the long run covariance matrix of B_i is assumed to be homogeneous across units. This is not necessary, and all the results derived hereafter could be obtained considering different long run covariance structures (under some mild summability conditions): thus this assumption is done merely for the sake of the notation and without loss of generality. All results are derived assuming only one regressor in (1) and only one factor in (5). This is done for the sake of notational simplicity, and extensions to the multivariate case can be done at the price of a more complicated algebra but essentially under the same assumptions.

Assumption 2 is needed only in order for the CLT and the LLN to hold for the α_i s and the β_i s such that e.g. $n^{-1} \sum_{i=1}^n (\beta_i - \beta)^2 \xrightarrow{p} \sigma_\beta^2$ and $(n\sigma_\beta^2)^{-1/2} \sum_{i=1}^n (\beta_i - \beta) \xrightarrow{d} N(0, 1)$; less strict assumptions could be considered as long as the CLT and the LLN hold in the weak form.

Consider the LSDV and the POLS estimators respectively for β in (3),

$$\begin{aligned}\hat{\beta}^{LSDV} &= \left[\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2 \right]^{-1} \left[\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it} \bar{y}_{it} \right], \\ \hat{\beta}^{POLS} &= \left[\sum_{i=1}^n \sum_{t=1}^T \tilde{x}_{it}^2 \right]^{-1} \left[\sum_{i=1}^n \sum_{t=1}^T \tilde{x}_{it} \tilde{y}_{it} \right],\end{aligned}$$

where $\bar{x}_{it} = x_{it} - T^{-1} \sum_t x_{it}$ and $\tilde{x}_{it} = x_{it} - (nT)^{-1} \sum_i \sum_t x_{it}$ and \bar{y}_{it} and \tilde{y}_{it} are defined similarly. The following theorem characterizes the limiting distribution of the two estimators $\hat{\beta}^{LSDV}$ and $\hat{\beta}^{POLS}$ and of their standard errors computed from (4).

Lemma 1 *Let Assumptions 1 and 2 hold. Then, as $(n, T) \rightarrow \infty$ with $n/T \rightarrow 0$, it holds that*

$$\sqrt{n} \left(\hat{\beta}^{LSDV} - \beta \right) \xrightarrow{d} \left[\frac{2}{5} \frac{\lambda \sigma_u^2}{\sigma_x^2} + \frac{4}{5} \sigma_\beta^2 + 6 \frac{\lambda \varphi^2 \sigma_F^2}{\sigma_x^2} D \right]^{1/2} \times Z, \quad (6)$$

$$\frac{1}{\sqrt{nT}} se \left(\hat{\beta}^{LSDV} \right) \xrightarrow{d} \sqrt{\frac{\lambda \sigma_u^2}{\sigma_x^2} + \sigma_\beta^2 + \frac{6 \lambda \varphi^2 \sigma_F^2}{\sigma_x^2} \left(\int \bar{W}_1^2 \right)}, \quad (7)$$

and

$$\sqrt{n} \left(\hat{\beta}^{POLLS} - \beta \right) \xrightarrow{d} \left[2 \frac{\lambda \sigma_u^2}{\sigma_x^2} + \frac{10}{3} \sigma_\beta^2 + 2 \frac{\lambda \sigma_F^2}{\sigma_x^2} D' \right]^{1/2} \times Z \quad (8)$$

$$\frac{1}{\sqrt{nT}} se \left(\hat{\beta}^{POLLS} \right) \xrightarrow{d} \sqrt{\frac{\lambda \sigma_u^2}{\sigma_x^2} + \sigma_\beta^2 + \frac{2\lambda \varphi^2 \sigma_F^2}{\sigma_x^2} \left(\int W_1^2 \right)}, \quad (9)$$

with $Z \sim N(0, 1)$ independent of all the other random variables; D and D' are defined as

$$D = \int_0^1 \int_0^1 \bar{W}(s) \bar{W}(r) \psi(r, s) dr ds,$$

$$\begin{aligned} D' &= \varphi^2 \int_0^1 \int_0^1 W(s) W(r) \min(r, s) dr ds \\ &\quad - \bar{\varphi}^2 \left(\int_0^1 W(r) dr \right) \int_0^1 \int_0^1 [W(r) + \bar{W}(r)] \min(r, s) dr ds, \end{aligned}$$

with $\psi(r, s) = \frac{1}{3} + \min(r, s) - [\frac{1}{2}r^2 + r(1-r)] - [\frac{1}{2}s^2 + s(1-s)]$, $\bar{\varphi} = n^{-1} \sum_i \varphi_i$ and $\varphi^2 = n^{-1} \sum_i \varphi_i^2$. For both estimators, it holds that $t_\beta = O_p(\sqrt{T})$ as $(n, T) \rightarrow \infty$ under $n/T \rightarrow 0$.

Lemma 1 states that $\hat{\beta}$ is estimated consistently at a rate \sqrt{n} . This result is typical in panel spurious regression as shown by Kao (1999) and Phillips and Moon (1999a). The novel result in (6) and (8) is the asymptotic distribution of $\hat{\beta}$ under the broadly general Assumptions 1 and 2. Note that the limiting distribution of $\hat{\beta}$ is mixed normal instead of normal, contrary to what is usually found in large nonstationary panel literature. This is due both to the presence of the common factors $F_t^{(\cdot)}$ across units and the nonstationarity of $F_t^{(\cdot)}$, which makes the asymptotic theory in Phillips and Moon (1999a) not applicable here; see also the discussion in Kao, Trapani and Urga (2008a, 2008b). Equations (7) and (9) show that, for similar reasons, $se(\hat{\beta})$ converges in distribution to a random variable rather than converging to a constant; this arises again from the presence of the $I(1)$ common factors $F_t^{(\cdot)}$ across units.

Equations (6)-(9) show that, as already proved by Kao (1999) under different assumptions, $t_\beta = O_p(\sqrt{T})$ both when computed using the LSDV and the POLS estimator. There-

fore, the ordinary t -test t_β can not be employed due to a twofold reason. First, t_β diverges as $T \rightarrow \infty$ and therefore, as $T \rightarrow \infty$, the size of the test converges in probability to 1; this finding is consistent with the theory derived by Kao (1999) for a more restrictive setting and with the simulations in Fuertes (2008). Second, even when suitably normalised by \sqrt{T} , t_β would not converge to a standard normal because $se(\hat{\beta})$ does not converge to the asymptotic variance of $\hat{\beta}$.

