On the asymptotic t-test for large nonstationary panel models

Lorenzo Trapani*
Cass Business School

September 20, 2010

Abstract

The asymptotic t-test for the long-run average in a heterogeneous nonstationary panel model is derived. 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, strong cross sectional dependence in the errors and in the regressors and mixed stationary/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 power properties,and the correct size, even for n as small as 25.

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

*Cass Business School, Faculty of Finance, 106 Bunhill Row, London EC1Y 8TZ, Tel.: +44 (0) 207 040 5260; email: L.Trapani@city.ac.uk.

1 Introduction

Consider the heterogeneous panel regression model

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}, \tag{1}$$

where $i = 1, ..., n$, $t = 1, ..., T$ and the variables y_{it} and x_{it} are both $I(1)$ for each i. It is well known, since the seminal contributions by Engle and Granger (1987) and Phillips (1986), that when the error term u_{it} is $I(1)$, there is no cointegration between y_{it} and x_{it}. However, this does not imply that there is no relationship of any kind between y_{it} and x_{it}: this would be the case if and only if $\beta_i = 0$, which is genuine spurious regression. Finding $u_{it} \sim I(1)$ could be due to various other reasons, which have been summarised, *inter alia*, by Fuertes (2008) and Choi, Hu and Ogaki (2008); these include e.g. neglected nonlinearity in the form of lumpy adjustment costs or Markov switching, nonstationary measurement error, or other forms of misspecification such as omitting an important nonstationary variable from the specification of (1). Of course, even when $\beta_i \neq 0$, β_i cannot be interpreted as a cointegration coefficient, due to the lack of cointegration. However, β_i may be interpreted as a statistical long-run correlation coefficient between y_{it} and x_{it}, and thus the information contained in β_i would not be entirely worthless. The question then arises as to how to estimate consistently and conduct inference on β_i.

In a pure time series setting, it is well known that when u_{it} is $I(1)$, then β_i may not be estimated consistently - see e.g. Phillips (1986). Conversely, in a large panel framework (where both n and T are large), the seminal contributions by Kao (1999) and Phillips and Moon (1999b) show that even when (some or all of) the u_{it}s are nonstationary, it is still possible to estimate consistently, at a rate $O_p\left(n^{-1/2}\right)$, the average long run correlation between y_{it} and x_{it}, i.e. the coefficient β customarily known as the "long-run average" - see Phillips and Moon (1999b, 2000). The issue of estimating average elasticities has been paid great attention in the empirical literature. For example, in the context of growth studies, as Fuertes (2008, p. 3356) points out "[...] output is regressed against the stock of physical capital and/or education, and the error term captures technical progress. The individual equations are not long-run equilibrium relations if technology is $I(1)$, but the average capital (or education) elasticity $E(\beta_i)$ is still of interest for economists". See also the discussions in Temple (1999), Baltagi and Kao (2000), Sun (2004) and Smith and Fuertes (2009). An important potential shortcoming of the long run average coefficient β is that it is equal to the average β_i, say $E(\beta_i)$, only under certain assumptions. For example, $\beta = E(\beta_i)$

when the β_is and the x_{it}s are independent, although other assumptions would grant the same result - see e.g. the discussion in Phillips and Moon (2000). In principle, $E(\beta_i)$ could be estimated using the Mean-Group estimator (henceforth $\hat{\beta}^{MG}$) proposed by Pesaran and Smith (1995); if all units cointegrate, then it is well known that $\hat{\beta}^{MG}$ is $\sqrt{n}T$-consistent (see also Smith and Fuertes, 2009, p. 52-53). If, however, the panel is mixed, with u_{it} being $I(0)$ for some units and $I(1)$ for some other units, then $\hat{\beta}^{MG}$ becomes inconsistent. Conversely, the estimator of the long run average parameter β, say $\hat{\beta}$, is $\sqrt{n}$-consistent irrespective of whether all, some or no units cointegrate in (1). Thus, $\hat{\beta}$ is a more robust estimator of the long-run average relationship between y_{it} and x_{it} (see Phillips and Moon, 2000, p. 272). In recent years, the literature on nonstationary panels has developed various techniques to estimate β. Phillips and Moon (1999a, 1999b) discuss the properties of the Pooled OLS (POLS) and of the Least Squares Dummy Variables (LSDV) estimators, finding that these are $\sqrt{n}-$consistent as $(n, T) \to \infty$ under the restriction $\frac{n}{T} \to 0$; a similar analysis is contained in Kao (1999). Sun (2004) proposes an estimator of β based on the estimators of the long run variance/covariance matrices in (1), showing that this is $\sqrt{n}-$consistent as $(n, T) \to \infty$ under the milder restriction $\frac{\sqrt{n}}{T} \to 0$.

The asymptotic t-test for long-run averages

Despite the availability of consistent estimators for β in a panel setting, asymptotic theory and inference differ from the time series case; particularly, Kao (1999) shows that t-tests based on estimates of β diverge as $T \to \infty$, thereby leading to null rejection frequency asymptotically equal to 1. Hence, 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 behavior 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

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

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

$$y_{it} = \alpha_i + \beta x_{it} + v_{it}, \tag{3}$$

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

$$se\left(\hat{\beta}\right) = \sqrt{\frac{1}{nT}\frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{v}_{it}^{2}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}^{2}}}, \tag{4}$$

where $\hat{v}_{it}$ is the regression residual and $\bar{x}_{it}$ is the demeaned version of x_{it}, i.e. $\bar{x}_{it} = x_{it} - T^{-1}\sum_{t=1}^{T} x_{it}$. 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 $t_\beta = O_p\left(\sqrt{T}\right)$ - 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): this is a consequence of $se\left(\hat{\beta}\right)$ 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 the bootstrap may be computationally burdensome, the simulations in Fuertes (2008) show that size distortion is reduced. However, as $T \to \infty$, t_β diverges and thus even the bootstrap becomes inapplicable.

The main contribution of this paper is to propose a general methodology to consistently estimate the standard error of the estimated β, obtaining $\widetilde{se}\left(\hat{\beta}\right)$. Thus, $\hat{\beta} - \beta_0$ can be properly studentized by $\widetilde{se}(\cdot)$, thereby obtaining a nuisance free statistic, whose limiting distribution is standard Normal as $(n, T) \to \infty$ jointly with $\frac{n}{T} \to 0$. As an illustration, we calculate the standard errors of the LSDV and POLS estimators for β, defined respectively as $\widetilde{se}\left(\hat{\beta}^{POLS}\right)$ and $\widetilde{se}\left(\hat{\beta}^{LSDV}\right)$, and study the asymptotic t-test based on the two estimators.

The empirical relevance of conducting a t-test for the long-run average β in a panel that is a mixture of cointegrating and spurious regressions can be motivated by noting that in empirical applications it is rather frequent to find evidence of a panel which is a mixture of spurious and cointegrating regressions. As an example, consider the literature on PPP. Whilst empirical evidence shows that very often some units are not cointegrated and/or that the individual slopes are not homogeneous as prescribed by the strong version of the PPP (see e.g. Pedroni, 2001, and the references therein), it is still interesting to verify whether on average, in the panel, there is a one-to-one correspondence between the movements of nominal exchange rates and aggregate price ratios. Denoting the elasticity of the former to the latter as β, this would correspond to testing for $H_0 : \beta = 1$. As another example, in the literature on the Feldstein-Horioka puzzle, estimation of the long-run average is a routine exercise, also motivated by the systematic failure of cointegration in the unit specific equations (Phillips and Moon, 2000, p. 279; Coakley *et al.*, 2004); in this context, a customarily tested hypothesis is that of perfect capital mobility, which is a t-test for the null $H_0 : \beta = 0$, where β represents the elasticity of investments to savings.

From a statistical point of view, tests based on the modified t-statistic proposed here, $\tilde{t}_\beta$, are a useful complement to bootstrap for at least three reasons. First, the test is shown to be asymptotically $N(0, 1)$ under serial dependence, in presence of strong cross sectional dependence arising from the presence of a factor-loading structure in in the regressor x_{it} and/or in the error term u_{it} and for any value of the percentage of spurious regressions in the panel. Thus, the test is robust to all the cases considered in Fuertes (2008), and can be used with no need for pre-testing, thereby avoiding e.g. having to apply unit root tests to the residuals $\hat{u}_{it}$. Note that some bootstrapping algorithms (e.g. the pre-testing sieve bootstrap, discussed in Fuertes, 2008) do require pre-testing. Also, the estimation of the fraction of spurious regressions is not required either (see also Ng, 2008, and Trapani, 2010). A second advantage is that 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 $\tilde{t}_\beta$ instead of t_β directly should be enhanced, also allowing for asymptotic refinements. Thus, the results of this paper can be viewed as both an alternative or a complement to bootstrap. 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 of the estimated β; the modified t-test is studied in Section 3. Numerical evidence from synthetic data is reported in Section 4. Section 5 concludes. Proofs and Tables are in Appendices A and B respectively. 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 $\lfloor \cdot \rfloor$ the integer part; for a generic Brownian motion W, we define $\bar{W} = W - \int W$ to denote the demeaned version of W.

2 Assumptions and preliminary 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

$$x_{it} = x_{it-1} + e_{it}^{x}, \tag{5}$$

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}:

$$u_{it} = \left[\varphi_i F_t^{(\lambda)} + e_{it}^{u(\lambda)}\right] d_{\lambda,i} + \left[\varphi_i F_t^{(1-\lambda)} + e_{it}^{u(1-\lambda)}\right] (1 - d_{\lambda,i}), \tag{6}$$

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

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

with $\left|\rho_i^F\right| < 1$ and $|\rho_i^e| < 1$; $\left\{\varepsilon_t^F, \varepsilon_{it}^e\right\}$ is assumed to be a linear stationary process in both cases. 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. Allowing for $\lambda \in [0, 1]$ means that the boundary cases where 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. 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 (1999b, p. 1080) - and also possibly because $u_{it} \sim I(1)$ for some i. All results are derived assuming only one regressor in (1) and only one factor in (6). 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.

Consider the following assumptions.

