A Supplement to "Panel Unit Root Tests in the Presence of a Multifactor Error Structure" by M.H. Pesaran, L.V. Smith and T. Yamagata (2012)

17 October 2012

This supplement provides proofs of the main theoretical results in Pesaran, Smith and Yagamata (2012, PSY) for the case of models with linear trends, and models with intercepts and serially correlated idiosyncratic errors. It also provides theoretical results for the cross sectionally augmented Sargan-Bhargava statistics, gives the details of a number of different panel unit root tests used in the empirical application, and provides comparative Monte Carlo results of the proposed tests and other panel unit root tests. This supplement should be consulted in conjunction with the paper.

S1 Proof of Theorem 2.1 in PSY in the Case of Models with Linear Trends

Under the unit root null hypothesis we have

$$\mathbf{z}_{it} = \mathbf{z}_{i0} + \mathbf{A}_i\mathbf{d}_t + \mathbf{\Gamma}_i\mathbf{s}_{ft} + \mathbf{s}_{it}, \tag{S1}$$

where $\mathbf{s}_{it} = (s_{iyt}, \mathbf{s}'_{ixt})'$, $s_{iyt} = \sum_{s=1}^{t}\varepsilon_{iys}$, $\mathbf{s}_{ixt} = \sum_{s=1}^{t}\boldsymbol{\varepsilon}_{ixs}$ and $\mathbf{s}_{ft} = \sum_{s=1}^{t}\mathbf{f}_s$. For $\mathbf{d}_t = (1,t)'$, (recall that we define $\mathbf{d}_0 \equiv \mathbf{0}$ and $\Delta\mathbf{d}_1 = (0,1)'$) and partitioning the $(k+1)\times 2$ matrix $\mathbf{A}_i = (\boldsymbol{\alpha}_{i0}, \boldsymbol{\alpha}_{i1})$ conformably with $\mathbf{d}_t$ from (S1) we have that

$$\mathbf{z}_{it} = \mathbf{z}_{i0} + \boldsymbol{\alpha}_{i0} + \boldsymbol{\alpha}_{i1}t + \mathbf{\Gamma}_i\mathbf{s}_{ft} + \mathbf{s}_{it},\ t = 1,2,...,T. \tag{S2}$$

Averaging (S2) across i we obtain

$$\bar{\mathbf{z}}_t = \bar{\mathbf{z}}_0 + \bar{\boldsymbol{\alpha}}_0 + \bar{\boldsymbol{\alpha}}_1 t + \bar{\mathbf{\Gamma}}\mathbf{s}_{ft} + \bar{\mathbf{s}}_t,\ t = 1,2,...,T. \tag{S3}$$

Under the null hypothesis writing (S2) in matrix notation, we have

$$\Delta\mathbf{Z}_i = \boldsymbol{\tau}_T\boldsymbol{\alpha}'_{i1} + \mathbf{F}\mathbf{\Gamma}'_i + \mathbf{E}_i, \tag{S4}$$

where $\mathbf{E}_i = (\boldsymbol{\varepsilon}_{i1}, \boldsymbol{\varepsilon}_{i2}, ..., \boldsymbol{\varepsilon}_{iT})'$, and $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{iyt}, \boldsymbol{\varepsilon}'_{ixt})'$. Similarly, we can write (S3) as

$$\Delta\bar{\mathbf{Z}} = \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}'_1 + \mathbf{F}\bar{\mathbf{\Gamma}}' + \bar{\mathbf{E}}, \tag{S5}$$

where $\bar{\mathbf{Z}} = \mathbf{N}^{-1}\sum_{\mathbf{i=1}}^{\mathbf{N}}\mathbf{Z}_i$, $\bar{\mathbf{E}} = N^{-1}\sum_{i=1}^{N}\mathbf{E}_i$, and etc. as in PSY. From (S4) and (S5) it follows, respectively, that

$$\mathbf{Z}_{i,-1} = \boldsymbol{\tau}_T\mathbf{z}'_{i0} + \mathbf{t}_{T-1}\boldsymbol{\alpha}'_{i1} + \mathbf{S}_{f,-1}\mathbf{\Gamma}'_i + \mathbf{S}_{i,-1},$$
$$\bar{\mathbf{Z}}_{-1} = \boldsymbol{\tau}_T\bar{\mathbf{z}}'_0 + \mathbf{t}_{T-1}\bar{\boldsymbol{\alpha}}'_1 + \mathbf{S}_{f,-1}\bar{\mathbf{\Gamma}}' + \bar{\mathbf{S}}_{-1}, \tag{S6}$$

where $\mathbf{t}_{T-1} = (0,1,...,T-1)'$. Recall thnder the null we have

$$\Delta\mathbf{y}_i = \boldsymbol{\tau}_T\tilde{\alpha}_{i1} + \Delta\bar{\mathbf{Z}}\boldsymbol{\delta}_i + \sigma_i\boldsymbol{\upsilon}_i, \tag{S7}$$

where

$$\tilde{\alpha}_{i1} = \alpha_{iy1} - \bar{\boldsymbol{\alpha}}'_1\boldsymbol{\delta}_i.$$

From (S7) it follows that

$$\mathbf{y}_{i,-1} = \boldsymbol{\tau}_T\mathring{y}_{i0} + \mathbf{t}_{T-1}\tilde{\alpha}_{i1} + \bar{\mathbf{Z}}_{-1}\boldsymbol{\delta}_i + \sigma_i\mathring{\mathbf{s}}_{i,-1}, \tag{S8}$$

where

$$\mathring{y}_{i0} = y_{i0} - \bar{\mathbf{z}}'_0\boldsymbol{\delta}_i,\ \mathring{\mathbf{s}}_{i,-1} = (\mathbf{s}_{iy,-1} - \bar{\mathbf{S}}_{-1}\boldsymbol{\delta}_i)/\sigma_i.$$

Now consider the augmented regression for testing the panel unit root hypothesis, which in the linear trend case is given by

$$\Delta y_{it} = g_{i0} + g_{i1}\mathbf{t}_{T-1} + b_i y_{it-1} + \mathbf{c}'_i\bar{\mathbf{z}}_{t-1} + \mathbf{h}'_i\Delta\bar{\mathbf{z}}_t + \epsilon_{it}. \tag{S9}$$

From (S7) and (S8), we have $\bar{\mathbf{M}}\boldsymbol{\Delta}\mathbf{y}_i = \bar{\mathbf{M}}\boldsymbol{\upsilon}_i$ and $\bar{\mathbf{M}}\mathbf{y}_{i,-1} = \bar{\mathbf{M}}\mathring{\mathbf{s}}_{i,-1}$, where $\bar{\mathbf{M}} = \mathbf{I}_T - \bar{\mathbf{W}}\left(\bar{\mathbf{W}}'\bar{\mathbf{W}}\right)^{-1}\bar{\mathbf{W}}'$, with $\bar{\mathbf{W}} = \left(\Delta\bar{\mathbf{Z}}, \boldsymbol{\tau}_T, \bar{\mathbf{Z}}_{-1}, \mathbf{t}_{T-1}\right)$. Note that $\mathbf{t}_{T-1}$ in (S9) could be replaced by $\mathbf{t}_T = \mathbf{t}_{T-1} + \boldsymbol{\tau}_T = (1, 2, ..., T)'$, without loss of generality, since $\bar{\mathbf{M}}\mathbf{t}_T = \mathbf{0}$ because $\bar{\mathbf{W}}$ contains both $\boldsymbol{\tau}_T$ and $\mathbf{t}_{T-1}$. However, to be consistent with (S2) we use $\mathbf{t}_{T-1}$. Then the t-ratio of b_i is given by

$$t_i(N,T) = \frac{\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\mathring{\mathbf{s}}_{i,-1}}{T}}{\left(\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i}{T-2k-5}\right)^{1/2}\left(\frac{\mathring{\mathbf{s}}_{i,-1}'\bar{\mathbf{M}}\mathring{\mathbf{s}}_{i,-1}}{T^2}\right)^{1/2}}. \tag{S10}$$

Theorem S1.1 *Suppose the series $\mathbf{z}_{it}$, for $i = 1, 2, ..., N$, $t = 1, 2, ..., T$, is generated under (5) according to (11) and $\mathbf{d}_t = (1, t)'$. Then under Assumptions 1-5 in PSY and as N and $T \to \infty$, such that $\sqrt{T}/N \to 0$, $t_i(N,T)$ given by (S10) has the same sequential $(N \to \infty, T \to \infty)$ and joint $[(N,T)_j \to \infty]$ limit distribution, is free of nuisance parameters, and is given by*

$$CADF_i = \frac{\int_0^1 W_i(r)dW_i(r) - \boldsymbol{\omega}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v}2}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}}{\left(\int_0^1 W_i^2(r)dr - \boldsymbol{\pi}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v}2}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}\right)^{1/2}}, \tag{S11}$$

where

$$\boldsymbol{\omega}_{i2\mathbf{v}} = \begin{pmatrix} W_i(1) \\ \int_0^1 [\mathbf{W}_{\mathbf{v}}(r)]\, dW_i(r) \\ W_i(1) - \int_0^1 W_i(r)dr \end{pmatrix}, \; \boldsymbol{\pi}_{i2\mathbf{v}} = \begin{pmatrix} \int_0^1 W_i(r)dr \\ \int_0^1 [\mathbf{W}_{\mathbf{v}}(r)]\, W_i(r)dr \\ \int_0^1 rW_i(r)dr \end{pmatrix},$$

$$\mathbf{G}_{\mathbf{v}2} = \begin{pmatrix} 1 & [\int_0^1 \mathbf{W}_{\mathbf{v}}(r)dr]' & 1/2 \\ \int_0^1 \mathbf{W}_{\mathbf{v}}(r)dr & \int_0^1 [\mathbf{W}_{\mathbf{v}}(r)]\,[\mathbf{W}_{\mathbf{v}}(r)]'\, dr & \int_0^1 r\mathbf{W}_{\mathbf{v}}(r)dr \\ 1/2 & \int_0^1 r\mathbf{W}_{\mathbf{v}}(r)'dr & 1/3 \end{pmatrix}$$

$W_i(r)$ is a scalar standard Brownian motion, and $\mathbf{W}_{\mathbf{v}}(r)$ is m^0-dimensional standard Brownian motion defined on [0,1] corresponding to ε_{iyt} and $\mathbf{v}_t$, respectively. $W_i(r)$ and $\mathbf{W}_{\mathbf{v}}(r)$ are mutually independent.