In order to solve this problem, we propose an alternative estimator of the variances of $\hat{\beta}^{LSDV}$ and $\hat{\beta}^{POLS}$.

Let $\hat{v}_{it} = \bar{y}_{it} - \hat{\beta}^{LSDV} \bar{x}_{it}$ and $\hat{v}_{it} = \tilde{y}_{it} - \hat{\beta}^{POLS} \tilde{x}_{it}$ and define

$$\tilde{se}(\hat{\beta}^{LSDV}) = \frac{\sqrt{\sum_{i=1}^n \left(\sum_{t=1}^T \bar{x}_{it} \hat{v}_{it} \right)^2}}{\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2}, \quad (10)$$

$$\tilde{se}(\hat{\beta}^{POLS}) = \frac{\sqrt{\sum_{i=1}^n \left(\sum_{t=1}^T \tilde{x}_{it} \hat{v}_{it} \right)^2}}{\sum_{i=1}^n \sum_{t=1}^T \tilde{x}_{it}^2}. \quad (11)$$

Under $H_0 : \beta = \beta_0$, the modified t -statistic is defined as

$$\tilde{t}_\beta^{LSDV} = \frac{\hat{\beta}^{LSDV} - \beta_0}{\tilde{se}(\hat{\beta}^{LSDV})},$$

and \tilde{t}_β^{POLS} can be defined similarly using $\hat{\beta}^{POLS}$. The following theorem characterizes the limiting distribution of \tilde{t}_β under H_0 and under local alternatives $H_A^{(n)} : \beta = \beta_0 + c/\sqrt{n}$.

Theorem 1 *Let Assumptions 1 and 2 hold. Then as $(n, T) \rightarrow \infty$ with $n/T \rightarrow 0$,*

$$\tilde{t}_\beta^{LSDV} \xrightarrow[H_0]{d} N(0, 1), \quad (12)$$

$$\tilde{t}_\beta^{LSDV} \xrightarrow[H_A^{(n)}]{d} N(c, 1). \quad (13)$$

The same results hold for \tilde{t}_β^{POLS} .

Theorem 1 shows the distribution of the modified t -statistics \tilde{t}_β^{LSDV} and \tilde{t}_β^{POLS} under the null and under local alternatives. The studentization in (10) ensures that \tilde{t}_β has a standard normal distribution as $(n, T) \rightarrow \infty$, under the setup laid out in Assumptions 1 and 2.

Since Lemma 1 shows that $\hat{\beta}$ has a mixed Gaussian distribution, appropriate studentization ensures standard normality - see also the discussion in Andrews (2005). This is achieved with no need to estimate nuisance parameters such as the fraction of error terms u_{it} that are $I(1)$, the degree of heterogeneity σ_β^2 or φ_i and $F_t^{(\cdot)}$ in (5). The statistic is robust under various circumstances, e.g. when $\lambda = 0$ or 1 , and regardless of the presence and the extent of cross-sectional dependence. Thus, no pretesting is required prior to implementing tests based on \tilde{t}_β . Note that since the limiting distribution of \tilde{t}_β is nuisance free, bootstrapping \tilde{t}_β could lead to asymptotic refinements as opposed to bootstrapping t_β which does depend upon nuisance parameters as shown in Lemma 1. Finally, note that Theorem 1 holds for the joint limit case, provided that $n/T \rightarrow 0$. The practical consequence of this restriction is that the number of units n cannot be "too large" with respect to T .

The test is shown to be powerful versus local alternatives of magnitude $O_p(1/\sqrt{n})$. This result is not surprising and, as shown in the proof, it follows from the consistency of $\sqrt{n}\tilde{s}\tilde{e}(\hat{\beta})$ as an estimator of $Var(\hat{\beta})$ under $H_A^{(n)}$.

3 Monte Carlo evidence

The null rejection frequencies of tests based on \tilde{t}_β^{LSDV} are reported in Tables 1 and 2¹. The Monte Carlo exercises have been conducted with 2000 replications. When assessing the size of the test at a confidence level c_α , this entails a standard error given by $\sqrt{c_\alpha(1-c_\alpha)/2000}$; for a confidence level $c_\alpha = 0.05$, this means that the reported rejection frequencies lie within a 95% confidence interval of width ± 0.01 .

The experiments reported here were carried out for $(n, T) = (25, 100)$. Other experiments with different combinations of (n, T) showed that (1) different values of T do not alter the size of the test, and (2) size distortion does arise for $n < 25$. Thus, $n = 25$ seems to be the threshold below which the asymptotic approximation of \tilde{t}_β is no longer accurate.

Synthetic data are generated using equation (1)

$$y_{it} = \alpha_i + \beta_i x_{it} + \left[\varphi_i F_t^{(\lambda)} + e_{it}^{u(\lambda)} \right] d_\lambda + \left[\varphi_i F_t^{(1-\lambda)} + e_{it}^{u(1-\lambda)} \right] (1 - d_\lambda),$$

with $x_{it} = x_{it-1} + e_{it}^x$, $\alpha_i \sim iidN(1, 1)$, $\beta_i \sim iidN(1, \sigma_\beta^2)$ and $\varphi_i \sim iidN(\bar{\varphi}, \sigma_\varphi)$. We consider the cases $\sigma_\beta^2 = \{0, 1, 2\}$ and $(\bar{\varphi}, \sigma_\varphi) = \{(0, 0), (1, 1)\}$. The fraction of spurious regressions, λ , is set as $\lambda \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$.

¹Results for \tilde{t}_β^{POLs} showed a similar pattern and thus were not reported.