Assumption 1: [*time series properties*] *(i)* for all i it holds that $|\varphi_i|^{2+\delta} < \infty$ for some $\delta > 0$; *(ii)* letting $\omega_{it} = \left[e^x_{it}, \varepsilon^F_t, \varepsilon^e_{it}\right]'$, 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 (IP) holds for the partial sums of ω_{it} such that for all $r \in [0,1]$, $T^{-1/2}\sum_{t=1}^{\lfloor Tr \rfloor}\omega_{it} \xrightarrow{d} B_i(r)$, where $B_i(r) = [B_{xi}(r), B_{\varepsilon i}(r), B_{ei}(r)]$ has covariance matrix $\Omega_i = diag\{\sigma^2_x, \sigma^2_F, \sigma^2_u\}$.
Assumption 2: [*heterogeneous coefficients*] *(i)* for all i, the coefficients α_i and β_i are *i.i.d.* with $E(\alpha_i) = \alpha$, $Var(\alpha_i) = \sigma^2_\alpha < \infty$ and $E(\beta_i) = \beta$, $Var(\beta_i) = \sigma^2_\beta < \infty$ 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.

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 IP 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 (6)) - 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 $F^{(\cdot)}_t$ apart from the validity of the IP and $|\varphi_i|^{2+\delta} < \infty$. The latter is needed in order for the variance of $\varphi_i F^{(\cdot)}_t$ 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 $F^{(\cdot)}_t$ need to be identified or estimated, as they will be treated as nuisance parameters. Correlation between ε^F_t and ε^e_{it} is assumed to be zero; while this is common in panel factor models (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. 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).

Assumption 2 is needed only in order for the Central Limit Theorem (CLT) and the Law of Large Numbers (LLN) to hold for the α_is and the β_is such that e.g. $n^{-1}\sum_{i=1}^{n}(\beta_i-\beta)^2 \xrightarrow{p} \sigma_\beta^2$ and $\left(n\sigma_\beta^2\right)^{-1/2}\sum_{i=1}^{n}(\beta_i-\beta) \xrightarrow{d} N(0,1)$; less strict assumptions could be considered as long as the CLT and the LLN hold in the weak form. A consequence of Assumption 2*(ii)* is that $E(\beta_i)=\beta$.

As well as considering a framework where the regressors x_{it} are cross-sectionally independent, typical in the early panel literature (Phillips and Moon, 1999b; Kao, 1999), recently the literature has also considered the case of strong cross-dependence in the regressors (see e.g. the analysis and the comments in Pesaran, 2006; and Kapetanios *et al.*, 2009). This case, arising e.g. from the presence of common factors in the DGP of the x_{it}s, can be accommodated using exactly the same theoretical framework as developed for the case of cross-sectionally independent regressors, and it is explored separately for the sake of notational simplicity. Consider the following DGP, alternative to (5)

$$x_{it}=\vartheta_i G_t, \tag{7}$$

where the nonstationary common factor G_t is defined by $G_t=G_{t-1}+\varepsilon_t^G$, with ε_t^G a stationary, linear process. Equation (7) could be generalised e.g. to include idiosyncratic nonstationary shocks as well, i.e. $x_{it}=\vartheta_i G_t+\phi_{it}$ with $\phi_{it}\sim I(1)$, which we do not consider for brevity. Consider the following additional assumptions.

Assumption 3: *[cross-correlated regressors] (i)* either (a) $E(\varphi_i\vartheta_i)=O\left(T^{-1/2}\right)$ for all i, or (b) $u_{it}=e_{it}^{u(\lambda)}d_{\lambda,i}+e_{it}^{u(1-\lambda)}(1-d_{\lambda,i})+\varphi_i F_t^{(1-\lambda)}$; *(ii)* letting $\tilde{\omega}_{it}=\left[\varepsilon_{it}^G,\varepsilon_t^F,\varepsilon_{it}^e\right]'$, it holds that $\tilde{\omega}_{it}$ is a linear stationary process with $E\left\|\tilde{\omega}_{it}\right\|^{2+\delta}<\infty$ for some $\delta>0$ and an IP holds whereby $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\omega_{it}\xrightarrow{d}\tilde{B}_i(r)$, where $\tilde{B}_i(r)$ has covariance $\tilde{\Omega}_i=diag\left\{\sigma_G^2,\sigma_F^2,\sigma_u^2\right\}$; *(iii)* $E\left|\varphi_i\right|^{4+\delta}<\infty$ and $E\left|\vartheta_i\right|^{4+\delta}<\infty$ for some $\delta>0$ with $\{\varphi_i,\vartheta_i\}$ is independent of β_i.

Assumption 3 is made of three parts. Parts *(ii)* and *(iii)* are a generalisation of Assumptions 1*(ii)* and 2. Particularly, part *(iii)* is a sufficient condition to show that $E\left|\varphi_i\vartheta_i\right|^{2+\delta}<\infty$. The assumption that G_t and F_t have no long-run correlation could be relaxed but is maintained here for simplicity. Part *(i)* controls the amount of cross-dependence allowed for. Part *(i)*(a) is needed in order to establish that the correlation between x_{it} and u_{it} (conditional on the common shocks G_t and F_t) is weak; a similar (in spirit) requirement is Assumption

CU in Andrews (2005, p. 1559). If $E(\varphi_i \vartheta_i) \neq 0$ and finite, it would no longer be possible to have a consistent estimator of the long-run average β using an ordinary OLS framework (see also, in a cross sectional setting, Andrews, 2005); consistency could be achieved using the estimation technique proposed by Pesaran (2006), although in order for this to be applied without incurring in overdifferencing, one needs either $\lambda = 0$ or $\lambda = 1$. Alternatively, it can be shown that OLS is consistent even if $E(\varphi_i \vartheta_i) \neq 0$, as long as the common factors in the error term are stationary, as required in part (b) of Assumption 3*(i)*. Further explanations as to how part *(i)* works are below, after Lemma 2.

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

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

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

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 notation will be used henceforth: we let $\varphi^k = \lim_{n\to\infty} n^{-1}\sum_i \varphi_i^k$ for $k = 1, 2, ...$; ϑ^k is defined similarly. Letting W_j for $j = 1, 2, ...$ be independent standard Brownian motions, we define $D_j = \int_0^1\int_0^1 \bar{W}_j(s)\bar{W}_j(r)\ \psi(r,s)\,drds$ with $\psi(r,s) = \frac{1}{3} + \min(r,s) - \left[\frac{1}{2}r^2 + r(1-r)\right] - \left[\frac{1}{2}s^2 + s(1-s)\right]$ and,

$$D_1' = \varphi^2 \int_0^1\int_0^1 W_1(r)\, W_1(s) \min(r,s)\, drds$$
$$- (\varphi^1)^2 \left[\int_0^1 W_1(r)\,dr\right] \int_0^1\int_0^1 \left[W_1(r) + \bar{W}_1(r)\right] \min(r,s)\, drds.$$

The following theorem characterizes the limiting distribution of $\hat{\beta}^{LSDV}$ and $\hat{\beta}^{POLS}$ and of their standard errors computed from (4).

Lemma 1 *Let the DGP of x_{it} be the one in (5), and let Assumptions 1 and 2 hold. Then,* *as* $(n, T) \to \infty$ *with* $\frac{n}{T} \to 0$, *it holds that*

$$\sqrt{n}\left(\hat{\beta}^{LSDV} - \beta\right) \xrightarrow{d} \left[\frac{2}{5}\frac{\lambda\sigma_u^2}{\sigma_x^2} + \frac{9}{5}\sigma_\beta^2 + 6\frac{\lambda\varphi^2\sigma_F^2}{\sigma_x^2}D_1\right]^{1/2} \times Z, \tag{8}$$

$$\sqrt{nT}se\left(\hat{\beta}^{LSDV}\right) \xrightarrow{d} \sqrt{\frac{\lambda\sigma_u^2}{\sigma_x^2} + \sigma_\beta^2 + \frac{6\lambda\varphi^2\sigma_F^2}{\sigma_x^2}\left(\int \bar{W}^2\right)}, \tag{9}$$

and

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

$$\sqrt{nT}se\left(\hat{\beta}^{POLS}\right) \xrightarrow{d} \sqrt{\frac{\lambda\sigma_u^2}{\sigma_x^2} + \sigma_\beta^2 + \frac{2\lambda\varphi^2\sigma_F^2}{\sigma_x^2}\left(\int W^2\right)}, \tag{11}$$

with $Z \sim N(0,1)$ *independent of all the other random variables. For both estimators, it holds that* $t_\beta = O_p\left(\sqrt{T}\right)$ *as* $(n, T) \to \infty$ *under* $\frac{n}{T} \to 0$.

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, 1999b). The novel result in (8) and (10) 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 non-stationarity of $F_t^{(\cdot)}$, which makes the asymptotic theory in Phillips and Moon (1999b) not applicable here; see also the discussion in Kao, Trapani and Urga (2008, 2010). Equations (9) and (11) show that, for similar reasons, $se\left(\hat{\beta}\right)$ 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.

Equation (8) stipulates the consistency of $\hat{\beta}^{LSDV}$ under various circumstances: mixed presence of cointegration and spurious regressions (so that $\lambda \in [0, 1]$), various levels of slope heterogeneity (so that $\sigma_\beta^2 \in [0, +\infty)$), and presence of strong cross-sectional dependence. Since the asymptotics of the LSDV estimator (as well as that of the POLS one) is robust to the values of various nuisance parameters, it can be used, for the purpose of carrying out a t-test on β, without the need of pre-testing for, e.g. panel cointegration or homogeneity. An alternative approach to the estimation of β is using some Mean Group procedure. Particularly, in the boundary cases $\lambda = 0$ and $\lambda = 1$, as an alternative estimation method one could use the Common Correlated Effects Mean Group (CCE-MG) estimator, discussed in Kapetanios *et al.* (2009) and Pesaran (2006, applying it to first differenced data) respectively. The CCE-MG estimator of β, say $\hat{\beta}^{CCEMG}$, is $\sqrt{n}$-consistent and it holds that (see e.g. Theorem 2 in Pesaran, 2006, p. 985) $\sqrt{n}\left(\hat{\beta}^{CCEMG}-\beta\right)\xrightarrow{d}N\left(0,\sigma_\beta^2\right)$. Thus, $\hat{\beta}^{MG}$ is asymptotically more efficient, which could result in higher power for the t-test. However, the range of applicability of Lemma 1 states that the LSDV and POLS estimators are more robust, since they can be used for all values of $\lambda \in [0,1]$ and not only at the boundaries.

When the DGP of the x_{it}s contains common factors, as in (7), the results are conceptually very similar, and the method of proof is exactly the same. We illustrate this by presenting the asymptotics of the LSDV estimator only, in order to save space - the limiting distribution of the Pooled-OLS estimator can be calculated following exactly the same steps as in the proof of Lemma 1.