Proof. Let $\mathbf{W}_{f2} = (\mathbf{F}, \boldsymbol{\tau}_T, \mathbf{S}_{f,-1}, \mathbf{t}_{T-1})$ and $\bar{\boldsymbol{\Xi}}_2 = \left(\bar{\mathbf{E}}, \mathbf{0}_T, \bar{\mathbf{S}}_{-1}, \mathbf{0}_T\right)$, and note that $\bar{\mathbf{W}} = \left(\Delta\bar{\mathbf{Z}}, \boldsymbol{\tau}_T, \bar{\mathbf{Z}}_{-1}, \mathbf{t}_{T-1}\right)$ can be written as

$$\bar{\mathbf{W}}' = \mathbf{Q}_{2N}\mathbf{W}_{f2}' + \bar{\boldsymbol{\Xi}}_2', \text{ where } \underset{(2k+4)\times(2m^0+2)}{\mathbf{Q}_{2N}} = \begin{pmatrix} \bar{\boldsymbol{\Gamma}} & \bar{\boldsymbol{\alpha}}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{z}}_0 & \bar{\boldsymbol{\Gamma}} & \bar{\boldsymbol{\alpha}}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}. \tag{S12}$$

Expanding $\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\mathring{\mathbf{s}}_{i,-1}/T$ gives

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\mathring{\mathbf{s}}_{i,-1}}{T} = \frac{\boldsymbol{\upsilon}_i'\mathring{\mathbf{s}}_{i,-1}}{T} - \left(\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}\mathbf{B}_2\right)\left(\mathbf{B}_2\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{B}_2\right)^{-1}\left(\frac{\mathbf{B}_2\bar{\mathbf{W}}'\mathring{\mathbf{s}}_{i,-1}}{T}\right), \tag{S13}$$

where

$$\underset{(2k+4)\times(2k+4)}{\mathbf{B}_2} = \begin{pmatrix} \frac{1}{\sqrt{T}}\mathbf{I}_{k+2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{T}\mathbf{I}_{k+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{T^{3/2}} \end{pmatrix}.$$

Using Lemma A.1 together with the results in Proposition 17.1 of Hamilton (1994; p.486) we have

$$\begin{aligned} \frac{\mathring{\mathbf{s}}_{i,-1}'\boldsymbol{\upsilon}_i}{T^{3/2}} &= \frac{\mathbf{s}_{iy,-1}'\boldsymbol{\varepsilon}_{iy}}{\sigma_i^2 T^{3/2}} + O_p\left(\frac{1}{\sqrt{NT}}\right) \\ &\overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i(r)dW_i(r), \end{aligned} \tag{S14}$$

where $W_i(r)$ is a standard Brownian motion defined on [0,1], associated with ε_{iyt}. From (S12) it follows that

$$\mathbf{B}_2\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i = \mathbf{B}_2\mathbf{Q}_{N2}\mathbf{W}'_{f2}\boldsymbol{\upsilon} + \mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\boldsymbol{\upsilon}_i, \tag{S15}$$

$$\frac{\mathbf{B}_2\bar{\mathbf{W}}'\hat{\mathbb{s}}_{i,-1}}{T} = \frac{\mathbf{B}_2\mathbf{Q}_{N2}\mathbf{W}'_{f2}\hat{\mathbb{s}}_{i,-1}}{T} + \frac{\mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\hat{\mathbb{s}}_{i,-1}}{T},$$

$$\begin{aligned}\mathbf{B}_2\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{B}_2 =& \mathbf{B}_2\mathbf{Q}_{N2}\mathbf{W}'_{f2}\mathbf{W}_{f2}\mathbf{Q}'_{N2}\mathbf{B}_2 \\ &+\mathbf{B}_2\mathbf{Q}_{N2}\mathbf{W}'_{f2}\bar{\boldsymbol{\Xi}}_2\mathbf{B}_2+\mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\mathbf{W}_{f2}\mathbf{Q}'_{N2}\mathbf{B}_2+\mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\bar{\boldsymbol{\Xi}}_2\mathbf{B}_2.\end{aligned}$$

Using Lemma A.1, it is easily seen that, as $(T,N) \xrightarrow{j} \infty$ with $\sqrt{T}/N \to 0$,

$$\mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\boldsymbol{\upsilon}_i \overset{(N,T)_j}{\to} \mathbf{0},\ \frac{\mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\hat{\mathbb{s}}_{i,-1}}{T} \overset{(N,T)_j}{\to} \mathbf{0},\ \mathbf{B}_2\bar{\boldsymbol{\Xi}}'_2\bar{\boldsymbol{\Xi}}_2\mathbf{B}_2 \overset{(N,T)_j}{\to} \mathbf{0}, \text{ and } \mathbf{B}_2\mathbf{Q}_{N2}\mathbf{W}'_{f2}\bar{\boldsymbol{\Xi}}_2\mathbf{B}_2 \overset{(N,T)_j}{\to} \mathbf{0}. \tag{S16}$$

Under Assumptions 1-5 in PSY, following a similar derivation of Lemma A.1 in PSY, we have

$$\frac{\mathbf{t}'_T\bar{\mathbf{E}}}{T^{3/2}} = O_p\left(\frac{1}{\sqrt{N}}\right),\ \frac{\mathbf{t}'_T\bar{\mathbf{S}}_{-1}}{T^{5/2}} = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{S17}$$

Define

$$\mathbf{C}_2 = \begin{pmatrix} \frac{1}{\sqrt{T}}\mathbf{I}_{m^0+1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{T}\mathbf{I}_{m^0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{T^{3/2}} \end{pmatrix},$$

so that, using Lemma A.1 and the results in Proposition 17.1 and 18.1 of Hamilton (1994; p.486, p.547-8) such as

$$\frac{\mathbf{F}'\mathbf{t}_T}{T^{3/2}} \overset{T}{\Longrightarrow} \boldsymbol{\Lambda}_f\left[\mathbf{W}_{\mathbf{v}}(1) - \int_0^1 \mathbf{W}_{\mathbf{v}}(r)dr\right],\ \frac{\boldsymbol{\tau}'_T\mathbf{t}_T}{T^2} \overset{T}{\to} \frac{1}{2}, \frac{\mathbf{S}'_{f,-1}\mathbf{t}_T}{T^{5/2}} \overset{T}{\Longrightarrow} \boldsymbol{\Lambda}_f\int_0^1 r\mathbf{W}_{\mathbf{v}}(r)dr,\ \frac{\mathbf{t}'_T\mathbf{t}_T}{T^3} \overset{T}{\to} \frac{1}{3},$$

as $(T,N) \xrightarrow{j} \infty$ with $\sqrt{T}/N \to 0$ we have

$$\mathbf{B}_2\mathbf{Q}_{2N}\mathbf{W}'_{f2}\boldsymbol{\upsilon}_i = \mathbf{Q}_{2N}\mathbf{C}_2\mathbf{W}'_{f2}\boldsymbol{\upsilon}_i \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}_2\boldsymbol{\vartheta}_{i2f}, \tag{S18}$$

$$\frac{\mathbf{B}_2\mathbf{Q}_N\mathbf{W}'_f\hat{\mathbb{s}}_{i,-1}}{T} = \frac{\mathbf{Q}_N\mathbf{C}\mathbf{W}'_f\hat{\mathbb{s}}_{i,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}_2\boldsymbol{\kappa}_{i2f}, \tag{S19}$$

$$\mathbf{B}_2\mathbf{Q}_{N2}\mathbf{W}'_{f2}\mathbf{W}_{f2}\mathbf{Q}'_{2N}\mathbf{B}_2 = \mathbf{Q}_{2N}\mathbf{C}_2\mathbf{W}'_{f2}\mathbf{W}_{f2}\mathbf{C}_2\mathbf{Q}'_{2N} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}_2\boldsymbol{\Upsilon}_{f2}\mathbf{Q}'_2, \tag{S20}$$

where

$$\mathbf{Q}_2 = \underset{N\to\infty}{\text{plim}}\,\mathbf{Q}_{2N},\ \boldsymbol{\vartheta}_{i2f} = \begin{pmatrix} \boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\ \boldsymbol{\Lambda}^*_{f2}\boldsymbol{\omega}_{i2\mathbf{v}} \end{pmatrix},\ \boldsymbol{\kappa}_{i2f} = \begin{pmatrix} \mathbf{0}_{m^0} \\ \boldsymbol{\Lambda}^*_{f2}\boldsymbol{\pi}_{i2\mathbf{v}} \end{pmatrix},\ \boldsymbol{\Upsilon}_{f2} = \begin{pmatrix} \mathbf{I}_{m^0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^*_{f2}\mathbf{G}_{\mathbf{v}2}\boldsymbol{\Lambda}^{*\prime}_{f2} \end{pmatrix},\ \boldsymbol{\Lambda}^*_{f2} = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_f & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

$$\boldsymbol{\omega}_{i2\mathbf{v}} = \begin{pmatrix} W_i(1) \\ \int_0^1 [\mathbf{W}_{\mathbf{v}}(r)]\,dW_i(r) \\ W_i(1) - \int_0^1 W_i(r)dr \end{pmatrix},\ \boldsymbol{\pi}_{i2\mathbf{v}} = \begin{pmatrix} \int_0^1 W_i(r)dr \\ \int_0^1 [\mathbf{W}_{\mathbf{v}}(r)]\,W_i(r)dr \\ \int_0^1 rW_i(r)dr \end{pmatrix},$$

$$\mathbf{G}_{\mathbf{v}2} = \begin{pmatrix} 1 & [\int_0^1 \mathbf{W}_{\mathbf{v}}(r)dr]' & 1/2 \\ \int_0^1 \mathbf{W}_{\mathbf{v}}(r)dr & \int_0^1 [\mathbf{W}_{\mathbf{v}}(r)]\,[\mathbf{W}_{\mathbf{v}}(r)]'\,dr & \int_0^1 r\mathbf{W}_{\mathbf{v}}(r)dr \\ 1/2 & \int_0^1 r\mathbf{W}_{\mathbf{v}}(r)'dr & 1/3 \end{pmatrix},$$