The innovations $e_t = [e_{it}^x, \varepsilon_t^F, e_{it}^u]'$ are generated by creating a Gaussian *i.i.d.* sequence $\dot{e}_t = [\dot{e}_{it}^x, \dot{e}_t^F, \dot{e}_{it}^u]'$ with $Var(\dot{e}_{it}^x) = \sigma_x^2$ and $Var(\dot{e}_t^F) = Var(\dot{e}_{it}^u) = 1$; the signal to noise ratio is $\sigma_x = 0.5$. Time dependence is induced as

$$e_t = \rho_e e_{t-1} + \dot{e}_t + \theta_e \dot{e}_{t-1},$$

with $(\rho_e, \theta_e) \in \{(0, 0), (0.75, 0), (0, 0.75), (0, -0.75)\}$.

[Insert Tables 1 and 2 somewhere here]

Tables 1 and 2 report the null rejection frequencies for the cases $(\bar{\varphi}, \sigma_\varphi) = (0, 0)$ and $(1, 1)$ - thereby considering absence and presence of cross-dependence respectively. The following stylized facts emerge:

- the presence of *strong cross-sectional dependence* has a tendency to worsen the results. Although a factor structure in the error term alters only slightly the numbers in Table 2 (when compared to those in Table 1), including cross-dependence makes the empirical rejection frequencies less close to the nominal level 0.05. This is particularly evident as λ increases, which is expected since the asymptotics are driven by those units with $I(1)$ error and therefore $I(1)$ common factors in their error term. A possible explanation is that presence of cross sectional dependence makes the convergence of the (cross-sectional) CLT slower, thereby marring the accuracy of asymptotic approximations;
- the role played by *the fraction of spurious regressions*, λ , is evident in both cases, with or without cross dependence. Particularly, *ceteris paribus*, as λ increases, the empirical rejection frequencies increase as well, with a tendency towards oversizement. It is worth noticing that whilst for $\lambda < 0.5$ the observed rejection frequencies are usually not significantly different to the nominal level, for $\lambda > 0.5$ there is evidence of significant (although slight) oversizement;
- the *degree of heterogeneity* σ_β^2 does not seem to affect size results in a regular way. Although rejection frequencies do change, and this is particularly evident as λ grows, there is no clear pattern across the various results in the tables. Note that for $(\lambda, \sigma_\beta^2) = (0, 0)$ there is a tendency towards undersizement; this can be explained in light of equations (6) and (8), from which it is evident that in the case of a cointegrated and

homogeneous panel, $t_\beta = O_p(1)$ and thus $\tilde{t}_\beta = O_p(T^{-1/2})$. As a consequence, when T grows, \tilde{t}_β tends to zero, whence the small size of the test;

- the presence of *serial dependence* in the errors does not play an important role in affecting the size, at least apart from the case of a negative MA root. This is in accord with the way in which the statistic is computed, i.e. by studentising. Whilst the cases $(\rho_e, \theta_e) = (0.75, 0)$ and $(\rho_e, \theta_e) = (0, 0.75)$ do not differ substantially from the white noise case, negative MA root seems to lead to undersizement. This is fairly common in the literature - see e.g. Leybourne and Newbold (1999) and the references cited therein.

4 Conclusions

This note proposes an asymptotic t -test for the long-run average parameter β in mixed panels. Standardization using the estimated standard deviation leads to t -test that are asymptotically standard normal. Thus, as $(n, T) \rightarrow \infty$, t -tests based on this standardization are free from nuisance parameters, and can therefore be employed using the critical values from the standard normal distribution. From the point of view of the features of the model, the test is robust to any degree of parameter heterogeneity, to any level of serial and cross-sectional dependence and to any values of the fraction of spurious regressions in the sample. Due to the broad generality of the assumptions under which the asymptotic t -test can be computed, the test can be applied without any pretesting being necessary.

Thus, the contribution of this paper can be seen as both a substitute and a complement to the bootstrap, as discussed in Fuertes (2008). The asymptotic t -test is a substitute to the bootstrap and it has the advantage of being computationally straightforward. Monte Carlo evidence shows that the size distortion is small when n is as small as 25, and it also reinforces the conclusion that the test is robust to various values of parameters heterogeneity, fraction of spurious regression, and presence of serial and cross dependence. The asymptotic t -test is also a complement to sieve bootstrap, in that the test statistic \tilde{t}_β is asymptotically pivotal, and thus the bootstrap version of \tilde{t}_β^{LSDV} and \tilde{t}_β^{POLS} could yield asymptotic refinements.

References

- [1] Andrews, D.W.K. (2005), "Cross-Section Regression with Common Shocks," *Econometrica*, 73, 1551-1586.
- [2] Bai, J. (2003), "Inferential Theory for Structural Models of Large Dimensions", *Econometrica*, 71, 135-171.
- [3] Bai, J. (2004), "Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data," *Journal of Econometrics*, 122, 137-183.
- [4] Bai, J., Kao, C. (2006), "On the Estimation and Inference of a Panel Cointegration Model with Cross-Sectional Dependence", in *Panel Data Econometrics: Theoretical Contributions and Empirical Applications*, edited by Badi H. Baltagi.
- [5] Bai, J., and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191-221.
- [6] Bai, J., and Ng, S. (2004), "A Panic Attack on Unit Roots and Cointegration", *Econometrica*, 72, 1127-1178.
- [7] Baltagi, B.H., Kao, C. and Liu, (2008), "Asymptotic Properties of Estimators for the Linear Panel Regression Model with Individual Effects and Serially Correlated Errors: The Case of Stationary and Non-Stationary Regressors and Residuals", forthcoming at *The Econometrics Journal*.
- [8] Choi, I., Hu, L., Ogaki, M. (2008), "Robust Estimation for Structural Spurious Regressions and A Hausman-type Cointegration Test", *Journal of Econometrics*, 142, 327-351.
- [9] Fuertes, A.-M. (2008), "Sieve Bootstrap t -test on Long-Run Average Parameters", *Computational Statistics and Data Analysis*, 52, 3354-70.
- [10] Hall, P., Heyde, C. C. (1980), *Martingale Limit Theory and Its Applications*, New York: Academic Press.
- [11] Kao, C. (1999), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data", *Journal of Econometrics*, 90, 1-44.