Lemma 2 *Let the DGP of x_{it} be the one in (7), and let Assumptions 1-2 and 3(ii) hold. Then, under either Assumption 3(i)(a) or 3(i)(b), as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$, it holds that*

$$\sqrt{n}\left(\hat{\beta}^{LSDV}-\beta\right)\xrightarrow{d}\Sigma_G^{-1}\times \tag{12}$$

$$\left[\sigma_\beta^2\Sigma_G^2+\lambda\sigma_u^2\sigma_G^2\vartheta^2 D_2+d_b\lambda\sigma_u^2\sigma_G^2\kappa^2\left(\int \bar{W}_1\bar{W}_2\right)^2\right]^{1/2}\times Z,$$

$$\sqrt{nT}se\left(\hat{\beta}^{LSDV}\right)\xrightarrow{d}\Sigma_1^{-1/2}\left[\frac{\lambda\sigma_u^2}{6}+\sigma_\beta^2\Sigma_G+d_b\lambda\varphi^2\sigma_F^2\left(\int \bar{W}_1^2\right)\right]^{1/2}, \tag{13}$$

where $\Sigma_G=\sigma_G^2\vartheta^2\int\bar{W}_2^2$, $\kappa^2=\lim_{n\to\infty}n^{-1}\sum_i\varphi_i^2\vartheta_i^2$ and the dummy variable d_b is equal to 1 under Assumption 3(i)(a) and zero otherwise. It holds that $t_\beta=O_p\left(\sqrt{T}\right)$ as $(n,T)\to\infty$ under $\frac{n}{T}\to 0$.

Lemma 2 states a very similar result to Lemma 1, in terms of rate of convergence of $\hat{\beta}^{LSDV}$ and also as far as its standard error is concerned. The method of proof, as shown in Appendix A, is the same. As in Lemma 1, the presence of common factors introduces a randomness which is not smoothed away by cross-sectional averaging: the limiting distribution of $\hat{\beta}^{LSDV}$ is mixed Normal instead of Normal. In this case, this is due to the presence of common factors in both the error term u_{it} and in the regressor x_{it}, as it can be seen by noting that the denominator of $\hat{\beta}^{LSDV}-\beta$ does not converge in probability to a number as in the case of Lemma 1, but it converges weakly to a random variable, Σ_G.

As equation (24) in Appendix A shows, Assumption 3*(i)* is needed in order to ensure consistency of $\hat{\beta}^{LSDV}$; this is because, in the expression of the denominator of $\hat{\beta}^{LSDV}-\beta$, there is a term equal to $\left(n^{-1/2}\sum_{i=1}^{n}\varphi_i\vartheta_i\right)\left(T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t\right)$. As $(n,T)\to\infty$ this has expectation (conditional on F_t and G_t) given by $\Delta=\sqrt{n}E\left(\varphi_i\vartheta_i\right)\left(T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t\right)$; this is negligible (and therefore there is no asymptotic bias) if either $E\left(\varphi_i\vartheta_i\right)$ or $T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t$ converge to zero in some sense. This happens under Assumption 3*(i)*(b), as the FCLT entails that $E\left(T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t\right)=O_p\left(T^{-1}\right)$, and therefore $\Delta=O_p\left(\frac{\sqrt{n}}{T}\right)$; alternatively, Assumption 3*(i)*(a) suffices to show that $\Delta=O_p\left(\sqrt{\frac{n}{T}}\right)$. In both cases, the restriction $\frac{n}{T}\to 0$ as $(n,T)\to\infty$ entails that $\Delta=o_p\left(1\right)$.

3 The asymptotic t-test

Equations (8)-(11) in Lemma 1, and (12)-(13) in Lemma 2, show that, as already proved by Kao (1999) under different assumptions, $t_\beta=O_p\left(\sqrt{T}\right)$ both when computed using the LSDV and the POLS estimator. Therefore, the ordinary t-test t_β can not be employed due to two reasons. First, t_β diverges as $T\to\infty$ and therefore, as $T\to\infty$, 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 $se\left(\hat{\beta}\right)$ does not converge to the asymptotic variance of $\hat{\beta}$.

In order to solve this problem, in this section we propose an alternative approach to the estimation of the variances of $\hat{\beta}^{LSDV}$ and $\hat{\beta}^{POLS}$, in a similar spirit to Bai and Kao (2006).

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

$$\widetilde{se}\left(\hat{\beta}^{LSDV}\right) = \frac{\sqrt{\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\bar{x}_{it}\hat{v}_{it}\right)^2}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}^2}, \tag{14}$$

$$\widetilde{se}\left(\hat{\beta}^{POLS}\right) = \frac{\sqrt{\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\tilde{x}_{it}\check{v}_{it}\right)^2}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{x}_{it}^2}. \tag{15}$$

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

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

and $\tilde{t}_\beta^{POLS}$ can be defined similarly.

The following theorem characterizes the limiting distribution of $\tilde{t}_\beta$ under H_0 and under local alternatives $H_A^{(n)} : \beta = \beta_0 + c/\sqrt{n}$.

Theorem 1 *Consider (5), and let Assumptions 1 and 2 hold. Then as* $(n, T) \to \infty$ *with* $\frac{n}{T} \to 0$,

$$\tilde{t}_\beta^{LSDV} \xrightarrow[H_0]{d} N(0,1), \tag{16}$$

$$\tilde{t}_\beta^{LSDV} \xrightarrow[H_A^{(n)}]{d} N(c,1). \tag{17}$$

The same results hold for $\tilde{t}_\beta^{POLS}$, *and also under (7) and Assumption 3.*

Theorem 1 shows the distribution of the modified t-statistics $\tilde{t}_\beta^{LSDV}$ and $\tilde{t}_\beta^{LSDV}$ under the null and under local alternatives. The studentization in (14) ensures that $\tilde{t}_\beta$ has a standard Normal distribution as $(n, T) \to \infty$, 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 (6). The statistic is robust to the values of nuisance parameters: for example, when $\lambda = 0$ or 1, and regardless of the presence and the extent of cross-sectional dependence. Also, the limiting distribution of $\tilde{t}_\beta$ is nuisance free, and therefore bootstrapping $\tilde{t}_\beta$ could yield asymptotic refinements. This holds true also if there are common factors in the DGP of the x_{it}s, in spite of the different limiting distribution of $\hat{\beta}^{LSDV}$ as illustrated by Lemma 2. Finally, note that Theorem 1 holds for the joint limit case, provided that $\frac{n}{T} \to 0$. 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 $\sqrt{n}\widetilde{se}\left(\hat{\beta}\right)$ as an estimator of $Var\left(\hat{\beta}\right)$ under $H_A^{(n)}$.

Theorem 1 states that both $\left[\widetilde{se}\left(\hat{\beta}^{LSDV}\right)\right]^2$ and $\left[\widetilde{se}\left(\hat{\beta}^{POLS}\right)\right]^2$ are consistent estimators for the variance of $\hat{\beta}^{LSDV}$ and $\hat{\beta}^{POLS}$ respectively. Whilst details are in Appendix A, it is worth providing an intuition of this result here. Consider the numerator of (14), $n^{-1}\sum_{i=1}^{n}\left(T^{-2}\sum_{t=1}^{T}\bar{x}_{it}\hat{v}_{it}\right)^{2}$. Given that $\hat{\beta}^{LSDV}$ is a consistent estimator, $n^{-1}\sum_{i=1}^{n}\left(T^{-2}\sum_{t=1}^{T}\bar{x}_{it}\hat{v}_{it}\right)^{2}=n^{-1}\sum_{i=1}^{n}\left(T^{-2}\sum_{t=1}^{T}\bar{x}_{it}v_{it}\right)^{2}+o_{p}(1)$. The quantities $\left(T^{-2}\sum_{t=1}^{T}\bar{x}_{it}v_{it}\right)^{2}$ are (after conditioning) independent across i and have finite expectation. Thus, a cross-sectional LLN holds as $n\rightarrow\infty$, for any T, ensuring that $n^{-1}\sum_{i=1}^{n}\left(T^{-2}\sum_{t=1}^{T}\bar{x}_{it}v_{it}\right)^{2}\xrightarrow{p}\lim_{n\rightarrow\infty}n^{-1}\sum_{i=1}^{n}E\left[\lim_{T\rightarrow\infty}\left(T^{-2}\sum_{t=1}^{T}\bar{x}_{it}v_{it}\right)^{2}\right]$, which, as $T\rightarrow\infty$, corresponds to the asymptotic variance of the numerator of $\hat{\beta}^{LSDV}$ by definition - see equation (20) in Appendix A. Thus, similarly to the Phillips and Moon (1999b) approach, the fact that T passes to infinity is just ancillary to the main argument of the proof. This is another example that illustrates how it is possible, with panel data, to estimate long run variances using the full time series sample due to the availability of a large amount of cross sectional information - see also the discussion in Sun (2004).

4 Monte Carlo evidence

The null rejection frequency and the power of tests based on $\tilde{t}_{\beta}^{LSDV}$ are reported in Tables 1-3 in Appendix B; results for $\tilde{t}_{\beta}^{POLS}$ showed a similar pattern and thus were not reported. 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 $\sqrt{c_{\alpha}\left(1-c_{\alpha}\right)/2000}$; 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 $(n,T)=(25,100)$, $(50,200)$ and $(100,400)$; tables are reported in Appendix B. Other experiments, not reported here, showed that size distortion does arise for $n<25$. Thus, $n=25$ seems to be the threshold below which the asymptotic approximation of $\tilde{t}_{\beta}$ is no longer accurate.

Synthetic data are generated using equation (1)

$$y_{it}=\alpha_{i}+\beta_{i}x_{it}+\left[\varphi_{i}F_{t}^{(\lambda)}+e_{it}^{u(\lambda)}\right]d_{\lambda,i}+\left[\varphi_{i}F_{t}^{(1-\lambda)}+e_{it}^{u(1-\lambda)}\right]\left(1-d_{\lambda,i}\right),$$

with $x_{it}=x_{it-1}+e_{it}^{x}$, $\alpha_{i}\sim i.i.d.N(1,1)$, $\beta_{i}\sim i.i.d.N(1,\sigma_{\beta}^{2})$ and $\varphi_{i}\sim i.i.d.N(\bar{\varphi},\sigma_{\varphi})$. We consider the cases $\sigma_{\beta}^{2}=\{0,0.25,0.5\}$, based on "typical" values of heterogeneity (see the discussion in Trapani and Urga, 2009) and also $\sigma_{\beta}^{2}=4$ to investigate the robustness of the test versus extreme levels of heterogeneity; and we set $(\bar{\varphi}, \sigma_\varphi) = \{(0,0), (1,1)\}$. The fraction of spurious regressions, λ, is set as $\lambda \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. When common factors are present in the DGP of the x_{it} as in (7), we set $\vartheta_i \sim i.i.d.N(\bar{\vartheta}, \sigma_\vartheta)$ with $(\bar{\vartheta}, \sigma_\vartheta) = (1,1)$.