$\boldsymbol{\Lambda}_f$ is defined by (3), $\mathbf{W}_{\mathbf{v},i}(1)$ is defined such that $T^{-1/2}\sum_{t=1}^T \mathbf{v}_t\varepsilon_{iyt}/\sigma_i \overset{T}{\Longrightarrow} \mathbf{W}_{\mathbf{v},i}(1)$, with $\mathbf{v}_t$ defined as in Assumption 2, $\mathbf{W}_{\mathbf{v}}(r)$ is an m^0-dimensional standard Brownian motion associated with $\mathbf{v}_t$ defined on [0,1], and

$W_i(r)$ is defined as above. These two groups of Brownian motions $(\mathbf{W}_{\mathbf{v}}(r), W_i(r))$ are independent of each other. Collecting the results from (S15) to (S20), as well as using Lemma A.2 (since $\mathbf{Q_2}$ has full column rank) we have

$$\left(\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}\mathbf{B}_2\right)\left(\mathbf{B}_2\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{B}_2\right)^{-1}\left(T^{-1}\mathbf{B}_2\bar{\mathbf{W}}'\hat{\mathbf{s}}_{i,-1}\right) \overset{(N,T)_j}{\Longrightarrow} \boldsymbol{\vartheta}_{i2f}'\mathbf{Q}_2'\left(\mathbf{Q_2}\boldsymbol{\Upsilon}_{f2}\mathbf{Q}_2'\right)^{+}\mathbf{Q_2}\boldsymbol{\kappa}_{i2f} \tag{S21}$$
$$= \boldsymbol{\vartheta}_{i2f}'\boldsymbol{\Upsilon}_{f2}^{-1}\boldsymbol{\kappa}_{i2f} = \boldsymbol{\omega}_{i2\mathbf{v}}'\boldsymbol{\Lambda}_{f2}^{*\prime}\left(\boldsymbol{\Lambda}_{f2}^{*}\mathbf{G}_{\mathbf{v}2}\boldsymbol{\Lambda}_{f2}^{*\prime}\right)^{-1}\boldsymbol{\Lambda}_{f2}^{*}\boldsymbol{\pi}_{i2\mathbf{v}} = \boldsymbol{\omega}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v2}}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}.$$

Therefore, together with (S13), (S14) and (S21), as $(T, N) \xrightarrow{j} \infty$ with $\sqrt{T}/N \to 0$ we have

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\hat{\mathbf{s}}_{i,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i(r)dW_i(r) - \boldsymbol{\omega}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v2}}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}. \tag{S22}$$

In a similar manner, noting that as $(T, N) \xrightarrow{j} \infty$, with $\sqrt{T}/N \to 0$

$$\frac{\hat{\mathbf{s}}_{i,-1}'\hat{\mathbf{s}}_{i,-1}}{T^2} = \frac{\mathbf{s}_{iy,-1}'\mathbf{s}_{iy,-1}}{\sigma_i^2 T^2} + O_p\left(\frac{1}{\sqrt{N}}\right) \tag{S23}$$
$$\overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i^2(r)dr,$$

and so we have that

$$\frac{\hat{\mathbf{s}}_{i,-1}'\bar{\mathbf{M}}\hat{\mathbf{s}}_{i,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i^2(r)dr - \boldsymbol{\pi}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v2}}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}. \tag{S24}$$

Next, consider $\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i/(T-2k-5)$. Note that $\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i$ are the residuals from the regression of $\boldsymbol{\upsilon}_i$ on $\bar{\mathbf{W}}_i = (\bar{\mathbf{W}}, \mathbf{y}_{i,-1})$, but from equation (S8) $\mathbf{y}_{i,-1}$ has components $(\bar{\mathbf{Z}}_{-1}, \boldsymbol{\tau}_T, \mathbf{t}_{T-1}, \hat{\mathbf{s}}_{i,-1})$. As $(\bar{\mathbf{Z}}_{-1}, \boldsymbol{\tau}_T, \mathbf{t}_{T-1}) \subset \bar{\mathbf{W}}$, but $\hat{\mathbf{s}}_{i,-1}$ is not contained in $\bar{\mathbf{W}}$, by regression theory $\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i = \bar{\mathbf{M}}_i^*\boldsymbol{\upsilon}_i$, where $\bar{\mathbf{M}}_i^* = \mathbf{I} - \bar{\mathbf{H}}_i\left(\bar{\mathbf{H}}_i'\bar{\mathbf{H}}_i\right)^{-1}\bar{\mathbf{H}}_i'$, with $\bar{\mathbf{H}}_i = (\bar{\mathbf{W}},\hat{\mathbf{s}}_{i,-1})$. Thus,

$$\frac{\boldsymbol{\upsilon}_i\bar{\mathbf{M}}_i^*\boldsymbol{\upsilon}_i}{T-2k-5} = \frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T-2k-5} - \frac{\left(\boldsymbol{\upsilon}_i'\bar{\mathbf{H}}_i\mathbf{B}_*\right)\left(\mathbf{B}_*\bar{\mathbf{H}}_i'\bar{\mathbf{H}}_i\mathbf{B}_*\right)^{-1}\left(\mathbf{B}_*\bar{\mathbf{H}}_i'\boldsymbol{\upsilon}_i\right)}{T-2k-5}, \tag{S25}$$

where

$$\underset{(2k+5)\times(2k+5)}{\mathbf{B}_*} = \begin{pmatrix} \mathbf{B}_2 & \mathbf{0} \\ \mathbf{0} & T^{-1} \end{pmatrix}.$$

First note that using Lemma A.1 we have

$$\frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T-2k-5} \overset{(N,T)_j}{\to} 1. \tag{S26}$$

We also have that

$$\mathbf{B}_*\bar{\mathbf{H}}_i'\boldsymbol{\upsilon}_i = \begin{pmatrix} \mathbf{B}_2\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i \\ \hat{\mathbf{s}}_{i,-1}'\boldsymbol{\upsilon}_i/T \end{pmatrix}, \mathbf{B}_*\bar{\mathbf{H}}_i'\bar{\mathbf{H}}_i'\mathbf{B}_* = \begin{pmatrix} \mathbf{B}_2\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{B}_2 & \mathbf{B}_2\bar{\mathbf{W}}'\hat{\mathbf{s}}_{i,-1}/T \\ \hat{\mathbf{s}}_{i,-1}'\bar{\mathbf{W}}\mathbf{B}_2/T & \hat{\mathbf{s}}_{i,-1}'\hat{\mathbf{s}}_{i,-1}/T^2 \end{pmatrix},$$

so then using (S14), (S23), and following the same line of analysis as for the results in (S21), it can be seen that $\left(\boldsymbol{\upsilon}_i'\bar{\mathbf{H}}_i\mathbf{B}_*\right)\left(\mathbf{B}_*\bar{\mathbf{H}}_i'\bar{\mathbf{H}}_i\mathbf{B}_*\right)^{-1}\left(\mathbf{B}_*\bar{\mathbf{H}}_i'\boldsymbol{\upsilon}_i\right)$ in (S25) will tend to a function of standard Brownian motions as $(T, N) \xrightarrow{j} \infty$ with $\sqrt{T}/N \to 0$. Thus, dividing by $T-2k-5$ makes the second term of (S25) asymptotically negligible, and together with the results in (S25) and (S26) we have $\frac{\boldsymbol{\upsilon}_i\bar{\mathbf{M}}_i^*\boldsymbol{\upsilon}_i}{T-2k-5} \overset{(N,T)_j}{\to} 1$ therefore, as $(T, N) \xrightarrow{j} \infty$ with $\sqrt{T}/N \to 0$,

$$\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i/(T-2k-5) \overset{(N,T)_j}{\to} 1. \tag{S27}$$

Finally, from the results in (S10), (S22), (S24) and (S27), we have, as $\sqrt{T}/N \to 0$,

$$t_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \frac{\int_0^1 W_i(r)dW_i(r) - \boldsymbol{\omega}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v2}}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}}{\left(\int_0^1 W_i^2(r)dr - \boldsymbol{\pi}_{i2\mathbf{v}}'\mathbf{G}_{\mathbf{v2}}^{-1}\boldsymbol{\pi}_{i2\mathbf{v}}\right)^{1/2}}, \tag{S28}$$

as required. Condition $\sqrt{T}/N \to 0$ is satisfied so long as $T/N \to \delta$, where δ is a fixed finite non-zero positive constant. For sequential asymptotics, with $N \to \infty$, first, we note that for a fixed T and as $N \to \infty$, $\mathbf{Q} = \text{plim}_{N\to\infty}\mathbf{Q}_N$ and by Lemma A.1, (S16) continues to hold (replacing '$\overset{(N,T)_j}{\to}$' by '$\overset{N}{\to}$'). Then, letting $T \to \infty$ yields (S28). ■