- [12] Kao, C., Trapani, L. and Urga, G. (2008a), "Asymptotics for Panel Models with Common Shocks", WP-CEA-01-2006, Centre for Econometric Analysis, Cass Business School.
- [13] Kao, C., Trapani, L. and Urga, G. (2008b), "Modelling and Testing for Structural Changes in Panel Cointegration Models with Common and Idiosyncratic Stochastic Trends", mimeo, Cass Business School.
- [14] Kapetanios, G., Pesaran, M.H., Yamagata, T. (2006), "Panels with Nonstationary Multifactor Error Structures", mimeo.
- [15] Leybourne, S., Newbold, P. (1999), "On the Size Properties of Phillips-Perron Tests", *Journal of Time Series Analysis*, 20, 51-61.
- [16] Ng, S. (2008), "A Simple Test for Non-Stationarity in Mixed Panels", forthcoming in *Journal of Business and Economics Statistics*.
- [17] Pesaran, M. H. (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", *Econometrica*, 74, 967-1012.
- [18] Phillips, P. C. B., and Moon, H. R. (1999a), "Linear Regression Limit Theory for Nonstationary Panel Data", Yale University, mimeographed.
- [19] Phillips, P. C. B., and Moon, H. R. (1999b), "Linear Regression Limit Theory for Nonstationary Panel Data" *Econometrica*, 67, 1057-1112.
- [20] Phillips, P. C. B., and Solo, V. (1992), "Asymptotics for linear processes", *Annals of Statistics*, 20, 971–1001.
- [21] Trapani, L. (2008), "Two \sqrt{n} -Consistent Estimators for Mixed Heterogeneous Panels", mimeo.

Appendix: proofs and derivations

Proof of Lemma 1. We shall firstly prove (6) and (7); the method of proof builds on the theory developed by Kao, Trapani and Urga (2008b). Henceforth, we will define $F_t = F_t^{(\lambda)} d_\lambda + F_t^{(1-\lambda)} (1 - d_\lambda)$ and $e_{it}^u = e_{it}^{u(\lambda)} d_\lambda + e_{it}^{u(1-\lambda)} (1 - d_\lambda)$.

Let us firstly consider the LSDV estimator. The estimation error is given by

$$\hat{\beta}^{LSDV} - \beta = \frac{\sum_{i=1}^n \sum_{t=1}^T [(\beta_i - \beta) \bar{x}_{it} + \bar{u}_{it}] \bar{x}_{it}}{\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2}.$$

Kao (1999) proved that

$$\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2 \xrightarrow{p} \frac{1}{6} \sigma_x^2, \quad (14)$$

and this result holds for all values of λ since it only depends on the features of \bar{x}_{it} . As far as the numerator of $\hat{\beta}^{LSDV} - \beta$ is concerned, let

$$\begin{aligned} \xi_{iT} &= \frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T [(\beta_i - \beta) \bar{x}_{it} + \bar{u}_{it}] \bar{x}_{it} \\ &= \frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) \bar{x}_{it}^2 + \frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u + \frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T \varphi_i \bar{F}_t \bar{x}_{it}, \end{aligned}$$

and let C be the σ -field generated by the F_t s. Then, conditional on C , the sequence ξ_{iT} is *i.i.d.* across i and

$$\begin{aligned} E(\xi_{iT} | C) &= E \left[\frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) \bar{x}_{it}^2 \right] + E \left[\frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right] + E \left[\frac{1}{T^2} \sum_{i=1}^n \sum_{t=1}^T \varphi_i \bar{F}_t \bar{x}_{it} \right] \\ &= I + II + III. \end{aligned}$$

Since, by Assumption 2, β_i is independent of \bar{x}_{it}^2 and $E(\beta_i) = \beta$, it follows that $I = 0$ for all T . As far as II is concerned, it follows from Phillips and Moon (1999b, p. 1101, eq. 8.17) that $II = O_p(T^{-1/2})$. Similar arguments would lead to $III = O_p(T^{-1/2})$. Thus, $E(\xi_{iT} | C) = O_p(T^{-1/2})$ for all i . Let $\bar{\xi}_{iT} = \xi_{iT} - E(\xi_{iT})$ and let I_i be the σ -field generated by C and $\{\xi_{1T}, \dots, \xi_{iT}\}$. Then $\{\bar{\xi}_{iT}, I_i; i \geq 1\}$ is a martingale difference sequence (MDS). As $T \rightarrow \infty$, a conditional Liapunov condition holds whereby $E|\bar{\xi}_{iT} | C|^{2+\delta} < \infty$ for all i (this follows from $\bar{\xi}_{iT}$ being *i.i.d.* across i conditional on C). This can be verified by noting that,

for some constant M_δ which only depends on δ

$$\begin{aligned} |\bar{\xi}_{iT}|^{2+\delta} &\leq M_\delta \left[|\beta_i - \beta|^{2+\delta} \left| \frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it}^2 \right|^{2+\delta} + \left| \frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right|^{2+\delta} + |\varphi_i|^{2+\delta} \left| \frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right|^{2+\delta} \right] \\ &= I + II + III. \end{aligned}$$

Assumption 2 ensures that $E|\beta_i - \beta|^{2+\delta}$ and $E|\varphi_i|^{2+\delta}$ are finite. Also, defining \bar{W}_{1i} , \bar{W}_{2i} and \bar{W}_3 as independent standard demeaned Brownian motions, the Continuous Mapping Theorem (CMT) ensures that as $T \rightarrow \infty$, $I \xrightarrow{d} \sigma_x^{2(2+\delta)} \left| \int \bar{W}_{1i}^2 \right|^{2+\delta}$, $II \xrightarrow{d} \sigma_e^{2+\delta} \sigma_x^{2+\delta} \left| \int \bar{W}_{1i} \bar{W}_{2i} \right|^{2+\delta}$ and $III \xrightarrow{d} \sigma_F^{2+\delta} \sigma_x^{2+\delta} \left| \int \bar{W}_{1i} \bar{W}_3 \right|^{2+\delta}$; all these quantities have finite expectation. Since $E|\bar{\xi}_{iT}|^C |C|^{2+\delta} < \infty$ as $T \rightarrow \infty$, it is possible to apply a Central Limit Theorem for Martingale Difference Sequences; as $(n, T) \rightarrow \infty$ with $n/T \rightarrow 0$