The innovations $e_t = \left[e^x_{it}, \varepsilon^F_t, e^u_{it}, \varepsilon^G_t\right]'$ are generated by creating a Gaussian *i.i.d.* sequence $\dot{e}_t = \left[\dot{e}^x_{it}, \dot{\varepsilon}^F_t, \dot{e}^u_{it}, \dot{\varepsilon}^G_t\right]'$ with $Var\left(\dot{e}^x_{it}\right) = Var\left(\dot{\varepsilon}^G_t\right) = \sigma^2_x$ and $Var\left(\dot{\varepsilon}^F_t\right) = Var\left(\dot{e}^u_{it}\right) = 1$; 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 $(\rho_e, \theta_e) \in \{(0,0), (0.75, 0), (0, 0.75), (0, -0.75)\}$.

Tables 1a-1c, 2a-2c and 3a-3c in Appendix B contain the null rejection frequencies for the cases $(\bar{\varphi}, \sigma_\varphi) = (0,0)$ and $(1,1)$, and $(\bar{\vartheta}, \sigma_\vartheta) = (1,1)$ respectively.

We comment on the main results of Tables 1 and 2; Tables 3 are commented in the last bullet point. The following stylized facts emerge:

- The presence of *strong cross-sectional dependence* in the error term u_{it} has a tendency to worsen the results, at least in small samples. Although a factor structure in the error term alters only slightly the numbers in Tables 2a and 2b (when compared to those in Tables 1a and1b) for the cases $(n, T) = (25, 100)$ and $(50, 200)$, 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 is driven by units with $I(1)$ errors and therefore with $I(1)$ 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. However, as the sample size increases, this discrepancy is mitigated, and for $(n, T) = (100, 400)$ the figures are very close, as one could realise comparing Tables 1c and 2c.

- 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. This is attenuated for larger sample sizes, and when $(n, T) = (100, 400)$ the null rejection frequencies are closer to their nominal values (Tables 1c and 2c), although there still is some tendency to oversizement for $\lambda = 1$.

- 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, particularly in small samples with $(n, T) = (25, 100)$ - see Tables 1a and 2a. Note that for $(\lambda, \sigma_\beta^2) = (0, 0)$ there is a tendency towards undersizement; this can be explained in light of equations (8) and (10), 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, $\tilde{t}_\beta$ tends to zero, whence the small size of the test. This is reinforced for $(n, T) = (50, 200)$ (Tables 1b and 2b) and $(n, T) = (100, 400)$ (Tables 1c and 2c). Also, the size is not correct for "large" values of σ_β^2 ($\sigma_\beta^2 = 4$ in the simulations), and particularly the test seems to be undersized in finite sample; of course, if σ_β^2 were not finite, this would make the CLT invalid, and therefore the asymptotics for $\hat{\beta}$ derived above would not hold. This could explain why "large" values of σ_β^2 entail size distortion; however, the undersizement disappears as the sample size increases to $(100, 400)$, see Tables 1c and 2c.

- 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 in small samples. This is fairly common in the literature - see e.g. Leybourne and Newbold (1999) and the references cited therein. As the sample size increases to $(n, T) = (50, 200)$ and $(100, 400)$, the impact of negative MA roots decreases, and the test is less and less undersized.

- The presence of *common factors in the regressor* x_{it} has a very strong impact on the empirical rejection frequency, making it depart sensibly from the 5% nominal value for $(n, T) = (25, 100)$ and also $(50, 200)$. There is no clear pattern of oversizement or undersizement. As (n, T) increases to $(100, 400)$, the size gets closer to its nominal value, suggesting that the convergence of $\hat{\beta}$ to β is slowed down by the presence of common factors. As λ increases, contrary to the other experiments, the empirical rejection frequency improves; a possible explanation is that the asymptotics is driven by the units that are spurious regressions, as the error term in that case is $I(1)$, as thus as λ increases the actual sample size, $\lfloor n\lambda \rfloor$, also increases, thus speeding up the convergence. The impact of the other nuisance parameters (mainly σ_β^2) is in line with the other experiments.

As well as running a set of experiments to assess the size of the test, the power versus the alternative $H_A : \beta = 2$ (under the null that $\beta = 1$) has been evaluated, and it is reported in Tables 1d, 2d and 3d. The tables show that the test has good power properties even for the case $(n, T) = (25, 100)$ - the cases $(n, T) = (50, 200)$ and $(100, 400)$ are not reported to save space, and show, as expected, an improvement in the power. The power does not seem to be strongly affected by the presence of various types of serial dependence; conversely, it is affected by other nuisance parameters. Particularly, as it emerges from considering the case $\sigma_\beta^2 = 4$, the power becomes lower as slope heterogeneity increases; this can be explained by noting that, as heterogeneity increases, the notion of "average slope" becomes fuzzy, thereby hampering inference conducted on β. Table 1d also shows that as the percentage of spurious regressions λ increases, the power seems to decrease, although this is less clear when common factors are present in the error term u_{it} and/or in the regressors x_{it} (Tables 2d and 3d respectively). Finally, the presence of common factors (Tables 2d and 3d) enhances the power, although it should be pointed out that the size of the test becomes worse in this case.

5 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 $(n, T) \to \infty$, t-tests based on this standardization are free from nuisance parameters, and can therefore be employed using the critical values from the 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 in the regressors and in the error term 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 pre-testing 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 $\tilde{t}_\beta$ is asymptotically pivotal, and thus the bootstrap version of $\tilde{t}_\beta^{LSDV}$ and $\tilde{t}_\beta^{POLS}$ could yield asymptotic refinements.

Finally, it is important to point out that although calculations have been explicitly carried out for the cases of the LSDV and the POLS estimators for illustrative purposes, the framework described here can be readily extended and applied to other estimators. As the proof of Theorem 1, and particularly equation (25), shows, the asymptotic standard error of an estimator of β can always be estimated, by construction, using similar formulas to (14) and (15). The main methodological guidelines, when calculating the standard error of an estimator, are to use a formula where: (1) at the numerator, in order to have the correct probability limit, the sum of the squared residuals $\hat{v}_{it}^2$ is replaced with the cross-sectional summation of the squares of the time series summations of the product between the residuals and the regressors; and (2) at the denominator, in order to have the appropriate order of magnitude, the sum of the squared regressors is not under square root.

Acknowledgements

I wish to thank Giovanni Urga for carefully reading an earlier version of this work and providing inspiring comments. I am grateful with the Editor, Prof. Erricos Kontoghiorghes, and anonymous Associate Editor and two anonymous Referees whose comments greatly improved the contents and presentation of this work. The usual disclaimer applies.

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. J. 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 B. Baltagi (ed), Panel Data Econometrics: Theoretical Contributions and Empirical Applications, Elsevier.

[5] Baltagi, B.H., Kao, C. 2000. Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey. Adv. Econometrics 15, 7-51.

[6] Baltagi, B.H., Kao, C. and Liu, L. 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. Econometrics J. 11, 554-572.

[7] Coakley, J., Fuertes, A.-M., Spagnolo, F. 2004. Is the Feldstein-Horioka Puzzle History?. Manchester School 72, 569-590.

[8] Choi, I., Hu, L., Ogaki, M. 2008. Robust Estimation for Structural Spurious Regressions and A Hausman-type Cointegration Test. J. Econometrics 142, 327-351.

[9] Engle, R.F., Granger, C.W.J. 1987. Co-integration and Error Correction: Representation, Estimation and Testing. Econometrica 55, 251-276.

[10] Fuertes, A.-M. 2008. Sieve Bootstrap t-test on Long-Run Average Parameters. Comput. Statist. and Data Anal. 52, 3354-70.

[11] Kao, C. 1999. Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. J. Econometrics 90, 1-44.

[12] Kao, C., Trapani, L. and Urga, G. 2008. Modelling and Testing for Structural Changes in Panel Cointegration Models with Common and Idiosyncratic Stochastic Trends. Mimeo, Cass Business School.

[13] Kao, C., Trapani, L. and Urga, G. 2010. Asymptotics for Panel Models with Common Shocks. Econometric Reviews (forthcoming)

[14] Kapetanios, G., Pesaran, M.H., Yamagata, T. 2009. Panels with Nonstationary Multifactor Error Structures. Mimeo.

[15] Leybourne, S., Newbold, P. 1999. On the Size Properties of Phillips-Perron Tests. J. Time Series Anal. 20, 51-61.

[16] Ng, S. 2008. A Simple Test for Non-Stationarity in Mixed Panels. J. Bus. Econom. Statist. 26, 113-127.

[17] Pedroni, P. 2001. Purchasing Power Parity Tests in Cointegrated Panels. Rev. Econom. Statist. 83, 727-731.

[18] Pesaran, M. H. 2006. Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. Econometrica 74, 967-1012.

[19] Pesaran, M. H., Smith, R.P. 1995. Estimating long-run relationships from dynamic heterogeneous panels. J. Econometrics 68, 79-113.

[20] Phillips, P. C. B. 1986. Understanding Spurious Regression in Econometrics. J. Econometrics 33, 311-340.

[21] Phillips, P. C. B., and Moon, H. R. 1999a. Linear Regression Limit Theory for Nonstationary Panel Data. Yale University, mimeographed.

[22] Phillips, P. C. B., and Moon, H. R. 1999b. Linear Regression Limit Theory for Nonstationary Panel Data. Econometrica 67, 1057-1112.

[23] Phillips, P. C. B., and Moon, H. R. 2000. Nonstationary Panel Data analysis: an Overview of Some Recent Developments. Econometric Reviews 19, 263-286.

[24] Phillips, P. C. B., and Solo, V. 1992. Asymptotics for linear processes. Ann. Stat. 20, 971–1001.

[25] Smith, R.P., Fuertes, A.-M. 2009. Panel Time Series. Cemmap mimeo 2009.

[26] Sun, Y. 2004. Estimation of the Long-run Average Relationship in Nonstationary Panel Time Series. Econometric Theory 20, 1227-1260.

[27] Trapani, L. 2010. Inferential Theory for Heterogeneity and Cointegration in Large Panels. Mimeo.

[28] Trapani, L. and Urga, G. 2009. Optimal Forecasting with Heterogeneous Panels: A Monte Carlo Study. Int. J. Forecasting 25, 567-586.

Appendix A: proofs and derivations

Henceforth, M_δ denotes a finite constant which only depends on a constant $\delta > 0$.