S2 Proof of Theorem 2.2 in PSY: The Case of Serially Correlated Errors

The t-ratio for this case is given by (42) in PSY which can be written as

$$t_i(N,T) = \frac{\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\hat{\mathbf{s}}_{i\zeta,-1}}{T}}{\left(\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1,p}\boldsymbol{\upsilon}_i}{T-3k-6}\right)^{1/2}\left(\frac{\hat{\mathbf{s}}_{i\zeta,-1}'\bar{\mathbf{M}}_{i1}\hat{\mathbf{s}}_{i\zeta,-1}}{T^2}\right)^{1/2}}, \tag{S29}$$

where $\boldsymbol{\upsilon}_i = [\boldsymbol{\eta}_{iy} - (\bar{\mathbf{E}} - \theta\bar{\mathbf{E}}_{-1})\boldsymbol{\delta}_i]/\sigma_{i\eta}$, $\hat{\mathbf{s}}_{i\zeta,-1} = (\mathbf{s}_{i\zeta,-1} - \bar{\mathbf{S}}_{-1}\boldsymbol{\delta}_i)/\sigma_{i\eta}$, and $\bar{\mathbf{M}}_{i1} = \mathbf{I}_T - \bar{\mathbf{W}}_{i1}\left(\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}\right)^{-1}\bar{\mathbf{W}}_{i1}'$ with $\bar{\mathbf{W}}_{i1} = (\Delta\mathbf{y}_{i,-1}, \Delta\bar{\mathbf{Z}}, \Delta\bar{\mathbf{Z}}_{-1}, \boldsymbol{\tau}_T, \bar{\mathbf{Z}}_{-1})$. Define the matrices $\mathbf{W}_{i1f} = (\mathbf{F}_{-1}\gamma_{iy} + \boldsymbol{\zeta}_{iy,-1}, \mathbf{F}, \mathbf{F}_{-1}, \boldsymbol{\tau}_T, \mathbf{S}_{f,-1})$ and $\bar{\boldsymbol{\Xi}}_1 = \left(\mathbf{0}_T, \bar{\mathbf{E}}, \bar{\mathbf{E}}_{-1}, \mathbf{0}_T, \bar{\mathbf{S}}_{-1}\right)$, so that

$$\bar{\mathbf{W}}_{i1}' = \mathbf{Q}_{1N}\mathbf{W}_{i1f}' + \bar{\boldsymbol{\Xi}}_1', \text{ with } \mathbf{Q}_{1N} = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\boldsymbol{\Gamma}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_N \end{pmatrix}, \tag{S30}$$

where $\mathbf{Q}_N$ is defined by (A.2) in PSY. Expanding $\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\hat{\mathbf{s}}_{i\zeta,-1}/T$ gives

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\hat{\mathbf{s}}_{i\zeta,-1}}{T} = \frac{\boldsymbol{\upsilon}_i'\hat{\mathbf{s}}_{i\zeta,-1}}{T} - \left(\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_{i1}\mathbf{B}_1\right)\left(\mathbf{B}_1\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}\mathbf{B}_1\right)^{-1}\left(\frac{\mathbf{B}_1\bar{\mathbf{W}}_{i1}'\hat{\mathbf{s}}_{i\zeta,-1}}{T}\right), \tag{S31}$$

where

$$\underset{(3k+5)\times(3k+5)}{\mathbf{B}_1} = \begin{pmatrix} \frac{1}{\sqrt{T}}\mathbf{I}_{2k+4} & \mathbf{0} \\ \mathbf{0} & \frac{1}{T}\mathbf{I}_{k+1} \end{pmatrix}.$$

Using Lemma A.1 together with the results in Proposition 17.1 of Hamilton (1994; p.486) we have

$$\frac{\boldsymbol{\upsilon}_i'\hat{\mathbf{s}}_{i\zeta,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \frac{1}{1-\theta}\int_0^1 W_i(r)dW_i(r). \tag{S32}$$

From (S30) it follows that

$$\begin{aligned} \mathbf{B}_1\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i &= \mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{if}'\boldsymbol{\upsilon}_i + \mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\boldsymbol{\upsilon}_i, \\ \frac{\mathbf{B}_1\bar{\mathbf{W}}_{i1}'\hat{\mathbf{s}}_{\zeta i,-1}}{T} &= \frac{\mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{if}'\hat{\mathbf{s}}_{\zeta i,-1}}{T} + \frac{\mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\hat{\mathbf{s}}_{\zeta i,-1}}{T}, \end{aligned} \tag{S33}$$

$$\begin{aligned} \mathbf{B}_1\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}\mathbf{B}_1 = {} & \mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{if}'\mathbf{W}_{if}\mathbf{Q}_{1N}'\mathbf{B}_1 \\ & + \mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{if}'\bar{\boldsymbol{\Xi}}_1\mathbf{B}_1 + \mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\mathbf{W}_{if}\mathbf{Q}_{1N}'\mathbf{B}_1 + \mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\bar{\boldsymbol{\Xi}}_1\mathbf{B}_1. \end{aligned}$$

Using Lemma A.1, as $(T,N) \overset{j}{\to} \infty$ with $\sqrt{T}/N \to 0$,

$$\mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\boldsymbol{\upsilon}_i \overset{(N,T)_j}{\to} \mathbf{0},\ \frac{\mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\hat{\mathbf{s}}_{\zeta i,-1}}{T} \overset{(N,T)_j}{\to} \mathbf{0},\ \mathbf{B}_1\bar{\boldsymbol{\Xi}}_1'\bar{\boldsymbol{\Xi}}_1\mathbf{B}_1 \overset{(N,T)_j}{\to} \mathbf{0}, \text{ and } \mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{if}'\bar{\boldsymbol{\Xi}}_1\mathbf{B}_1 \overset{(N,T)_j}{\to} \mathbf{0}. \tag{S34}$$

Define

$$\underset{(3m^0+2)\times(3m^0+2)}{\mathbf{C}_1} = \begin{pmatrix} \frac{1}{\sqrt{T}}\mathbf{I}_{2m^0+2} & \mathbf{0} \\ \mathbf{0} & \frac{1}{T}\mathbf{I}_{m^0} \end{pmatrix},$$

so that, using Lemma A.1 in PSY and the results in Proposition 17.1 and 18.1 of Hamilton (1994; p.486, p.547-8), as $(T,N) \overset{j}{\to} \infty$ with $\sqrt{T}/N \to \infty$ we have

$$\mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{i1f}'\boldsymbol{\upsilon}_i = \mathbf{Q}_{1N}\mathbf{C}_1\mathbf{W}_{i1f}'\boldsymbol{\upsilon}_i \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}_1\boldsymbol{\vartheta}_{i1f}, \tag{S35}$$

$$\frac{\mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{i1f}'\hat{\mathbf{s}}_{\zeta i,-1}}{T} = \frac{\mathbf{Q}_{1N}\mathbf{C}_1\mathbf{W}_{i1f}'\hat{\mathbf{s}}_{\zeta i,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}_1\boldsymbol{\kappa}_{i1f}, \tag{S36}$$

$$\mathbf{B}_1\mathbf{Q}_{1N}\mathbf{W}_{i1f}'\mathbf{W}_{i1f}\mathbf{Q}_{1N}'\mathbf{B}_1 = \mathbf{Q}_{1N}\mathbf{C}_1\mathbf{W}_{i1f}'\mathbf{W}_{i1f}\mathbf{C}_1\mathbf{Q}_{1N}' \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}_1\boldsymbol{\Upsilon}_{i1f}\mathbf{Q}_1', \tag{S37}$$

where

$$\mathbf{Q}_1 = \underset{N\to\infty}{\text{plim}}\ \mathbf{Q}_{1N},\ \boldsymbol{\vartheta}_{if1} = \begin{pmatrix} \gamma_{iy}'\boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) + \sqrt{\frac{\sigma_{i\eta}^2}{1-\theta^2}}W_i(1) \\ \boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\ \boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\ \boldsymbol{\Lambda}_f^*\boldsymbol{\omega}_{i\mathbf{v}} \end{pmatrix},\ \boldsymbol{\kappa}_{i1f} = \begin{pmatrix} \mathbf{0}_{2m^0+1} \\ \frac{1}{1-\theta}\boldsymbol{\Lambda}_f^*\boldsymbol{\pi}_{i\mathbf{v}} \end{pmatrix},$$

$$\boldsymbol{\Upsilon}_{i1f}=\left(\begin{array}{cc}\varkappa_{if1} & \mathbf{0}_{2m^0+1\times m^0+1}\\ \mathbf{0}'_{2m^0+1\times m^0+1} & \boldsymbol{\Lambda}^*_f\mathbf{G}_{\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\end{array}\right),\ \varkappa_{i1f}=\left(\begin{array}{ccc}\gamma'_{iy}\gamma_{iy}+\frac{\sigma^2_{\eta i}}{1-\theta^2} & \gamma'_{iy}\boldsymbol{\Sigma}'_{f1} & \gamma'_{iy}\\ \boldsymbol{\Sigma}_{f1}\gamma_{iy} & \mathbf{I}_{m^0} & \boldsymbol{\Sigma}_{f1}\\ \gamma_{iy} & \boldsymbol{\Sigma}'_{f1} & \mathbf{I}_{m^0}\end{array}\right),$$

$\boldsymbol{\Lambda}_f$ and $\boldsymbol{\Lambda}^*_f$ are defined by (3) and (A.12), respectively, $\mathbf{W}_{\mathbf{v},i}(1)$ is defined such that $T^{-1/2}\sum_{t=1}^{T}\mathbf{v}_t\eta_{iyt}/\sigma_{i\eta}\overset{T}{\Longrightarrow}\mathbf{W}_{\mathbf{v},i}(1)$, with $\mathbf{v}_t$ defined as in Assumption 2, $W_i(r)$ is a standard Brownian motion and $\mathbf{W}_{\mathbf{v}}(r)$ is an m^0-dimensional standard Brownian motion defined on [0,1], $\boldsymbol{\omega}_{i\mathbf{v}}$, $\boldsymbol{\pi}_{i\mathbf{v}}$ and $\mathbf{G}_{\mathbf{v}}$ are defined by (A.12), and $\boldsymbol{\Sigma}_{f\ell}=E(\mathbf{f}_t\mathbf{f}'_{t-\ell})$. Collecting the results from (S33) to (S37), as well as using Lemma A.2, (S32) and (S31) we have

$$\frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}_{i1}\mathring{\mathbf{s}}_{i\zeta,-1}}{T}=\frac{1}{1-\theta}\int_0^1 W_i(r)dW_i(r)-\boldsymbol{\omega}'_{i\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\left(\boldsymbol{\Lambda}^*_f\mathbf{G}_{\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\right)^{-1}\frac{1}{1-\theta}\boldsymbol{\Lambda}^*_f\boldsymbol{\pi}_{i\mathbf{v}}. \tag{S38}$$