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_{iT} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{\xi}_{iT} + \frac{1}{\sqrt{n}} \sum_{i=1}^n E(\xi_{iT}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{\xi}_{iT} + O_p\left(\sqrt{\frac{n}{T}}\right) \\ &\xrightarrow{d} \left[\frac{1}{n} \sum_{i=1}^n E(\bar{\xi}_{iT}^2 | C) \right]^{1/2} \times Z, \end{aligned}$$

where $Z \sim N(0, 1)$ and independent of $E(\bar{\xi}_{iT}^2 | C)$. Note that $E(\bar{\xi}_{iT}^2 | C) = E(\xi_{iT}^2 | C) + o_p(1)$ and consider $E(\xi_{iT}^2 | C)$; it holds

$$\begin{aligned} \xi_{iT}^2 &= (\beta_i - \beta)^2 \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it}^2 \right)^2 + \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right)^2 + \varphi_i^2 \left(\frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right)^2 \\ &\quad + 2(\beta_i - \beta) \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it}^2 \right) \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right) + 2(\beta_i - \beta) \varphi_i \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it}^2 \right) \left(\frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right) \\ &\quad + 2\varphi_i \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right) \left(\frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right) \\ &= a + b + c + d + e + f. \end{aligned} \tag{15}$$

Assumption 2 ensures that $E(d) = E(e) = 0$. In light of Assumption 1, $a \xrightarrow{d} (\beta_i - \beta)^2 \left(\int \bar{B}_{xi}^2 \right)^2$ as $T \rightarrow \infty$ and $E\left[(\beta_i - \beta)^2 \left(\int \bar{B}_{xi}^2 \right)^2\right] = E(\beta_i - \beta)^2 E\left(\int \bar{B}_{xi}^2\right)^2$ for all units, with $E(\beta_i - \beta)^2 = \sigma_\beta^2$ by definition and $E\left(\int \bar{B}_{xi}^2\right)^2 = \sigma_x^4/45$ - see Kao (1999) for the latter result. Assumption 1 also ensures that $b \xrightarrow{d} \left(\int \bar{B}_{xi} \bar{B}_{ei}\right)^2$ for $i = 1, \dots, [n\lambda]$ and $b = O_p(T^{-1})$ for the other units;

since $\int \bar{B}_{xi}\bar{B}_{ei}$ has zero mean, it holds that $E(\int \bar{B}_{xi}\bar{B}_{ei})^2 = Var(\int \bar{B}_{xi}\bar{B}_{ei}) = \sigma_e^2\sigma_x^2/90$ - see Kao (1999). This also entails $E(f) = 0$ as $T \rightarrow \infty$. Last, considering c

$$E \left[\left(\frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right)^2 \middle| C \right] = \frac{1}{T^4} \sum_{t=1}^T \sum_{s=1}^T \bar{F}_t \bar{F}_s E(\bar{x}_{it}\bar{x}_{is}),$$

and after some algebra it can be proved that

$$\begin{aligned} \frac{E(\bar{x}_{it}\bar{x}_{is})}{\sigma_x^2} &= \min(s, t) + \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \min(s, t) \\ &\quad - \frac{t(t+1)}{2T} - \frac{t(T-t)}{T} - \frac{s(s+1)}{2T} - \frac{s(T-s)}{T} \\ &= T\psi(r, s). \end{aligned}$$

As $T \rightarrow \infty$ we have $T^{-3} \sum_{t=1}^T \sum_{s=1}^T \min(s, t) = 1/3$ and, due to the CMT,

$$\frac{1}{T^4} \sum_{t=1}^T \sum_{s=1}^T \bar{F}_t \bar{F}_s E(\bar{x}_{it}\bar{x}_{is}) \xrightarrow{d} \sigma_x^2 \int_0^1 \int_0^1 \bar{B}_F(r) \bar{B}_F(s) \psi(r, s) dr ds \equiv \sigma_x^2 \sigma_F^2 D,$$

for $i = 1, \dots, \lfloor n\lambda \rfloor$ and $b = O_p(T^{-1})$ for the other units. To calculate $n^{-1} \sum_{i=1}^n E(\xi_{iT}^2 | C)$, note that the only terms that differ across units and do not vanish asymptotically are b and c in (15) and

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right)^2 &\stackrel{a.s.}{=} \frac{\lambda}{n\lambda} \sum_{i=1}^{\lfloor n\lambda \rfloor} \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \right)^2 \xrightarrow{p} \lambda \frac{\sigma_e^2 \sigma_x^2}{90}, \\ \frac{1}{n} \sum_{i=1}^n \varphi_i^2 \left(\frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right)^2 &\stackrel{a.s.}{=} \frac{\lambda}{n\lambda} \sum_{i=1}^{\lfloor n\lambda \rfloor} \varphi_i^2 \left(\frac{1}{T^2} \sum_{t=1}^T \bar{F}_t \bar{x}_{it} \right)^2 \xrightarrow{d} \lambda \varphi^2 \sigma_x^2 \sigma_F^2 D. \end{aligned}$$

Combining these results together we obtain (6).

Now we turn our attention to (7). Recall

$$se(\hat{\beta}) = \sqrt{\frac{1}{nT} \frac{\sum_{i=1}^n \sum_{t=1}^T \hat{v}_{it}^2}{\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2}},$$

and let us focus on the numerator. Since $\hat{v}_{it} = (\beta_i - \hat{\beta}) \bar{x}_{it} + \varphi_i \bar{F}_t + \bar{e}_{it}^u$, it holds that

$$\begin{aligned}
\sum_{i=1}^n \sum_{t=1}^T \hat{v}_{it}^2 &= \sum_{i=1}^n \sum_{t=1}^T \bar{e}_{it}^{u2} + \sum_{i=1}^n \varphi_i^2 \sum_{t=1}^T \bar{F}_t^2 + \sum_{i=1}^n (\beta_i - \beta)^2 \sum_{t=1}^T \bar{x}_{it}^2 \\
&+ (\hat{\beta} - \beta)^2 \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2 + 2 (\hat{\beta} - \beta) \sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) \bar{x}_{it}^2 \\
&+ 2 \sum_{i=1}^n (\beta_i - \beta) \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u + 2 (\hat{\beta} - \beta) \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u \\
&+ 2 \sum_{i=1}^n \varphi_i (\beta_i - \beta) \sum_{t=1}^T \bar{x}_{it} \bar{F}_t + 2 (\hat{\beta} - \beta) \sum_{i=1}^n \varphi_i \sum_{t=1}^T \bar{x}_{it} \bar{F}_t \\
&+ 2 \sum_{i=1}^n \varphi_i \sum_{t=1}^T \bar{F}_t \bar{e}_{it}^u \\
&= I + II + III + IV + V + VI + VII + VIII + IX + X. \tag{16}
\end{aligned}$$