Proof of Lemma 1. We firstly prove (8) and (9); the method of proof builds on the theory developed by Kao, Trapani and Urga (2010). Henceforth, we define $F_t = F_t^{(\lambda)} d_{\lambda,i} + F_t^{(1-\lambda)} (1 - d_{\lambda,i})$ and $e_{it}^u = e_{it}^{u(\lambda)} d_{\lambda,i} + e_{it}^{u(1-\lambda)} (1 - d_{\lambda,i})$.

Consider the LSDV estimator. The estimation error is

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

Kao (1999) proved that

$$\frac{1}{nT^2} \sum_{i=1}^n \sum_{t=1}^T \bar{x}_{it}^2 \xrightarrow{p} \frac{1}{6} \sigma_x^2, \tag{18}$$

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

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

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

$$\begin{aligned} E(\xi_{iT} | C) &= E\left[\frac{1}{T^2} \sum_{t=1}^T (\beta_i - \beta) \bar{x}_{it}^2\right] + E\left[\frac{1}{T^2} \sum_{t=1}^T \bar{x}_{it} \bar{e}_{it}^u\right] + E\left[\frac{1}{T^2} \sum_{t=1}^T \varphi_i \bar{F}_t \bar{x}_{it}\right] \\ &= I + II + III. \end{aligned} \tag{19}$$

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 $II = O\left(T^{-1/2}\right)$. Similar arguments would lead to $III = O_p\left(T^{-1/2}\right)$. Thus, $E(\xi_{iT} | C) = O_p\left(T^{-1/2}\right)$ for all i. Let $\bar{\xi}_{iT} = \xi_{iT} - E(\xi_{iT})$ and let I_i be the σ-field generated by C and $\left\{\bar{\xi}_{1T}, ..., \bar{\xi}_{iT}\right\}$. Then $\left\{\bar{\xi}_{iT}, I_i; i \geq 1\right\}$ is a martingale difference sequence (MDS). As $T \to \infty$, a conditional Liapunov condition holds whereby $E\left|\bar{\xi}_{iT}\right| C\right|^{2+\delta} < \infty$ for all i (this follows from $\bar{\xi}_{iT}$ being *i.i.d.* across i conditional on C): using the C_r-inequality

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

Assumption 2 ensures that $E\left|\beta_i-\beta\right|^{2+\delta}$ and $E\left|\varphi_i\right|^{2+\delta}$ are finite. Also, define $\bar{W}_{1i}$, $\bar{W}_{2i}$ and $\bar{W}_3$ as independent standard demeaned Brownian motions. Then, as $T\to\infty$, the Continuous Mapping Theorem (CMT) ensures that $\left|T^{-2}\sum_{t=1}^{T}\bar{x}_{it}^2\right|^{2+\delta} \xrightarrow{d} \sigma_x^{2(2+\delta)}\left|\int \bar{W}_{1i}^2\right|^{2+\delta}$, $\left|T^{-2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right|^{2+\delta} \xrightarrow{d} \sigma_u^{2+\delta}\sigma_x^{2+\delta}\left|\int \bar{W}_{1i}\bar{W}_{2i}\right|^{2+\delta}$ and $\left|T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{x}_{it}\right|^{2+\delta} \xrightarrow{d} \sigma_F^{2+\delta}\sigma_x^{2+\delta}\left|\int \bar{W}_3\bar{W}_{1i}\right|^{2+\delta}$; all these quantities have finite expectation. Since $E\left|\bar{\xi}_{iT}\right| C\Big|^{2+\delta}<\infty$ as $T\to\infty$, it is possible to apply a CLT for MDS; as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{iT} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\bar{\xi}_{iT}+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\left(\xi_{iT}\right)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\bar{\xi}_{iT}+O_p\left(\sqrt{\frac{n}{T}}\right) \\
&\xrightarrow{d} \left[\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E\left(\bar{\xi}_{iT}^2\right|C\Big)\right]^{1/2}\times Z, \qquad (20)
\end{aligned}$$

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

$$\begin{aligned}
\xi_{iT}^2 &= \left(\beta_i-\beta\right)^2\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}^2\right)^2+\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right)^2+\varphi_i^2\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{x}_{it}\right)^2 \\
&\quad +2\left(\beta_i-\beta\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}^2\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right)+2\left(\beta_i-\beta\right)\varphi_i\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}^2\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{x}_{it}\right) \\
&\quad +2\varphi_i\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right)\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{x}_{it}\right) \\
&= a+b+c+d+e+f. \qquad (21)
\end{aligned}$$

Assumption 2 ensures that $E(d)=E(e)=0$. In light of Assumption 1, $a\xrightarrow{d}\left(\beta_i-\beta\right)^2\sigma_x^2\left(\int\bar{W}_{1i}^2\right)^2$ as $T\to\infty$ and $E\left[\left(\beta_i-\beta\right)^2\sigma_x^2\left(\int\bar{W}_{1i}^2\right)^2\right]=E\left(\beta_i-\beta\right)^2\sigma_x^2E\left(\int\bar{W}_{1i}^2\right)^2$ for all units, with $E\left(\beta_i-\beta\right)^2=\sigma_\beta^2$ by definition and $E\left(\int\bar{W}_{1i}^2\right)^2=\sigma_x^4/20$ - see Kao (1999) for the latter result. Assumption 1 also ensures that $b\xrightarrow{d}\sigma_x^2\sigma_u^2\left(\int\bar{W}_{1i}\bar{W}_{2i}\right)^2$ for $i=1,...,\lfloor n\lambda\rfloor$ and $b=O_p\left(T^{-1}\right)$ for the other units; since $\int\bar{W}_{1i}\bar{W}_{2i}$ has mean zero, it holds that $E\left(\int\bar{W}_{1i}\bar{W}_{2i}\right)^2=$

$Var\left(\int \bar{W}_{1i}\bar{W}_{2i}\right) = \sigma_e^2\sigma_x^2/90$ - see Kao (1999). This also entails $E(f) = 0$ as $T \to \infty$. Last, considering c

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

and after some algebra, using a strong approximation it can be proved that

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

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

$$\frac{1}{T^4}\sum_{t=1}^{T}\sum_{s=1}^{T}\bar{F}_t\bar{F}_s E\left(\bar{x}_{it}\bar{x}_{is}\right) \xrightarrow{d} \sigma_x^2\sigma_F^2\int_0^1\int_0^1 \bar{W}_3(r)\bar{W}_3(s)\psi(r,s)\,drds \equiv \sigma_x^2\sigma_F^2 D_1, \tag{22}$$

for $i = 1, ..., \lfloor n\lambda \rfloor$ and $b = O_p\left(T^{-1}\right)$ for the other units. To calculate $n^{-1}\sum_{i=1}^{n}E\left(\xi_{iT}^2\middle| C\right)$, note that the only terms that differ across units and do not vanish asymptotically are b and c in (21) and

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right)^2 &= \frac{\lambda}{n\lambda}\sum_{i=1}^{\lfloor n\lambda\rfloor}\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right)^2 + o_p(1) \xrightarrow{p} \lambda\frac{\sigma_e^2\sigma_x^2}{90}, \\ \frac{1}{n}\sum_{i=1}^{n}\varphi_i^2\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{x}_{it}\right)^2 &= \frac{\lambda}{n\lambda}\sum_{i=1}^{\lfloor n\lambda\rfloor}\varphi_i^2\left(\frac{1}{T^2}\sum_{t=1}^{T}\bar{F}_t\bar{x}_{it}\right)^2 + o_p(1) \xrightarrow{d} \lambda\varphi^2\sigma_x^2\sigma_F^2 D_1.\end{aligned}$$

Combining these results together we obtain (8).

Now we turn our attention to (9). Recall

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

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

$$
\begin{aligned}
\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{v}_{it}^2 &= \sum_{i=1}^{n}\sum_{t=1}^{T}\bar{e}_{it}^{u2} + \sum_{i=1}^{n}\varphi_i^2\sum_{t=1}^{T}\bar{F}_t^2 + \sum_{i=1}^{n}\left(\beta_i - \beta\right)^2\sum_{t=1}^{T}\bar{x}_{it}^2 \\
&\quad + \left(\hat{\beta} - \beta\right)^2\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}^2 - 2\left(\hat{\beta} - \beta\right)\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\beta_i - \beta\right)\bar{x}_{it}^2 \\
&\quad + 2\sum_{i=1}^{n}\left(\beta_i - \beta\right)\sum_{t=1}^{T}\bar{x}_{it}\bar{e}^u_{it} - 2\left(\hat{\beta} - \beta\right)\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}^u_{it} \\
&\quad + 2\sum_{i=1}^{n}\varphi_i\left(\beta_i - \beta\right)\sum_{t=1}^{T}\bar{x}_{it}\bar{F}_t - 2\left(\hat{\beta} - \beta\right)\sum_{i=1}^{n}\varphi_i\sum_{t=1}^{T}\bar{x}_{it}\bar{F}_t \\
&\quad + 2\sum_{i=1}^{n}\varphi_i\sum_{t=1}^{T}\bar{F}_t\bar{e}^u_{it} \\
&= I + II + III + IV + V + VI + VII + VIII + IX + X. \qquad (23)
\end{aligned}
$$

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

$$
\begin{aligned}
&\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{e}_{it}^{u2} = \frac{\lambda}{\left(n\lambda\right)T^2}\sum_{i=1}^{\lfloor n\lambda\rfloor}\sum_{t=1}^{T}\bar{e}_{it}^{u2} + o_p\left(1\right) \xrightarrow{p} \frac{\lambda\sigma_u^2}{6}, \\
&\frac{1}{nT^2}\sum_{i=1}^{n}\varphi_i^2\sum_{t=1}^{T}\bar{F}_t^2 \xrightarrow{d} \lambda\varphi^2\sigma_F^2\int\bar{W}_3^2, \\
&\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\beta_i - \beta\right)^2\bar{x}_{it}^2 \xrightarrow{p} \frac{\sigma_\beta^2\sigma_x^2}{6}.
\end{aligned}
$$

It follows that

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

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

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

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

The proof of (10) and (11) follows similar lines as the proof of (8) and (9), 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

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

Then, conditional on C, ς_{iT} is *i.i.d.*; similar passages as above lead to $E\left(\varsigma_{iT}|\,C\right)=O_p\left(T^{-1/2}\right)$. Letting $\bar{\varsigma}_{iT}=\varsigma_{iT}-E\left(\varsigma_{iT}|\,C\right)$, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$ the MDS CLT entails

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

with $Z\sim N(0,1)$ independent of $E\left(\varsigma_{iT}^2|\,C\right)$. Note also that the term $\sum_{i=1}^{n}\sum_{t=1}^{T}(\alpha_i-\alpha)\tilde{x}_{it}=O_p\left(\sqrt{n}T^{3/2}\right)$ is negligible. In order to prove (10), it is necessary to find the value of $E\left(\varsigma_{iT}^2|\,C\right)$.