In a similar manner, noting that as $(T,N)\overset{j}{\to}\infty$ with $\sqrt{T}/N\to 0$

$$\frac{\mathbf{s}'_{i\zeta,-1}\mathbf{s}_{i\zeta,-1}}{\sigma^2_{i\eta}T^2}\overset{(N,T)_j}{\Longrightarrow}\frac{1}{(1-\theta)^2}\int_0^1 W_i^2(r)dr, \tag{S39}$$

we have that

$$\frac{\mathring{\mathbf{s}}'_{i\zeta,-1}\bar{\mathbf{M}}_{i1}\mathring{\mathbf{s}}_{i\zeta,-1}}{T^2}\overset{(N,T)_j}{\Longrightarrow}\frac{1}{(1-\theta)^2}\int_0^1 W_i^2(r)dr-\frac{1}{1-\theta}\boldsymbol{\pi}'_{i\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\left(\boldsymbol{\Lambda}^*_f\mathbf{G}_{\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\right)^{-1}\frac{1}{1-\theta}\boldsymbol{\Lambda}^*_f\boldsymbol{\pi}_{i\mathbf{v}}. \tag{S40}$$

For the term $\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i/(T-3k-6)$, following a similar reasoning as in the uncorrelated case we can write $\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i=\bar{\mathbf{M}}^*_{i1}\boldsymbol{\upsilon}_i$, where $\bar{\mathbf{M}}^*_{i1}=\mathbf{I}_T-\bar{\mathbf{H}}_{i1}\left(\bar{\mathbf{H}}'_{i1}\bar{\mathbf{H}}_{i1}\right)^{-1}\bar{\mathbf{H}}'_{i1}$ with $\bar{\mathbf{H}}_{i1}=\left(\bar{\mathbf{W}}_{i1},\mathring{\mathbf{s}}_{i\zeta,-1}\right)$. Thus

$$\frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}^*_{i1}\boldsymbol{\upsilon}_i}{T-3k-6}=\frac{\boldsymbol{\upsilon}'_i\boldsymbol{\upsilon}_i}{T-3k-6}-\frac{\left(\boldsymbol{\upsilon}'_i\bar{\mathbf{H}}_{i1}\mathbf{B}'_{1*}\right)\left(\mathbf{B}_{1*}\bar{\mathbf{H}}'_{i1}\bar{\mathbf{H}}_{i1}\mathbf{B}'_{1*}\right)^{-1}\left(\mathbf{B}_{1*}\bar{\mathbf{H}}'_{i1}\boldsymbol{\upsilon}_i\right)}{T-3k-6}, \tag{S41}$$

where

$$\underset{(3k+6)\times(3k+6)}{\mathbf{B}_{1*}}=\left(\begin{array}{cc}\mathbf{B}_1 & \mathbf{0}\\ \mathbf{0} & T^{-1}\end{array}\right),$$

Using Lemma A.1 in PSY, first note that

$$\boldsymbol{\upsilon}'_i\boldsymbol{\upsilon}_i/(T-3k-6)\overset{(N,T)_j}{\Longrightarrow}1. \tag{S42}$$

Also, since

$$\mathbf{B}_{1*}\bar{\mathbf{H}}'_{i1}\boldsymbol{\upsilon}_i=\left(\begin{array}{c}\mathbf{B}_1\bar{\mathbf{W}}'_{i1}\boldsymbol{\upsilon}_i\\ \mathring{\mathbf{s}}'_{i\zeta,-1}\boldsymbol{\upsilon}_i/T\end{array}\right),\ \mathbf{B}_{1*}\bar{\mathbf{H}}'_{i1}\bar{\mathbf{H}}_{i1}\mathbf{B}_{1*}=\left(\begin{array}{cc}\mathbf{B}_1\bar{\mathbf{W}}'_{i1}\bar{\mathbf{W}}_{i1}\mathbf{B}_1 & \mathbf{B}_1\bar{\mathbf{W}}'_{i1}\mathring{\mathbf{s}}_{i\zeta,-1}/T\\ \mathring{\mathbf{s}}'_{i\zeta,-1}\bar{\mathbf{W}}_{i1}\mathbf{B}_1/T & \mathring{\mathbf{s}}'_{i\zeta,-1}\mathring{\mathbf{s}}_{i\zeta,-1}/T^2\end{array}\right),$$

using (S32), (S39), and following a similar reasoning as for the results in (S38), it can be seen that $\left(\boldsymbol{\upsilon}'_i\bar{\mathbf{H}}_{i1}\mathbf{B}'_{1*}\right)\left(\mathbf{B}_{1*}\bar{\mathbf{H}}'_{i1}\bar{\mathbf{H}}_{i1}\mathbf{B}'_{1*}\right)^{-1}\left(\mathbf{B}_{1*}\bar{\mathbf{H}}'_{i1}\boldsymbol{\upsilon}_i\right)$ in (S41) will tend to a function of standard Brownian motions as $(T,N)\overset{j}{\to}\infty$ with $\sqrt{T}/N\to\infty$. Thus, dividing by $T-3k-6$ makes the second term of (S41) asymptotically negligible, and together with the results in (S41) and (S42) we have $\frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}^*_{i1}\boldsymbol{\upsilon}_i}{T}\overset{(N,T)_j}{\to}1$. Therefore, as $(T,N)\overset{j}{\to}\infty$ with $\sqrt{T}/N\to\infty$,

$$\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i/(T-3k-6)\overset{(N,T)_j}{\to}1. \tag{S43}$$

Finally, from the results in (S29), (S38), (S40) and (S43), we have, as $\sqrt{T}/N\to 0$,

$$t_i(N,T)\overset{(N,T)_j}{\Longrightarrow}\frac{\frac{1}{1-\theta}\int_0^1 W_i(r)dW_i(r)-\boldsymbol{\omega}'_{i\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\left(\boldsymbol{\Lambda}^*_f\mathbf{G}_{\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\right)^{-1}\frac{1}{1-\theta}\boldsymbol{\Lambda}^*_f\boldsymbol{\pi}_{i\mathbf{v}}}{\left(\frac{1}{(1-\theta)^2}\int_0^1 W_i^2(r)dr-\frac{1}{1-\theta}\boldsymbol{\pi}'_{i\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\left(\boldsymbol{\Lambda}^*_f\mathbf{G}_{\mathbf{v}}\boldsymbol{\Lambda}^{*\prime}_f\right)^{-1}\frac{1}{1-\theta}\boldsymbol{\Lambda}^*_f\boldsymbol{\pi}_{i\mathbf{v}}\right)^{1/2}} \tag{S44}$$

$$=\frac{\int_0^1 W_i(r)dW_i(r)-\boldsymbol{\omega}'_{i\mathbf{v}}\mathbf{G}_{\mathbf{v}}^{-1}\boldsymbol{\pi}_{i\mathbf{v}}}{\left(\int_0^1 W_i^2(r)dr-\boldsymbol{\pi}'_{i\mathbf{v}}\mathbf{G}_{\mathbf{v}}^{-1}\boldsymbol{\pi}_{i\mathbf{v}}\right)^{1/2}}$$

as required, which is identical to the limit distribution obtained for $\theta=0$. Condition $\sqrt{T}/N\to 0$ is satisfied so long as $T/N\to\delta$, where δ is a fixed finite non-zero positive constant. For sequential asymptotics, with $N\to\infty$ first, we note that for a fixed T and as $N\to\infty$, $\mathbf{Q}=\text{plim}_{N\to\infty}\mathbf{Q}_N$ and by Lemma A.1 in PSY, (S34) continues to hold (replacing '$\overset{(N,T)_j}{\to}$' by '$\overset{N}{\to}$'). Then, letting $T\to\infty$ yields (S44).

S3 The Limiting Distribution of the CSB_i Statistics

S3.1 The Case of Serially Uncorrelated Errors

Consider

$$\Delta y_{it} = \beta_i(y_{i,t-1} - \boldsymbol{\alpha}'_{iy}\mathbf{d}_{t-1}) + \boldsymbol{\alpha}'_{iy}\Delta\mathbf{d}_t + \gamma'_{iy}\mathbf{f}_t + \varepsilon_{iyt}, \tag{S45}$$

where $\mathbf{d}_t = (1,t)'$ and recall the expression for $\mathbf{z}_{it}$,

$$\mathbf{z}_{it} = \mathbf{z}_{i0} + \boldsymbol{\Gamma}_i\mathbf{s}_{ft} + \mathbf{A}_i\mathbf{d}_t + \mathbf{s}_{it}. \tag{S46}$$