Consider I . Assumption 1 ensures that $I = O_p(nT^2)$; also, in light of assumption 1 we have $\sum_{t=1}^T \bar{F}_t^2 = O_p(T^2)$ and therefore, since $\sum_{i=1}^n \varphi_i^2 = O(n)$, it holds that $II = O_p(nT^2)$. As far as III is concerned, $\sum_{t=1}^T \bar{x}_{it}^2 = O_p(T^2)$ and the sequence $\sum_{i=1}^n (\beta_i - \beta)^2 \sum_{t=1}^T \bar{x}_{it}^2$ is *i.i.d.* across i and has non-zero mean and finite variance; thus, $III = O_p(nT^2)$. Recalling that $\hat{\beta} - \beta = O_p(n^{-1/2})$, we have $IV = O_p(n^{-1}T^2)$. Since $\sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) \bar{x}_{it}^2 = O_p(\sqrt{n}T^2)$, $V = O_p(T^2)$. Also, as far as VI is concerned, $\sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u = O_p(T^2)$ and $\sum_{i=1}^n (\beta_i - \beta) \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u$ is *i.i.d.* across i and zero mean so that a cross-sectional CLT holds; thus $VI = O_p(\sqrt{n}T^2)$. Similar arguments lead to $VII = O_p(nT^2)$. Given that $\varphi_i (\beta_i - \beta) \sum_{t=1}^T \bar{x}_{it} \bar{F}_t$ has mean zero across i and $\sum_{t=1}^T \bar{x}_{it} \bar{F}_t = O_p(T^2)$, and we have $VIII = O_p(\sqrt{n}T^2)$, and $IX = O_p(T^2)$. Last, since $\sum_{t=1}^T \bar{F}_t \bar{e}_{it}^u = O_p(T^2)$ and this is an *i.i.d.* sequence which has zero mean across i , $X = O_p(\sqrt{n}T^2)$. Thus, the terms that dominate are I , II and III , all of order $O_p(nT^2)$. It holds that

$$\begin{aligned}
\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T \bar{e}_{it}^{u2} &\stackrel{a.s.}{=} \frac{\lambda}{(n\lambda)T^2} \sum_{i=1}^{\lfloor n\lambda \rfloor} \sum_{t=1}^T \bar{e}_{it}^{u2} + o_p(1) \xrightarrow{p} \frac{\lambda\sigma_u^2}{6}, \\
\frac{1}{nT^2} \sum_{i=1}^n \varphi_i^2 \sum_{t=1}^T \bar{F}_t^2 &\xrightarrow{d} \lambda\varphi^2\sigma_F^2 \int \bar{W}^2, \\
\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta)^2 \bar{x}_{it}^2 &\xrightarrow{p} \frac{\sigma_\beta^2\sigma_x^2}{6}.
\end{aligned}$$

It follows that

$$\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T \hat{v}_{it}^2 \xrightarrow{d} \frac{\lambda\sigma_u^2}{6} + \frac{\sigma_\beta^2\sigma_x^2}{6} + \varphi^2\sigma_F^2 \int \bar{W}^2,$$

and therefore, since the denominator $\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2/nT^2 \xrightarrow{p} \sigma_x^2/6$, we have

$$se(\hat{\beta}) \xrightarrow{d} \frac{\lambda\sigma_u^2}{\sigma_x^2} + \sigma_\beta^2 + \frac{6\varphi^2\sigma_F^2}{\sigma_x^2} \int \bar{W}^2.$$

Note that no restriction on the rate of expansion between n and T is required here to derive the joint limit.

The proof of (8) and (9) follows similar lines as the proof of (6) and (7), and thus some passages are omitted. In this case the estimation error is

$$\hat{\beta}^{POLs} - \beta = \frac{\sum_{i=1}^n \sum_{t=1}^T \left\{ (\alpha_i - \alpha) + (\beta_i - \beta) x_{it} - (nT)^{-1} \left[\sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) x_{it} \right] + \tilde{u}_{it} \right\} \tilde{x}_{it}}{\sum_{i=1}^n \sum_{t=1}^T \tilde{x}_{it}^2}.$$

Lemma 1.2 in Baltagi, Kao and Liu (2008) shows that $(nT^2)^{-1} \sum_{i=1}^n \sum_{t=1}^T \tilde{x}_{it}^2 \xrightarrow{p} \sigma_x^2/2$. As far as the numerator is concerned, let $(nT)^{-1} \sum_{i=1}^n \sum_{t=1}^T x_{it} = \ddot{x}$, $(nT)^{-1} \sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) x_{it} = \ddot{x}_\beta$ and define

$$\varsigma_{iT} = \frac{1}{T^2} \sum_{t=1}^T [(\beta_i - \beta) x_{it} \tilde{x}_{it} - \ddot{x}_\beta \tilde{x}_{it} + \tilde{u}_{it} \tilde{x}_{it}].$$

Then, conditional on C , ς_{iT} is *i.i.d.*; similar passages as above lead to $E(\varsigma_{iT}|C) = O_p(T^{-1/2})$.