Note that

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

Note that $\sum_{i=1}^n\sum_{t=1}^T\left(\beta_i-\beta\right)x_{it}=O_p\left(\sqrt{n}T^{3/2}\right)$ and therefore $\ddot{x}_\beta=O_p\left(n^{-1/2}T^{1/2}\right)$. This follows from $\sum_{t=1}^T x_{it}=O_p\left(T^{3/2}\right)$. Also,

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

The x_{it}s have zero mean for all t since their DGP has no drift term. Thus, $\sum_{i=1}^n\sum_{t=1}^T x_{it}=O_p\left(\sqrt{n}T^{3/2}\right)$ and $\ddot{x}=O_p\left(n^{-1/2}T^{1/2}\right)$. This entails that $II=O_p\left(n^{-1}\right)$. Consider IV; from Assumption 2

$$\begin{aligned}
E\left(-\frac{1}{2}IV\right) &= \frac{1}{n}\sigma_\beta^2E\left[\ddot{x}_\beta\left(\frac{1}{T^2}\sum_{t=1}^T \tilde{x}_{it}\right)\left(\frac{1}{T^2}\sum_{t=1}^T x_{it}\tilde{x}_{it}\right)\right] \\
&= \frac{1}{n}O\left(\sqrt{\frac{T}{n}}\frac{1}{\sqrt{T}}\right)=O\left(\frac{1}{n^{3/2}}\right).
\end{aligned}$$

Finally, Assumption 2 entails that $E\left(V\right)=E\left(VI\right)=0$ for all n and T. As far as I is concerned, note that as $T\to\infty$, $I\xrightarrow{d}\left(\beta_i-\beta\right)^2\left(\sigma_x^2\int W^2\right)^2$. It holds that $E\left[\left(\int W^2\right)^2\Big|\,C\right]=E\left(\int W^2\right)^2=Var\left(\int W^2\right)+\left[E\left(\int W^2\right)\right]^2$. Since $E\left(\int W^2\right)=1/2$ and Lemma 14 in Phillips and Moon (1999a) ensures that $Var\left(\int W^2\right)=7/12$, we get $E\left(I\right)=5\sigma_\beta^2\sigma_x^4/6$. Considering III, note

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

which is $O_p\left(1\right)$ for $i=1,...,\lfloor n\lambda\rfloor$ and $o_p\left(1\right)$ for the other units. Similarly to above,

$\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{e}_{it}^{u} = \sum_{i=1}^{n}\sum_{t=1}^{T}x_{it}e_{it}^{u} - nT\ddot{x}\ddot{e}^{u}$ with $\ddot{x}$ and $\ddot{e}^{u}$ both $O_p\left(n^{-1/2}T^{1/2}\right)$. Thus, as $T\to\infty$, $E\left(a^2\right)=\lambda\sigma_u^2\sigma_x^2/2$. As far as b is concerned, define $\bar{F}=T^{-1}\sum_{t=1}^{T}F_t$; then

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

Noting that $E\left(\tilde{x}_{it}\tilde{x}_{is}\right)=\sigma_x^2\min\left(t,s\right)$ and following similar passages as above, it follows that $n^{-1}\sum_{i=1}^{n}E\left(\tilde{\xi}_{iT}^2\middle| C\right)\xrightarrow{d}\lambda\sigma_F^2\sigma_u^2 D_1'$. Last

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

since $E\left(x_{it}\tilde{x}_{is}e_{it}^{u}\right)=O_p\left(T^{-1/2}\right)$ for all i, t and s. Combining all the above, (10) follows; (11) follows from similar passages as above, and the proof is thus omitted to save space. ■

Proof of Lemma 2. The proof is very similar to that of Lemma 1, and it is based on applying the MDS CLT after cross-sectional independence is achieved through conditioning. Many passages are therefore omitted. Consider first (12); as above,

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

Consider the denominator of this expression. In light of (7), it holds that $\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}^2 = \left(\sum_{i=1}^{n}\vartheta_i^2\right)\left(\sum_{t=1}^{T}\bar{G}_t^2\right)$. Note that, as $(n,T)\to\infty$, this is the product of two limits, not a joint limit. The LLN entails that $\sum_{i=1}^{n}\vartheta_i^2=O_p\left(n\right)$, and, by definition, $n^{-1}\sum_{i=1}^{n}\vartheta_i^2\xrightarrow{p}\vartheta^2$. Also, applying the FCLT yields $\sum_{t=1}^{T}\bar{G}_t^2=O_p\left(T^2\right)$ with $T^{-2}\sum_{t=1}^{T}\bar{G}_t^2\xrightarrow{d}\sigma_G^2\int\bar{W}_2^2$. Hence, $\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_{it}^2\xrightarrow{d}\sigma_G^2\vartheta^2\int\bar{W}_2^2$. Turning to the numerator, define, as above, $\xi_{iT} = T^{-2}\sum_{t=1}^{T}\left[\left(\beta_i-\beta\right)\bar{x}_{it}+\bar{u}_{it}\right]\bar{x}_{it}$, and let C^* be the σ-field generated by $\left\{F_t\right\}_{t=1}^{T}\cup\left\{G_t\right\}_{t=1}^{T}$. As C in the proof of Lemma 1, C^* is an invariant σ-field. The sequence ξ_{iT} has (conditional on

C^*) expected value given by, using (7)

$$E\left(\xi_{iT}\right|C^*) = E\left[\frac{1}{T^2}\sum_{t=1}^{T}\left(\beta_i-\beta\right)\bar{x}_{it}^2\right|C^*\left.\right] + E\left[\frac{1}{T^2}\sum_{t=1}^{T}\bar{x}_{it}\bar{e}_{it}^u\right|C^*\left.\right] \tag{24}$$
$$+E\left[\frac{1}{T^2}\sum_{t=1}^{T}\varphi_i\vartheta_i\bar{F}_t\bar{G}_t\right|C^*\left.\right]$$
$$= I+II+III.$$

Assumption 2 ensures that $I=0$ for all T. Similar considerations as for III in (19) entail $II=O_p\left(T^{-1/2}\right)$. Finally, consider $III=\left(T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t\right)E\left(\varphi_i\vartheta_i\right)$. Under Assumption 3*(i)*(a), since $T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t=O_p(1)$ for some units i (those where the common factors F_t are nonstationary), $III=O_p\left(T^{-1/2}\right)$; alternatively, if $E\left(\varphi_i\vartheta_i\right)$ is strictly not equal to zero, Assumption 3*(i)*(b) entails $T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t=O_p\left(T^{-1}\right)$ using the FCLT. In either case, as $T\to\infty$, III is bounded by $O_p\left(T^{-1/2}\right)$. Thus, defining $\bar{\xi}_{iT}=\xi_{iT}-E\left(\xi_{iT}\right)$ and $I_i^*=C^*\cup\left\{\bar{\xi}_{1T},...,\bar{\xi}_{iT}\right\}$, $\left\{\bar{\xi}_{iT},I_i^*;i\geq i\right\}$ is a zero mean MDS. Similar passages as in the proof of Lemma 1 would lead to show that, under Assumptions 1-3, a Liapunov condition holds whereby $E\left|\bar{\xi}_{iT}\right|C^*\left.\right|^{2+\delta}<\infty$. Thus, applying the MDS CLT, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$, $n^{-1/2}\sum_{i=1}^{n}\xi_{iT}=n^{-1/2}\sum_{i=1}^{n}\bar{\xi}_{iT}+n^{-1/2}\sum_{i=1}^{n}E\left(\xi_{iT}\right)=n^{-1/2}\sum_{i=1}^{n}\bar{\xi}_{iT}+O_p\left(\sqrt{\frac{n}{T}}\right)$ $\xrightarrow{d}\left[E\left(\xi_{iT}^2\right|C^*)\right]^{1/2}\times Z$. In order to calculate $E\left(\xi_{iT}^2\right|C^*)$, consider (21). Even under Assumption 3, the expectations of d, e and f are negligible. As far as a, b and c are concerned, note that $a=\left(\beta_i-\beta\right)^2\vartheta_i^4\left(T^{-2}\sum_{t=1}^{T}\bar{G}_t^2\right)^2$; as $T\to\infty$, the FCLT entails $a\xrightarrow{d}\left(\beta_i-\beta\right)^2\sigma_G^4\vartheta_i^4\left(\int\bar{W}_2^2\right)^2$; conditional on C^*, $E(a)=\sigma_\beta^2\sigma_G^4\vartheta^4\left(\int\bar{W}_2^2\right)^2$. As far as b is concerned, note $b=\left(T^{-2}\sum_{t=1}^{T}\vartheta_i\bar{G}_t\bar{e}_{it}^u\right)^2$ the passages are essentially the same as for the proof of (22), although only $\lfloor n\lambda\rfloor$ units have a non-negligible contribution; thus $b\xrightarrow{d}\sigma_u^2\sigma_G^2\vartheta^2 D_2$ for $i=1,...,\lfloor n\lambda\rfloor$. Finally, consider $c=\left(\varphi_i\vartheta_i\right)^2\left(T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t\right)^2$; the FCLT entails that c is negligible under Assumption 3*(i)*(b). If Assumption 3*(i)*(a) holds, for $i=1,...,\lfloor n\lambda\rfloor$, from $T^{-2}\sum_{t=1}^{T}\bar{F}_t\bar{G}_t\xrightarrow{d}\sigma_F\sigma_G\int\bar{W}_1\bar{W}_2$, so that $c\xrightarrow{d}\left(\varphi_i\vartheta_i\right)^2\sigma_F^2\sigma_G^2\left(\int\bar{W}_1\bar{W}_2\right)^2$. Putting all together, it follows that

$$\lim_{n\to\infty}E\left(\xi_{iT}\right|C^*)=\sigma_\beta^2\sigma_G^4\vartheta^4\left(\int\bar{W}_2^2\right)^2+\lambda\sigma_u^2\sigma_G^2\vartheta^2D_2+\lambda\sigma_F^2\sigma_G^2\kappa^2\left(\int\bar{W}_1\bar{W}_2\right)^2,$$

where the last term is present only under Assumption 3*(i)*(a). Turning to (13), the numerator has the same expansion as in (23); again, similar arguments as in the proof of (9) would show that the only terms which are asymptotically non-negligible are I, II and III. The presence of common factors in the DGP of x_{it} does not affect the limits of I and II, and therefore $I \xrightarrow{p} \lambda\sigma_u^2/6$ and, as long as Assumption 3*(i)*(a) holds, $II \xrightarrow{p} \lambda\varphi^2\sigma_F^2 \int \bar{W}_1^2$. As far as III is concerned, note that $III = \left[n^{-1}\sum_{i=1}^n (\beta_i - \beta)^2 \vartheta_i^2\right] \left(T^{-2}\sum_{t=1}^T \bar{G}_t^2\right)$; as $(n,T) \to \infty$ this is the product of two limits, and the LLN and the FCLT yield $III \xrightarrow{d} \sigma_\beta^2\sigma_G^2\vartheta^2 \left(\int \bar{W}_2^2\right)$. Putting all together, (13) follows. ■