In matrix notation, under the null hypothesis

$$H_0 : \beta_i = 0 \text{ for all } i, \tag{S47}$$

we have

$$\Delta\mathbf{y}_i = \alpha_{iy1}\boldsymbol{\tau}_T + \mathbf{F}\gamma_{iy} + \boldsymbol{\varepsilon}_{iy}, \tag{S48}$$

where $\Delta\mathbf{y}_i = (\Delta y_{i1}, \Delta y_{i2}, ..., \Delta y_{iT})'$, $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_T)'$, $\boldsymbol{\varepsilon}_{iy} = (\varepsilon_{iy1}, \varepsilon_{iy2}, ..., \varepsilon_{iyT})'$, and

$$\Delta\bar{\mathbf{Z}} = \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}'_1 + \mathbf{F}\bar{\boldsymbol{\Gamma}}' + \bar{\mathbf{E}}, \tag{S49}$$

where $\Delta\bar{\mathbf{Z}} = (\Delta\bar{\mathbf{z}}_1, \Delta\bar{\mathbf{z}}_2, ..., \Delta\bar{\mathbf{z}}_T)'$ with $\Delta\bar{\mathbf{z}}_t = N^{-1}\sum_{i=1}^N \Delta\mathbf{z}_{it}$, $\Delta\mathbf{z}_{it} = (\Delta y_{it}, \Delta\mathbf{x}'_{it})'$and $\bar{\mathbf{E}} = N^{-1}\sum_{i=1}^N \mathbf{E}_i$, $\mathbf{E}_i = (\boldsymbol{\varepsilon}_{i1}, \boldsymbol{\varepsilon}_{i2}, ..., \boldsymbol{\varepsilon}_{iT})'$ with $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{iyt}, \boldsymbol{\varepsilon}'_{ixt})'$. Substituting $\mathbf{F} = \left(\Delta\bar{\mathbf{Z}} - \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}'_1 - \bar{\mathbf{E}}\right)\bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{\Gamma}}'\bar{\boldsymbol{\Gamma}}\right)^{-1}$, which is obtained by (S49), in (S48) yields

$$\Delta\mathbf{y}_i = \tilde{\alpha}_{i1}\boldsymbol{\tau}_T + \Delta\bar{\mathbf{Z}}\boldsymbol{\delta}_i + \sigma_i\boldsymbol{\upsilon}_i,$$

where $\tilde{\alpha}_{i1} = \alpha_{iy1} - \bar{\boldsymbol{\alpha}}'_1\boldsymbol{\delta}_i$, $\boldsymbol{\delta}_i = \bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{\Gamma}}'\bar{\boldsymbol{\Gamma}}\right)^{-1}\gamma_{iy}$, $\boldsymbol{\upsilon}_i = (\boldsymbol{\varepsilon}_{iy} - \bar{\mathbf{E}}\boldsymbol{\delta}_i)/\sigma_i$.

The test of the panel unit root hypothesis using the Sargan-Bhargava statistic is based on the cross section augmented regression

$$\Delta y_{it} = g_{i0} + \mathbf{c}'_i\Delta\bar{\mathbf{z}}_t + \epsilon_{it},$$

where the cross section augmented Sargan-Bhargava statistic is given by

$$CSB_i(N,T) = T^{-2}\frac{\sum_{t=1}^T \hat{u}_{it}^2}{\hat{\sigma}_i^2}, \tag{S50}$$

with $\hat{u}_{it} = \sum_{s=1}^t \hat{\epsilon}_{is}$, and $\hat{\sigma}_i^2 = \sum_{t=1}^T \hat{\epsilon}_{it}^2/(T-k-2)$.

Theorem S3.1 *Suppose the series $\mathbf{z}_{it}$, for $i = 1,2,...,N$, $t = 1,2,...,T$, is generated under (S47) according to (S46) and $\mathbf{d}_t = (1,\mathbf{t}_T)'$. Then under Assumptions 1-5 and as N and $T \to \infty$, such that $\sqrt{T}/N \to 0$, the joint $[(N,T)_j \to \infty]$ limit distribution of $CSB_i(N,T)$ given by (S50), is free of nuisance parameters and is given by*

$$CSB_i = \int_0^1 W_i^2(r)dr + \frac{1}{3}[W_i(1)]^2 - 2W_i(1)\int_0^1 rW_i(r)dr \tag{S51}$$

where $W_i(r)$ is a scalar standard Brownian motion defined on [0,1], associated with ε_{iyt}.

Proof. In matrix notation

$$\begin{aligned}\hat{\mathbf{u}}_i &= (\hat{u}_{i1}, \hat{u}_{i2}, ..., \hat{u}_{iT})',\\ \hat{\boldsymbol{\epsilon}}_i &= (\hat{\epsilon}_{i1}, \hat{\epsilon}_{i2}, ..., \hat{\epsilon}_{iT})',\end{aligned}$$

$$\hat{\mathbf{u}}_i = \mathbf{H}\hat{\boldsymbol{\epsilon}}_i,$$

where

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix},$$

$$\hat{\sigma}_i^2 = \frac{\Delta\mathbf{y}'_i\bar{\mathbf{M}}\Delta\mathbf{y}_i}{T-k-2},$$

with $\bar{\mathbf{M}} = \mathbf{I}_T - \bar{\mathbf{W}}\left(\bar{\mathbf{W}}'\bar{\mathbf{W}}\right)^{+}\bar{\mathbf{W}}'$, $\bar{\mathbf{W}} = \left(\Delta\bar{\mathbf{Z}}, \boldsymbol{\tau}_T\right)$. It follows that

$$CSB_i(N,T) = T^{-2}\frac{\hat{\mathbf{u}}_i'\hat{\mathbf{u}}_i}{\hat{\sigma}_i^2} = T^{-2}\frac{\hat{\boldsymbol{\epsilon}}_i'\mathbf{H}'\mathbf{H}\hat{\boldsymbol{\epsilon}}_i}{\left(\frac{\Delta\mathbf{y}_i'\bar{\mathbf{M}}\Delta\mathbf{y}_i}{T-k-2}\right)}.$$

We also have that

$$\hat{\boldsymbol{\epsilon}}_i = \bar{\mathbf{M}}\Delta\mathbf{y}_i = \sigma_i\bar{\mathbf{M}}\boldsymbol{\upsilon}_i,$$

so then

$$CSB_i(N,T) = \frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\mathbf{H}'\mathbf{H}\bar{\mathbf{M}}\boldsymbol{\upsilon}_i/T^2}{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\boldsymbol{\upsilon}_i/\left(T-k-2\right)}. \tag{S52}$$

Consider first the denominator of (S52)

$$\frac{\boldsymbol{\upsilon}_i\bar{\mathbf{M}}\boldsymbol{\upsilon}_i}{T-k-2} = \frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T-k-2} - \frac{1}{T-k-2}\left(\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}}{\sqrt{T}}\right)\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\left(\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\right). \tag{S53}$$

Noting that $\boldsymbol{\upsilon}_i = (\boldsymbol{\varepsilon}_{iy} - \bar{\mathbf{E}}\boldsymbol{\delta}_i)/\sigma_i$ and using Lemma A.1 of PSY we have that

$$\frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T-k-2} = \frac{\boldsymbol{\varepsilon}_{iy}'\boldsymbol{\varepsilon}_{iy}}{\sigma_i^2(T-k-2)} + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right) \overset{(N,T)_j}{\to} 1. \tag{S54}$$

Let $\mathbf{W}_f = (\mathbf{F}, \boldsymbol{\tau}_T)$ and $\bar{\boldsymbol{\Xi}} = \left(\bar{\mathbf{E}}, \mathbf{0}_T\right)$ so that

$$\bar{\mathbf{W}}' = \mathbf{Q}_N\mathbf{W}_f' + \bar{\boldsymbol{\Xi}}', \text{ where } \underset{(k+2)\times(m^0+1)}{\mathbf{Q}_N} = \begin{pmatrix} \bar{\boldsymbol{\Gamma}} & \bar{\boldsymbol{\alpha}}_1 \\ \mathbf{0} & 1 \end{pmatrix}. \tag{S55}$$

Using (S55), by Lemma A.1 and noting that $\mathbf{Q}_N = O_p(1)$ we have

$$\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}} = \mathbf{Q}_N\frac{\mathbf{W}_f'\boldsymbol{\varepsilon}_{iy}}{\sigma_i\sqrt{T}} + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{\sqrt{T}}{N}\right)$$

$$\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T} = \mathbf{Q}_N\frac{\mathbf{W}_f'\mathbf{W}_f}{T}\mathbf{Q}_N' + O_p\left(\frac{1}{\sqrt{NT}}\right) + O_p\left(\frac{1}{N}\right).$$

Thus, as $(T,N)\overset{j}{\to}\infty$ with $\sqrt{T}/N \to 0$ we have

$$\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}\boldsymbol{\vartheta}_{if}, \tag{S56}$$

$$\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}\mathbf{Q}', \tag{S57}$$

where

$$\mathbf{Q} = \underset{N\to\infty}{\text{plim}}\,\mathbf{Q}_N,\ \boldsymbol{\vartheta}_{if} = \begin{pmatrix} \boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\ W_i(1) \end{pmatrix},$$

since

$$\frac{\mathbf{W}_f'\boldsymbol{\varepsilon}_{iy}}{\sigma_i\sqrt{T}} = \begin{pmatrix} \frac{\mathbf{F}'\boldsymbol{\varepsilon}_{iy}}{\sigma_i\sqrt{T}} \\ \frac{\boldsymbol{\tau}_T'\boldsymbol{\varepsilon}_{iy}}{\sigma_i\sqrt{T}} \end{pmatrix} \overset{T}{\Longrightarrow} \begin{pmatrix} \boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\ W_i(1) \end{pmatrix},\ \frac{\mathbf{W}_f'\mathbf{W}_f}{T} = \begin{pmatrix} \frac{\mathbf{F}'\mathbf{F}}{T} & \frac{\mathbf{F}'\boldsymbol{\tau}_T}{T} \\ \frac{\boldsymbol{\tau}_T'\mathbf{F}}{T} & \frac{\boldsymbol{\tau}_T'\boldsymbol{\tau}_T}{T} \end{pmatrix} \overset{T}{\to} \begin{pmatrix} \mathbf{I}_{m^0} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \tag{S58}$$

where $\boldsymbol{\Lambda}_f$ is defined by (3), $\mathbf{W}_{\mathbf{v},i}(1)$ is defined such that $T^{-1/2}\sum_{t=1}^{T}\mathbf{v}_t\varepsilon_{iyt}/\sigma_i \overset{T}{\Longrightarrow} \mathbf{W}_{\mathbf{v},i}(1)$, with $\mathbf{v}_t$ defined as in Assumption 2, $\mathbf{W}_{\mathbf{v}}(r)$ is an m^0-dimensional standard Brownian motion associated with $\mathbf{v}_t$ defined on [0,1], and $W_i(r)$ is defined as above. These two groups of Brownian motions $(\mathbf{W}_{\mathbf{v}}(r), W_i(r))$ are independent of each other. Collecting the above results, as well as using Lemma A.2 in PSY we have

$$\left(\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}}{\sqrt{T}}\right)\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\left(\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\right) \overset{(N,T)_j}{\Longrightarrow} \boldsymbol{\vartheta}_{if}'\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^{+}\mathbf{Q}\boldsymbol{\vartheta}_{if} = \boldsymbol{\vartheta}_{if}'\boldsymbol{\vartheta}_{if}.$$