Letting $\bar{\varsigma}_{iT} = \varsigma_{iT} - E(\varsigma_{iT}|C)$, as $(n, T) \rightarrow \infty$ with $n/T \rightarrow 0$ the MDS CLT entails

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \varsigma_{iT} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{\varsigma}_{iT} + \frac{1}{\sqrt{n}} \sum_{i=1}^n E(\varsigma_{iT}|C) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{\varsigma}_{iT} + O_p\left(\sqrt{\frac{n}{T}}\right) \\ &\xrightarrow{d} \left[\frac{1}{n} \sum_{i=1}^n E(\bar{\varsigma}_{iT}^2|C) \right]^{1/2} \times Z, \end{aligned}$$

with $Z \sim N(0, 1)$ independent of $E(\bar{\varsigma}_{iT}^2|C)$. Note also that the term $\sum_{i=1}^n \sum_{t=1}^T (\alpha_i - \alpha) \tilde{x}_{it} = O_p(\sqrt{n}T^{3/2})$ is negligible. In order to prove (8), it is necessary to find the value of $E(\bar{\varsigma}_{iT}^2|C)$.

Note that

$$\begin{aligned}
\varsigma_{iT}^2 &= (\beta_i - \beta)^2 \left(\frac{1}{T^2} \sum_{t=1}^T x_{it} \tilde{x}_{it} \right)^2 + \ddot{x}_\beta^2 \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \right)^2 + \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{u}_{it} \tilde{x}_{it} \right)^2 \\
&\quad + 2(\beta_i - \beta) \ddot{x}_\beta \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \right) \left(\frac{1}{T^2} \sum_{t=1}^T x_{it} \tilde{x}_{it} \right) + 2\ddot{x}_\beta \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \right) \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{u}_{it} \tilde{x}_{it} \right) \\
&\quad + 2(\beta_i - \beta) \left(\frac{1}{T^2} \sum_{t=1}^T x_{it} \tilde{x}_{it} \right) \left(\frac{1}{T^2} \sum_{t=1}^T \tilde{u}_{it} \tilde{x}_{it} \right) \\
&= I + II + III + IV + V + VI.
\end{aligned}$$

Note that $\sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) x_{it} = O_p(\sqrt{n}T^{3/2})$ and therefore $\ddot{x}_\beta = O_p(n^{-1/2}T^{1/2})$. This can be proved noting that (1) the FCLT ensures that $\sum_{t=1}^T x_{it} = O_p(T^{3/2})$, and (2) that the x_{it} s have zero mean uniformly in t since their DGP has no drift term. Also,

$$\sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) x_{it} \tilde{x}_{it} = \sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) x_{it}^2 + \ddot{x} \left[\sum_{i=1}^n \sum_{t=1}^T (\beta_i - \beta) x_{it} \right].$$

Due to similar arguments as above, we have $\sum_{i=1}^n \sum_{t=1}^T x_{it} = O_p(\sqrt{n}T^{3/2})$ and thus $\ddot{x} = O_p(n^{-1/2}T^{1/2})$; therefore, $II = O_p(n^{-1/2}T^{-1/2})$ and $V = O_p(n^{-1/2})$. Assumption 2 ensures that $E(IV) = E(VI) = 0$ for all n and T . As far as I is concerned, note that as $T \rightarrow \infty$, $I \xrightarrow{d} (\beta_i - \beta)^2 (\sigma_x^2 \int W^2)^2$. It holds that $E\left[(\int W^2)^2 \middle| C\right] = E(\int W^2)^2 = \text{Var}(\int W^2) + [E(\int W^2)]^2$. Since $E(\int W^2) = 1/2$ and Lemma 14 in Phillips and Moon (1999a) ensures that $\text{Var}(\int W^2) = 7/12$, we get $E(I) = 5\sigma_\beta^2 \sigma_x^4 / 6$. Considering III , note

$$\frac{1}{T^2} \sum_{t=1}^T \tilde{u}_{it} \tilde{x}_{it} = \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{e}_{it}^u + \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \left(\varphi_i F_t - \frac{1}{nT} \sum_{i=1}^n \varphi_i \sum_{t=1}^T F_t \right) = a + b,$$

which is $O_p(1)$ for $i = 1, \dots, [n\lambda]$ and $o_p(1)$ for the other units. Similarly to above, $\sum_{i=1}^n \sum_{t=1}^T \tilde{x}_{it} \tilde{e}_{it}^u = \sum_{i=1}^n \sum_{t=1}^T x_{it} e_{it}^u - nT \ddot{x} \tilde{e}^u$ with \ddot{x} and \tilde{e}^u both $O_p(n^{-1/2}T^{1/2})$. Thus,

as $T \rightarrow \infty$, $E(a^2) = \lambda \sigma_u^2 \sigma_x^2 / 2$. As far as b is concerned, define $\bar{F} = T^{-1} \sum_{t=1}^T F_t$; then

$$\begin{aligned} & E \left\{ \left[\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} (\varphi_i F_t - \bar{\varphi} \bar{F}) \right]^2 \middle| C \right\} \\ &= \frac{1}{T^4} \sum_{t=1}^T \sum_{s=1}^T (\varphi_i F_t - \bar{\varphi} \bar{F}) (\varphi_i F_s - \bar{\varphi} \bar{F}) E(\tilde{x}_{it} \tilde{x}_{is}). \end{aligned}$$

Noting that $E(\tilde{x}_{it} \tilde{x}_{is}) = \sigma_x^2 \min(t, s)$ and following similar passages as above, it follows that $n^{-1} \sum_{i=1}^n E(\hat{\xi}_{iT}^2 | C) \xrightarrow{d} D'$. Last

$$\begin{aligned} & E \left\{ \left[\frac{1}{T^2} \sum_{t=1}^T x_{it} e_{it}^u \right] \left[\frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} (\varphi_i F_t - \bar{\varphi} \bar{F}) \right] \middle| C \right\} \\ &= \frac{1}{T^4} \sum_{t=1}^T \sum_{s=1}^T [E(x_{it} \tilde{x}_{is} e_{it}^u)] [\varphi_i F_t - \bar{\varphi} \bar{F}] = 0 \end{aligned}$$

since $E(x_{it} \tilde{x}_{is} e_{it}^u) = 0$ for all i, t and s . Combining all the above, (8) follows; (9) follows from similar passages as above, and the proof is thus omitted to save space. ■