Proof of Theorem 1. In order to prove (16), it is sufficient to show that, under H_0, $\sqrt{n}\widetilde{se}\left(\hat{\beta}^{LSDV}\right)$ converges to the standard deviation of $\sqrt{n}\left(\hat{\beta}^{LSDV} - \beta\right)$. Consider the numerator of (14) and let $\hat{\xi}_{iT} = T^{-2}\sum_{t=1}^T \bar{x}_{it}\hat{v}_{it}$; after some algebra it holds that $\hat{\xi}_{iT} = \xi_{iT} - \left(\hat{\beta}^{LSDV} - \beta\right) T^{-2}\sum_{t=1}^T \bar{x}_{it}^2$. Therefore

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

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

$$\frac{1}{n}\sum_{i=1}^n \hat{\xi}_{iT}^2 = \frac{1}{n}\sum_{i=1}^n \xi_{iT}^2 + o_p(1). \tag{25}$$

Consider $n^{-1}\sum_{i=1}^n \xi_{iT}^2$. Conditioning on C, ξ_{iT} is an *i.i.d.* sequence for all T; thus, a cross-sectional LLN can be applied such that

$$\frac{1}{n}\sum_{i=1}^n \xi_{iT}^2 \xrightarrow{p} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E\left(\xi_{iT}^2 \middle| C\right);$$

thus, in light of (25), $n^{-1}\sum_{i=1}^n \hat{\xi}_{iT}^2 \xrightarrow{p} \lim_{n\to\infty} n^{-1}\sum_{i=1}^n E\left(\xi_{iT}^2 \middle| C\right)$, which is the variance of the numerator of $\sqrt{n}\left(\hat{\beta}^{LSDV} - \beta\right)$ as reported in equation (20). Note that the fact that $T \to \infty$ is just an aside, as what ensures consistency of $\sqrt{n}\widetilde{se}\left(\hat{\beta}^{LSDV}\right)$ is the validity of the cross-sectional LLN. Essentially the same calculations would show that the same results hold under (7).

To prove (17), two results need to be shown under $H_A^{(n)}$. First, that $\sqrt{n}\left(\hat{\beta}^{LSDV} - \beta_0\right) \xrightarrow{d}$

$N(c, V^2)$, where V is given in (8). Second, that $\sqrt{n}\widetilde{se}\left(\hat{\beta}^{LSDV}\right) \xrightarrow{d} V$. 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 (8) it holds that, under $H_A^{(n)}$, $\sqrt{n}\left(\hat{\beta}^{LSDV} - \beta_0\right) \xrightarrow{d} N(0, V^2) + c \sim N(c, V^2)$. As far as $\sqrt{n}\widetilde{se}\left(\hat{\beta}^{LSDV}\right)$ is concerned, consider (23). Since $\widetilde{se}\left(\hat{\beta}^{LSDV}\right)$ 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 $\sqrt{n}\widetilde{se}\left(\hat{\beta}^{LSDV}\right) \xrightarrow{d} V$. Equation (17) follows. The proof for the POLS estimator follows essentially the same lines and therefore it is omitted. ■

Appendix B: Tables

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.025	0.067	0.025	0.041
	0.25		0.050	0.052	0.061	0.028
	0.5		0.040	0.041	0.051	0.026
	4		0.027	0.026	0.030	0.024
0.2	0		0.032	0.057	0.033	0.039
	0.25		0.071	0.074	0.072	0.037
	0.5		0.064	0.067	0.074	0.029
	4		0.037	0.037	0.055	0.024
0.4	0		0.055	0.059	0.054	0.057
	0.25		0.068	0.067	0.066	0.044
	0.5		0.064	0.064	0.069	0.039
	4		0.046	0.044	0.057	0.035
0.6	0		0.073	0.074	0.071	0.072
	0.25		0.057	0.057	0.065	0.015
	0.5		0.047	0.045	0.060	0.009
	4		0.015	0.015	0.028	0.008
0.8	0		0.085	0.085	0.085	0.087
	0.25		0.069	0.071	0.076	0.040
	0.5		0.067	0.067	0.074	0.036
	4		0.042	0.042	0.051	0.026
1	0		0.074	0.074	0.074	0.077
	0.25		0.072	0.072	0.068	0.090
	0.5		0.072	0.072	0.072	0.084
	4		0.089	0.088	0.085	0.081

Table 1a: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T)=(25,100)$ - case with no common factors in u_{it} or x_{it}.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.007	0.041	0.007	0.007
	0.25		0.057	0.038	0.049	0.035
	0.5		0.056	0.046	0.047	0.034
	4		0.032	0.030	0.030	0.031
0.2	0		0.044	0.056	0.048	0.042
	0.25		0.048	0.039	0.040	0.040
	0.5		0.038	0.039	0.032	0.037
	4		0.035	0.035	0.039	0.035
0.4	0		0.052	0.054	0.052	0.052
	0.25		0.037	0.038	0.048	0.047
	0.5		0.038	0.029	0.039	0.046
	4		0.038	0.038	0.039	0.037
0.6	0		0.065	0.068	0.075	0.068
	0.25		0.064	0.065	0.068	0.040
	0.5		0.057	0.056	0.066	0.030
	4		0.028	0.030	0.028	0.028
0.8	0		0.062	0.066	0.062	0.063
	0.25		0.060	0.062	0.066	0.031
	0.5		0.054	0.055	0.063	0.026
	4		0.040	0.039	0.046	0.041
1	0		0.073	0.073	0.076	0.073
	0.25		0.069	0.069	0.069	0.071
	0.5		0.074	0.075	0.066	0.069
	4		0.078	0.078	0.077	0.078

Table 1b: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T)=(50,200)$ - case with no common factors in u_{it} or x_{it}.

λ	σ_β^2	(ρ, θ)	$(0, 0)$	$(0.75, 0)$	$(0, 0.75)$	$(0, -0.75)$
0	0		0.009	0.029	0.009	0.009
	0.25		0.055	0.046	0.055	0.053
	0.5		0.052	0.046	0.054	0.052
	4		0.044	0.048	0.044	0.042
0.2	0		0.048	0.047	0.046	0.049
	0.25		0.053	0.048	0.047	0.048
	0.5		0.051	0.046	0.051	0.053
	4		0.044	0.048	0.044	0.044
0.4	0		0.057	0.057	0.056	0.056
	0.25		0.042	0.044	0.044	0.046
	0.5		0.056	0.057	0.056	0.055
	4		0.048	0.051	0.048	0.047
0.6	0		0.053	0.055	0.055	0.055
	0.25		0.052	0.055	0.056	0.044
	0.5		0.055	0.057	0.056	0.050
	4		0.045	0.042	0.045	0.043
0.8	0		0.056	0.055	0.054	0.054
	0.25		0.055	0.054	0.054	0.056
	0.5		0.058	0.056	0.057	0.057
	4		0.048	0.048	0.049	0.046
1	0		0.060	0.061	0.061	0.060
	0.25		0.058	0.058	0.056	0.057
	0.5		0.061	0.061	0.061	0.063
	4		0.055	0.055	0.057	0.054

Table 1c: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n, T) = (100, 400)$ - case with no common factors in u_{it} or x_{it}.

λ	σ^2_β	(ρ, θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		1.000	1.000	0.997	1.000
	0.25		1.000	0.998	0.993	1.000
	0.5		0.986	0.979	0.996	1.000
	4		0.253	0.539	0.484	0.585
0.2	0		0.999	1.000	0.976	1.000
	0.25		0.998	0.996	0.988	1.000
	0.5		1.000	0.995	0.939	1.000
	4		0.489	0.898	0.857	0.427
0.4	0		0.999	0.998	0.916	1.000
	0.25		0.990	0.996	0.845	1.000
	0.5		0.996	0.994	0.851	0.999
	4		0.348	0.810	0.418	0.866
0.6	0		0.993	0.997	0.842	1.000
	0.25		0.993	0.982	0.839	1.000
	0.5		0.984	0.996	0.675	0.995
	4		0.374	0.687	0.451	0.110
0.8	0		0.986	0.985	0.738	1.000
	0.25		0.988	0.965	0.644	1.000
	0.5		0.996	0.988	0.660	0.970
	4		0.569	0.765	0.423	0.409
1	0		0.974	0.980	0.673	1.000
	0.25		0.870	0.956	0.630	1.000
	0.5		0.935	0.892	0.503	1.000
	4		0.934	0.491	0.611	0.972

Table 1d: empirical power $\tilde{t}_\beta$ for $H_0 : \beta = 1$ under $H_A : \beta = 2$. Sample size $(n, T) = (25, 100)$ - case with no common factors in u_{it} or x_{it}.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.025	0.064	0.026	0.044
	0.25		0.050	0.052	0.061	0.028
	0.5		0.040	0.041	0.051	0.026
	4		0.027	0.026	0.030	0.024
0.2	0		0.035	0.048	0.032	0.026
	0.25		0.071	0.073	0.072	0.037
	0.5		0.064	0.067	0.074	0.029
	4		0.050	0.050	0.060	0.046
0.4	0		0.060	0.059	0.060	0.066
	0.25		0.068	0.067	0.067	0.044
	0.5		0.064	0.064	0.069	0.039
	4		0.059	0.057	0.066	0.051
0.6	0		0.070	0.071	0.073	0.076
	0.25		0.057	0.057	0.065	0.015
	0.5		0.047	0.045	0.060	0.009
	4		0.029	0.030	0.038	0.026
0.8	0		0.079	0.077	0.076	0.086
	0.25		0.069	0.071	0.076	0.040
	0.5		0.067	0.068	0.074	0.036
	4		0.045	0.046	0.050	0.042
1	0		0.077	0.078	0.081	0.074
	0.25		0.072	0.072	0.068	0.089
	0.5		0.072	0.072	0.072	0.084
	4		0.093	0.094	0.086	0.094