Dividing by $T-k-2$ will make the second term of (S53) asymptotically negligible and so it follows that

$$\frac{\boldsymbol{\upsilon}_i\bar{\mathbf{M}}\boldsymbol{\upsilon}_i}{T-k-2} \overset{(N,T)_j}{\to} 1. \tag{S59}$$

Consider next the numerator of (S52). Noting that

$$\begin{aligned}\frac{\mathbf{H}\bar{\mathbf{M}}\boldsymbol{\upsilon}_i}{T} &= \frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}-\frac{\mathbf{H}\bar{\mathbf{W}}\left(\bar{\mathbf{W}}'\bar{\mathbf{W}}\right)^{+}\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{T}\\ &= \frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}-\frac{\mathbf{H}\bar{\mathbf{W}}}{T^{3/2}}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}},\end{aligned}$$

we have

$$\begin{aligned}\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}\mathbf{H}'\mathbf{H}\bar{\mathbf{M}}\boldsymbol{\upsilon}_i}{T^2} &= \frac{\boldsymbol{\upsilon}_i'\mathbf{H}'\mathbf{H}\boldsymbol{\upsilon}_i}{T^2}-\frac{\boldsymbol{\upsilon}_i'\mathbf{H}'}{T}\frac{\mathbf{H}\bar{\mathbf{W}}}{T^{3/2}}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\\ &\quad -\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\frac{\bar{\mathbf{W}}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}\\ &\quad +\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\frac{\bar{\mathbf{W}}'\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}}{T^3}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^{+}\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\\ &= I-2II+III. \end{aligned} \tag{S60}$$

We look at terms I, II and III in turn. Consider I. Noting that we can write

$$\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}=\frac{\mathbf{s}_{iy}-\bar{\mathbf{S}}\boldsymbol{\delta}_i}{T\sigma_i}=\frac{\mathring{\mathsf{s}}_i}{T}, \tag{S61}$$

where $\mathbf{s}_{iy}=(s_{iy1},...,s_{iy,T})'$ with $s_{iyt}=\sum_{s=1}^{t}\varepsilon_{iys}$ and $\bar{\mathbf{S}}=N^{-1}\sum_{i=1}^{N}\mathbf{S}_i$ with $\mathbf{S}_i=(\mathbf{s}_{i1},...,\mathbf{s}_{i,T})'$, using Lemma A.1 we have

$$\begin{aligned} I &= \frac{\boldsymbol{\upsilon}_i'\mathbf{H}'\mathbf{H}\boldsymbol{\upsilon}_i}{T^2}=\frac{(\mathbf{s}_{iy}'-\boldsymbol{\delta}_i'\bar{\mathbf{S}}')(\mathbf{s}_{iy}-\bar{\mathbf{S}}\boldsymbol{\delta}_i)}{\sigma_i^2T^2}\\ &= \frac{\mathbf{s}_{iy}'\mathbf{s}_{iy}}{\sigma_i^2T^2}+O_p(\frac{1}{\sqrt{N}})+O_p(\frac{1}{N}), \end{aligned}\tag{S62}$$

and as $(T,N)\xrightarrow{j}\infty$ it immediately follows that

$$I \overset{(N,T)_j}{\Longrightarrow}\int_0^1 W_i^2(r)dr,$$

since $\frac{\mathbf{s}_{iy}'\mathbf{s}_{iy}}{\sigma_i^2T^2}\overset{T}{\Longrightarrow}\int_0^1 W_i^2(r)dr$ as $T\to\infty$.

Now consider II. Firstly, using (S55) we can write

$$\mathbf{H}\bar{\mathbf{W}}=\mathbf{W}_{H,f}\mathbf{Q}_N'+\bar{\boldsymbol{\Xi}}_H, \tag{S63}$$

where $\mathbf{H}\bar{\mathbf{W}}=\left(\bar{\mathbf{Z}}-\boldsymbol{\tau}_T\bar{\mathbf{z}}_0',\mathbf{t}_T\right)$, $\mathbf{W}_{H,f}=(\mathbf{S}_f,\mathbf{t}_T)$ with $\mathbf{S}_f=(\mathbf{s}_{f1},...,\mathbf{s}_{fT})'$, $\bar{\boldsymbol{\Xi}}_H=\left(\bar{\mathbf{S}},\mathbf{0}_T\right)$. Together with (S61) we have

$$\frac{\bar{\mathbf{W}}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}=\frac{\mathbf{Q}_N\mathbf{W}_{H,f}'\mathring{\mathsf{s}}_i}{T^{5/2}}+\frac{\bar{\boldsymbol{\Xi}}_H'\mathring{\mathsf{s}}_i}{T^{5/2}}.$$

Using the expression for $\mathring{\mathsf{s}}_i$ given by (S61) and Lemma A.1 together with the result that $\mathbf{t}_T'\bar{\mathbf{S}}/T^{5/2}=O_p(N^{-1/2})$ which follows from a similar derivation of Lemma A.1, we have that

$$\frac{\mathbf{W}_{H,f}'\mathring{\mathsf{s}}_i}{T^{5/2}}=\begin{pmatrix}\frac{\mathbf{S}_f'\mathbf{s}_{iy}}{\sigma_iT^{5/2}}+O_p(\frac{1}{\sqrt{NT}})\\ \frac{\mathbf{t}_T'\mathbf{s}_{iy}}{\sigma_iT^{5/2}}+O_p\left(\frac{1}{\sqrt{N}}\right)\end{pmatrix},\ \frac{\bar{\boldsymbol{\Xi}}_H'\mathring{\mathsf{s}}_i}{T^{5/2}}=O_p(\frac{1}{\sqrt{NT}})+O_p(\frac{1}{\sqrt{N^2T}}). \tag{S64}$$

Noting that

$$\frac{\mathbf{S}_f'\mathbf{s}_{iy}}{\sigma_iT^2}\overset{T}{\Longrightarrow}\boldsymbol{\Lambda}_f\int_0^1\left[\mathbf{W}_{\mathbf{v}}(r)\right]W_i(r)dr\ ,\ \frac{\mathbf{t}_T'\mathbf{s}_{iy}}{\sigma_iT^{5/2}}\overset{T}{\Longrightarrow}\int_0^1 rW_i(r)dr, \tag{S65}$$

as $T\to\infty$, using Lemma A.1 together with the results in (S65) we have, as $(T,N)\xrightarrow{j}\infty$

$$\frac{\bar{\mathbf{W}}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}\overset{(N,T)_j}{\Longrightarrow}\mathbf{Q}\boldsymbol{\kappa}_{if},\text{ with }\boldsymbol{\kappa}_{if}=\begin{pmatrix}\mathbf{0}\\ \int_0^1 rW_i(r)dr\end{pmatrix}. \tag{S66}$$

Now, using (S56), (S57) and (S66) and Lemma A.2 it follows that

$$II = \frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^+ \frac{\bar{\mathbf{W}}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T} \overset{(N,T)_j}{\Longrightarrow} \boldsymbol{\vartheta}_{if}'\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^+\mathbf{Q}\boldsymbol{\kappa}_{if}$$

$$= \boldsymbol{\vartheta}_{if}'\boldsymbol{\kappa}_{if} = W_i(1)\int_0^1 rW_i(r)dr. \tag{S67}$$

Finally consider III. Using (S63) we have

$$\frac{\bar{\mathbf{W}}'\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}}{T^3} = \frac{\mathbf{Q}_N\mathbf{W}_{H,f}'\mathbf{W}_{H,f}\mathbf{Q}_N'}{T^3} + \frac{\mathbf{Q}_N\mathbf{W}_{H,f}'\bar{\boldsymbol{\Xi}}_H}{T^3} + \frac{\bar{\boldsymbol{\Xi}}_H'\mathbf{W}_{H,f}\mathbf{Q}_N'}{T^3} + \frac{\bar{\boldsymbol{\Xi}}_H'\bar{\boldsymbol{\Xi}}_H}{T^3},$$

and by Lemma A.1

$$\frac{\mathbf{W}_{H,f}'\bar{\boldsymbol{\Xi}}_H}{T^3} = \frac{\mathbf{S}_f'\bar{\mathbf{S}}}{T^3} = O_p(\frac{1}{\sqrt{T^2N}}),\ \frac{\bar{\boldsymbol{\Xi}}_H'\bar{\boldsymbol{\Xi}}_H}{T^3} = O_p(\frac{1}{NT}). \tag{S68}$$

Noting that $\mathbf{S}_f'\mathbf{S}_f/T^2 \overset{T}{\Longrightarrow} \boldsymbol{\Lambda}_f\left(\int_0^1[\mathbf{W}_{\mathbf{v}}(r)]\,[\mathbf{W}_{\mathbf{v}}(r)]'\,dr\right)\boldsymbol{\Lambda}_f'$, $\mathbf{S}_f'\mathbf{t}_T/T^{5/2} \overset{T}{\Longrightarrow} \boldsymbol{\Lambda}_f\int_0^1 r\,[\mathbf{W}_{\mathbf{v}}(r)]\,dr$ and $\mathbf{t}_T'\mathbf{t}_T/T^3 \to 1/3$ as $T\to\infty$, we have

$$\frac{\mathbf{W}_{H,f}'\mathbf{W}_{H,f}}{T^3} = \begin{pmatrix} \frac{\mathbf{S}_f'\mathbf{S}_f}{T^3} & \frac{\mathbf{S}_f'\mathbf{t}_T}{T^3} \\ \frac{\mathbf{t}_T'\mathbf{S}_f}{T^3} & \frac{\mathbf{t}_T'\mathbf{t}_T}{T^3}\end{pmatrix} \overset{T}{\to} \boldsymbol{\Upsilon} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1/3\end{pmatrix}. \tag{S69}$$