Proof of Theorem 1. In order to prove (12), it is sufficient to show that, under H_0 , $\sqrt{n} \tilde{s}e(\hat{\beta}^{LSDV})$ converges to the standard deviation of $\sqrt{n}(\hat{\beta}^{LSDV} - \beta)$. Consider the numerator of (10) and let $\hat{\xi}_{iT} = T^{-2} \sum_{t=1}^T \bar{x}_{it} \hat{v}_{it}$; after some algebra it holds that $\hat{\xi}_{iT} = \xi_{iT} - (\hat{\beta}^{LSDV} - \beta) T^{-2} \sum_{t=1}^T \bar{x}_{it}^2$. Conditioning on C , $\hat{\xi}_{iT}$ is an *i.i.d.* sequence and thus an ergodic theorem ensures that

$$\frac{1}{n} \sum_{i=1}^n \hat{\xi}_{iT}^2 \xrightarrow{p} E \left(\frac{1}{n} \sum_{i=1}^n \hat{\xi}_{iT}^2 \middle| C \right),$$

with

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \hat{\xi}_{iT}^2 &= \frac{1}{n} \sum_{i=1}^n \xi_{iT}^2 + (\hat{\beta}^{LSDV} - \beta)^2 \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it}^2 \right)^2 \\ &\quad - 2(\hat{\beta}^{LSDV} - \beta) \frac{1}{n} \sum_{i=1}^n \xi_{iT} \left(\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it}^2 \right) \\ &= I + II + III. \end{aligned}$$

Since $(\hat{\beta}^{LSDV} - \beta) = O_p(n^{-1/2})$, we have $II = o_p(1)$ and $III = o_p(1)$. Thus,

$$E\left(\frac{1}{n} \sum_{i=1}^n \hat{\xi}_{iT}^2 \middle| C\right) = E\left[\frac{1}{n} \sum_{i=1}^n \xi_{iT}^2 \middle| C\right] + o_p(1),$$

which is the variance of the numerator of $\sqrt{n}(\hat{\beta}^{LSDV} - \beta)$ as shown in Lemma 1.

To prove (13), two results need to be shown under $H_A^{(n)}$. First, that $\sqrt{n}(\hat{\beta}^{LSDV} - \beta_0) \xrightarrow{d} N(c, V^2)$, where V is given in (6). Second, that $\sqrt{n}\tilde{s}e(\hat{\beta}^{LSDV}) \xrightarrow{d} V$. Since under $H_A^{(n)}$: $\beta = \beta_0 + c/\sqrt{n}$

$$\begin{aligned} \hat{\beta}^{LSDV} - \beta_0 &= \left(\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2\right)^{-1} \left(\sum_{i=1}^n \xi_{iT} + \frac{c}{\sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2\right) \\ &= \left(\sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2\right)^{-1} \sum_{i=1}^n \xi_{iT} + \frac{c}{\sqrt{n}}, \end{aligned}$$

and recalling (6) it holds that, under $H_A^{(n)}$, $\sqrt{n}(\hat{\beta}^{LSDV} - \beta_0) \xrightarrow{d} N(0, V^2) + c \sim N(c, V^2)$.

As far as $\sqrt{n}\tilde{s}e(\hat{\beta}^{LSDV})$ is concerned, consider (16). Since $\tilde{s}e(\hat{\beta}^{LSDV})$ is computed using $\hat{\beta}^{LSDV}$ in $\hat{v}_{it} = \bar{y}_{it} - \hat{\beta}\bar{x}_{it}$, for any values of β it holds that $\sqrt{n}\tilde{s}e(\hat{\beta}^{LSDV}) \xrightarrow{d} V$. Equation (13) follows. The proof for the POLS estimator follows essentially the same lines and therefore it is omitted. ■

TABLES

λ	σ_β	(ρ_e, θ_e)			
		$(0, 0)$	$(0.75, 0)$	$(0, 0.75)$	$(0, -0.75)$
0	0	0.025	0.067	0.025	0.041
	1	0.034	0.040	0.045	0.025
	2	0.027	0.026	0.030	0.024
0.2	0	0.032	0.057	0.033	0.039
	1	0.057	0.059	0.071	0.029
	2	0.037	0.037	0.055	0.026
0.4	0	0.055	0.059	0.054	0.057
	1	0.060	0.062	0.069	0.034
	2	0.046	0.044	0.057	0.035
0.6	0	0.073	0.074	0.071	0.072
	1	0.033	0.035	0.053	0.009
	2	0.015	0.015	0.028	0.008
0.8	0	0.085	0.085	0.085	0.087
	1	0.054	0.055	0.069	0.031
	2	0.042	0.042	0.051	0.026
1	0	0.074	0.074	0.074	0.077
	1	0.079	0.079	0.072	0.085
	2	0.089	0.088	0.085	0.081

Table 1: empirical rejection frequencies for \tilde{t}_β at a nominal size 5% - case with no common factors, $(\bar{\varphi}, \sigma_\varphi^2) = (0, 0)$.

λ	σ_β	(ρ_e, θ_e)			
		$(0, 0)$	$(0.75, 0)$	$(0, 0.75)$	$(0, -0.75)$
0	0	0.025	0.064	0.026	0.044
	1	0.034	0.039	0.044	0.025
	2	0.027	0.026	0.030	0.024
0.2	0	0.035	0.048	0.032	0.026
	1	0.071	0.072	0.077	0.063
	2	0.050	0.050	0.060	0.046
0.4	0	0.060	0.059	0.060	0.066
	1	0.074	0.073	0.077	0.070
	2	0.059	0.057	0.066	0.051
0.6	0	0.070	0.071	0.073	0.076
	1	0.054	0.056	0.059	0.049
	2	0.029	0.030	0.038	0.026
0.8	0	0.079	0.077	0.076	0.086
	1	0.061	0.060	0.061	0.058
	2	0.045	0.046	0.050	0.042
1	0	0.077	0.078	0.081	0.074
	1	0.083	0.084	0.081	0.093
	2	0.093	0.094	0.086	0.094

Table 2: empirical rejection frequencies for \tilde{t}_β at a nominal size 5% - case with common factors, $(\bar{\varphi}, \sigma_\varphi^2) = (1, 1)$.