Table 2a: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T) = (25,100)$ - case with common factors in u_{it}, $\left(\bar{\varphi},\sigma^2_\varphi\right) = (1,1)$.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.009	0.026	0.009	0.002
	0.25		0.037	0.040	0.031	0.046
	0.5		0.047	0.038	0.037	0.035
	4		0.031	0.031	0.051	0.051
0.2	0		0.049	0.046	0.046	0.050
	0.25		0.038	0.037	0.039	0.043
	0.5		0.034	0.036	0.049	0.048
	4		0.042	0.042	0.046	0.041
0.4	0		0.055	0.056	0.060	0.056
	0.25		0.034	0.043	0.045	0.036
	0.5		0.038	0.033	0.042	0.047
	4		0.040	0.041	0.039	0.041
0.6	0		0.064	0.059	0.063	0.059
	0.25		0.058	0.062	0.062	0.056
	0.5		0.058	0.059	0.059	0.050
	4		0.044	0.045	0.050	0.037
0.8	0		0.073	0.072	0.063	0.062
	0.25		0.053	0.059	0.061	0.053
	0.5		0.054	0.053	0.059	0.048
	4		0.037	0.037	0.047	0.040
1	0		0.069	0.069	0.070	0.068
	0.25		0.075	0.074	0.066	0.071
	0.5		0.076	0.077	0.067	0.074
	4		0.072	0.072	0.075	0.072

Table 2b: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T) = (50,200)$ - case with common factors in u_{it}, $\left(\bar{\varphi}, \sigma^2_\varphi\right) = (1,1)$.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.006	0.012	0.009	0.002
	0.25		0.047	0.047	0.051	0.046
	0.5		0.049	0.049	0.046	0.045
	4		0.051	0.051	0.051	0.051
0.2	0		0.048	0.049	0.049	0.046
	0.25		0.051	0.053	0.050	0.050
	0.5		0.050	0.051	0.048	0.047
	4		0.047	0.046	0.048	0.044
0.4	0		0.054	0.055	0.054	0.058
	0.25		0.044	0.042	0.041	0.042
	0.5		0.053	0.054	0.052	0.052
	4		0.047	0.048	0.046	0.047
0.6	0		0.054	0.051	0.053	0.053
	0.25		0.051	0.053	0.050	0.050
	0.5		0.057	0.053	0.055	0.054
	4		0.047	0.045	0.046	0.048
0.8	0		0.056	0.056	0.056	0.056
	0.25		0.055	0.055	0.055	0.053
	0.5		0.056	0.056	0.054	0.055
	4		0.049	0.051	0.048	0.051
1	0		0.062	0.060	0.061	0.061
	0.25		0.054	0.057	0.056	0.058
	0.5		0.061	0.061	0.061	0.059
	4		0.054	0.054	0.055	0.053

Table 2c: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T)=(100,400)$ - case with common factors in u_{it}, $\left(\bar{\varphi},\sigma^2_\varphi\right)=(1,1)$.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		1.000	1.000	1.000	1.000
	0.25		1.000	1.000	1.000	1.000
	0.5		1.000	0.967	1.000	1.000
	4		0.100	0.755	0.903	0.987
0.2	0		1.000	1.000	1.000	1.000
	0.25		1.000	1.000	1.000	1.000
	0.5		1.000	1.000	1.000	1.000
	4		0.592	1.000	0.990	0.687
0.4	0		1.000	1.000	1.000	1.000
	0.25		1.000	1.000	1.000	1.000
	0.5		1.000	1.000	1.000	1.000
	4		0.404	0.611	0.724	0.990
0.6	0		1.000	1.000	1.000	1.000
	0.25		1.000	1.000	0.993	1.000
	0.5		1.000	1.000	0.998	1.000
	4		0.308	0.513	0.824	0.584
0.8	0		1.000	1.000	0.993	1.000
	0.25		1.000	1.000	0.999	1.000
	0.5		1.000	1.000	1.000	0.999
	4		0.881	0.960	0.459	0.601
1	0		1.000	1.000	0.997	1.000
	0.25		1.000	1.000	0.999	1.000
	0.5		1.000	1.000	0.997	1.000
	4		0.998	0.607	1.000	0.822

Table 2d: empirical power $\tilde{t}_\beta$ for $H_0 : \beta = 1$ under $H_A : \beta = 2$. Sample size $(n, T) = (25, 100)$ - case with common factors in u_{it}, $\left(\bar{\varphi}, \sigma^2_\varphi\right) = (1, 1)$.

λ	σ^2_β	(ρ, θ)	$(0, 0)$	$(0.75, 0)$	$(0, 0.75)$	$(0, -0.75)$
0	0		0.023	0.027	0.026	0.054
	0.25		0.100	0.017	0.046	0.046
	0.5		0.099	0.143	0.044	0.013
	4		0.093	0.170	0.019	0.008
0.2	0		0.030	0.038	0.030	0.043
	0.25		0.092	0.054	0.066	0.040
	0.5		0.092	0.122	0.029	0.078
	4		0.074	0.013	0.067	0.013
0.4	0		0.063	0.065	0.064	0.075
	0.25		0.041	0.100	0.048	0.004
	0.5		0.093	0.091	0.030	0.076
	4		0.098	0.047	0.021	0.001
0.6	0		0.077	0.068	0.077	0.071
	0.25		0.047	0.090	0.043	0.028
	0.5		0.026	0.041	0.108	0.005
	4		0.047	0.022	0.017	0.001
0.8	0		0.089	0.078	0.075	0.081
	0.25		0.067	0.070	0.082	0.026
	0.5		0.090	0.040	0.120	0.047
	4		0.095	0.027	0.019	0.001
1	0		0.089	0.091	0.082	0.068
	0.25		0.067	0.067	0.069	0.033
	0.5		0.049	0.057	0.039	0.021
	4		0.019	0.045	0.068	0.082

Table 3a: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n, T) = (25, 100)$ - case with common factors in x_{it}, $\left(\bar{\vartheta}, \sigma^2_\vartheta\right) = (1, 1)$.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.007	0.004	0.006	0.026
	0.25		0.073	0.009	0.004	0.001
	0.5		0.015	0.010	0.007	0.007
	4		0.006	0.085	0.020	0.098
0.2	0		0.038	0.041	0.046	0.052
	0.25		0.082	0.017	0.090	0.079
	0.5		0.076	0.010	0.044	0.039
	4		0.069	0.063	0.015	0.050
0.4	0		0.067	0.061	0.059	0.057
	0.25		0.022	0.064	0.097	0.026
	0.5		0.109	0.019	0.043	0.010
	4		0.042	0.072	0.101	0.004
0.6	0		0.055	0.061	0.075	0.059
	0.25		0.094	0.022	0.089	0.091
	0.5		0.090	0.031	0.065	0.008
	4		0.020	0.019	0.010	0.011
0.8	0		0.065	0.087	0.062	0.071
	0.25		0.039	0.042	0.040	0.065
	0.5		0.076	0.057	0.027	0.064
	4		0.015	0.092	0.050	0.055
1	0		0.057	0.060	0.069	0.065
	0.25		0.067	0.061	0.063	0.043
	0.5		0.057	0.041	0.061	0.052
	4		0.086	0.033	0.032	0.035

Table 3b: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T) = (50,200)$ - case with common factors in x_{it}, $\left(\bar{\vartheta},\sigma^2_\vartheta\right) = (1,1)$.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		0.004	0.007	0.007	0.012
	0.25		0.020	0.011	0.014	0.041
	0.5		0.010	0.023	0.020	0.029
	4		0.003	0.018	0.016	0.001
0.2	0		0.003	0.018	0.016	0.001
	0.25		0.047	0.055	0.061	0.046
	0.5		0.063	0.012	0.061	0.036
	4		0.027	0.026	0.034	0.060
0.4	0		0.009	0.008	0.011	0.001
	0.25		0.046	0.054	0.054	0.052
	0.5		0.061	0.040	0.042	0.061
	4		0.036	0.038	0.033	0.038
0.6	0		0.058	0.061	0.054	0.050
	0.25		0.041	0.042	0.060	0.043
	0.5		0.036	0.041	0.061	0.043
	4		0.029	0.052	0.034	0.038
0.8	0		0.057	0.058	0.060	0.052
	0.25		0.043	0.050	0.060	0.046
	0.5		0.046	0.062	0.047	0.061
	4		0.041	0.033	0.037	0.046
1	0		0.052	0.059	0.060	0.056
	0.25		0.060	0.040	0.046	0.050
	0.5		0.050	0.041	0.067	0.040
	4		0.043	0.057	0.037	0.042

Table 3c: empirical rejection frequencies for $\tilde{t}_\beta$ at a nominal size 5%. Sample size $(n,T)=(100,400)$ - case with common factors in x_{it}, $\left(\bar{\vartheta},\sigma^2_\vartheta\right)=(1,1)$.

λ	σ^2_β	(ρ,θ)	$(0,0)$	$(0.75,0)$	$(0,0.75)$	$(0,-0.75)$
0	0		1.000	1.000	1.000	1.000
	0.25		1.000	1.000	1.000	1.000
	0.5		0.999	1.000	0.994	1.000
	4		0.119	0.898	0.503	0.555
0.2	0		1.000	1.000	1.000	1.000
	0.25		1.000	0.995	1.000	1.000
	0.5		0.998	1.000	0.998	1.000
	4		0.871	0.706	0.700	0.275
0.4	0		1.000	1.000	1.000	1.000
	0.25		1.000	0.998	1.000	1.000
	0.5		0.995	0.972	0.998	1.000
	4		0.956	0.703	0.463	0.445
0.6	0		1.000	1.000	1.000	1.000
	0.25		1.000	0.992	1.000	1.000
	0.5		0.988	0.983	0.967	0.999
	4		0.760	0.228	0.741	0.679
0.8	0		0.999	0.998	1.000	1.000
	0.25		0.993	0.996	0.995	1.000
	0.5		0.914	0.962	0.991	1.000
	4		0.947	0.574	0.653	0.828
1	0		1.000	1.000	0.999	1.000
	0.25		0.999	0.997	0.997	1.000
	0.5		0.994	0.973	0.989	1.000
	4		0.773	0.725	0.830	0.605

Table 3d: empirical power $\tilde{t}_\beta$ for $H_0 : \beta = 1$ under $H_A : \beta = 2$. Sample size $(n,T) = (25,100)$ - case with common factors in x_{it}, $\left(\bar{\vartheta}, \sigma^2_\vartheta\right) = (1,1)$.