Using (S68) together with the results in (S69) it follows that

$$\frac{\bar{\mathbf{W}}'\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}}{T^3} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}\boldsymbol{\Upsilon}\mathbf{Q}'. \tag{S70}$$

From (S56), (S57) and (S70), together with Lemma A.2 we have

$$III = \frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^+\frac{\bar{\mathbf{W}}'\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}}{T^3}\left(\frac{\bar{\mathbf{W}}'\bar{\mathbf{W}}}{T}\right)^+\frac{\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i}{\sqrt{T}}$$

$$\overset{(N,T)_j}{\Longrightarrow} \boldsymbol{\vartheta}_{if}'\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^+\mathbf{Q}\boldsymbol{\Upsilon}\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^+\mathbf{Q}\boldsymbol{\vartheta}_{if}$$

$$= \boldsymbol{\vartheta}_{if}'\boldsymbol{\Upsilon}\boldsymbol{\vartheta}_{if} = \frac{1}{3}[W_i(1)]^2. \tag{S71}$$

Substituting (S62),(S67) and (S71) into (S60), together with (S59), we obtain

$$CSB_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i^2(r)dr + \frac{1}{3}[W_i(1)]^2 - 2W_i(1)\int_0^1 rW_i(r)dr$$

as required. ■

In the intercept only case, using a similar derivation as above it follows that

$$CSB_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i^2(r)dr.$$

S3.2 The Case of Serially Correlated Errors

Consider

$$\Delta y_{it} = \beta_i(y_{i,t-1} - \boldsymbol{\alpha}_{iy}'\mathbf{d}_{t-1}) + \boldsymbol{\alpha}_{iy}'\Delta\mathbf{d}_t + \gamma_{iy}'\mathbf{f}_t + \zeta_{iyt}(\theta_i), \tag{S72}$$

with $\boldsymbol{\alpha}_{iy} = (\alpha_{iy0}, \alpha_{iy1})'$, $\mathbf{d}_t = (1,t)'$ and

$$\zeta_{iyt} = \theta_i\zeta_{iy,t-1} + \eta_{iyt},\ \ |\theta_i| < 1,\ \text{for } i = 1,2,...,N;\ t = 1,2,...,T, \tag{S73}$$

where η_{iyt} is independently distributed across time, with zero mean and a finite positive variance, $\sigma_{i\eta}^2$.

Under the null that $\beta_i = 0$, with $\theta_i = \theta$ (S72) reduces to

$$\Delta y_{it} = \alpha_{iy1} + \gamma_{iy}'\mathbf{f}_t + \zeta_{iyt}(\theta). \tag{S74}$$

Using the lag operator we can write $\zeta_{iyt}(\theta) = (1-\theta L)^{-1}\eta_{iyt}$ so that

$$\Delta y_{it} = (1-\theta)\,\alpha_{iy1} + \theta\Delta y_{i,t-1} + \gamma_{iy}'(\mathbf{f}_t - \theta\mathbf{f}_{t-1}) + \eta_{iyt}. \tag{S75}$$

In matrix notation

$$\Delta\mathbf{y}_i = (1-\theta)\,\alpha_{iy1}\boldsymbol{\tau}_T + \theta\Delta\mathbf{y}_{i,-1} + (\mathbf{F}-\theta\mathbf{F}_{-1})\gamma_{iy} + \boldsymbol{\eta}_{iy}, \tag{S76}$$

where $\Delta\mathbf{y}_{i,-1} = (\Delta y_{i0}, \Delta y_{i1}, ..., \Delta y_{iT-1})'$, $\mathbf{F}_{-1}=(\mathbf{f}_0, \mathbf{f}_1, ..., \mathbf{f}_{T-1})'$, $\boldsymbol{\eta}_{iy} = \left(\eta_{iy1}, \eta_{iy2}, ..., \eta_{iyT}\right)'$ and

$$\Delta\bar{\mathbf{Z}} = \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}_1' + \mathbf{F}\bar{\boldsymbol{\Gamma}}' + \bar{\mathbf{E}},\ \Delta\bar{\mathbf{Z}}_{-1} = \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}_1' + \mathbf{F}_{-1}\bar{\boldsymbol{\Gamma}}' + \bar{\mathbf{E}}_{-1}, \tag{S77}$$

where $\bar{\mathbf{E}} = N^{-1}\sum_{i=1}^N \mathbf{E}_i$, with $\mathbf{E}_i = (\boldsymbol{\zeta}_{iy}'(\theta), \mathbf{E}_{ix}')'$, $\mathbf{E}_{ix} = (\boldsymbol{\varepsilon}_{ix1}, \boldsymbol{\varepsilon}_{ix2}, ..., \boldsymbol{\varepsilon}_{ixT})'$, and $\boldsymbol{\zeta}_{iy}(\theta) = \left(\zeta_{iy1}(\theta), \zeta_{iy2}(\theta), ..., \zeta_{iyT}(\theta)\right)'$, $\Delta\bar{\mathbf{Z}}_{-1} = (\Delta\bar{\mathbf{z}}_0, \Delta\bar{\mathbf{z}}_1, ..., \Delta\bar{\mathbf{z}}_{T-1})'$, and similarly for $\bar{\mathbf{E}}_{-1}$. Substituting $\mathbf{F} = \left(\Delta\bar{\mathbf{Z}} - \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}_1' - \bar{\mathbf{E}}\right)\bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{\Gamma}}'\bar{\boldsymbol{\Gamma}}\right)^{-1}$ and $\mathbf{F}_{-1}=\left(\Delta\bar{\mathbf{Z}}_{-1} - \boldsymbol{\tau}_T\bar{\boldsymbol{\alpha}}_1' - \bar{\mathbf{E}}_{-1}\right)\bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{\Gamma}}'\bar{\boldsymbol{\Gamma}}\right)^{-1}$, which are obtained by (S77), in (S76) yields

$$\Delta\mathbf{y}_i = \mathring{\alpha}_i\boldsymbol{\tau}_T + \theta\Delta\mathbf{y}_{i,-1} + (\Delta\bar{\mathbf{Z}}-\theta\Delta\bar{\mathbf{Z}}_{-1})\boldsymbol{\delta}_i + \sigma_{i\eta}\boldsymbol{\upsilon}_i, \tag{S78}$$

where $\mathring{\alpha}_i = (1-\theta)\left(\alpha_{iy1} - \bar{\boldsymbol{\alpha}}_1'\boldsymbol{\delta}_i\right)$, $\boldsymbol{\delta}_i = \bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{\Gamma}}'\bar{\boldsymbol{\Gamma}}\right)^{-1}\gamma_{iy}$, and

$$\boldsymbol{\upsilon}_i = [\boldsymbol{\eta}_{iy} - (\bar{\mathbf{E}} - \theta\bar{\mathbf{E}}_{-1})\boldsymbol{\delta}_i]/\sigma_{i\eta}.$$

The test of the panel unit root hypothesis using the Sargan-Bhargava statistic is based on the cross section augmented regression

$$\Delta y_{it} = g_{i0} + b_i\Delta y_{i,t-1} + \mathbf{c}_i'\Delta\bar{\mathbf{z}}_t + \mathbf{h}_i'\Delta\bar{\mathbf{z}}_{t-1} + \epsilon_{it},$$

where the cross section augmented Sargan-Bhargava statistic is given by

$$CSB_i(N,T) = T^{-2}\frac{\sum_{t=1}^T \hat{u}_{it}^2}{\hat{\sigma}_i^2}, \tag{S79}$$

with $\hat{u}_{it} = \sum_{s=1}^t \hat{\epsilon}_{is}$, and $\hat{\sigma}_i^2 = \sum_{t=1}^T \hat{\epsilon}_{it}^2/\left(T - 2(k+1) - 2\right)$.

Theorem S3.2 *Suppose the series $\mathbf{z}_{it}$, for $i = 1, 2, ..., N$, $t = 1, 2, ..., T$, is generated under (S47) according to (S77) and $\mathbf{d}_t = (1, \mathbf{t}_T)'$. Then under Assumptions 1-5, as N and $T \to \infty$ such that $\sqrt{T}/N \to 0$, $CSB_i(N,T)$ in (S79) has the same joint $[(N,T)_j \to \infty]$ limit distribution given by (S51) obtained for $\theta = 0$.*

Proof. We have that

$$\hat{\sigma}_i^2 = \frac{\Delta\mathbf{y}_i'\bar{\mathbf{M}}_{i1}\Delta\mathbf{y}_i}{T - 2(k+1) - 2},$$

with $\bar{\mathbf{M}}_{i1}= \mathbf{I}_T - \bar{\mathbf{W}}_{i1}\left(\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}\right)^+\bar{\mathbf{W}}_{i1}'$, $\bar{\mathbf{W}}_{i1} = \left(\Delta\mathbf{y}_{i,-1}, \Delta\bar{\mathbf{Z}}, \Delta\bar{\mathbf{Z}}_{-1}, \boldsymbol{\tau}_T\right)$. Noting that

$$\hat{\boldsymbol{\varepsilon}}_i = \bar{\mathbf{M}}_{i1}\Delta\mathbf{y}_i = \sigma_i\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i,$$

we have

$$CSB_i(N,T) = \frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\mathbf{H}'\mathbf{H}\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i/T^2}{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i/\left(T - 2(k+1) - 2\right)}.$$

Define the matrices $\mathbf{W}_{i1f} = (\boldsymbol{\zeta}_{iy,-1}, \mathbf{F}, \mathbf{F}_{-1}, \boldsymbol{\tau}_T)$ and $\bar{\boldsymbol{\Xi}}_1 = \left(\mathbf{0}_T, \bar{\mathbf{E}}, \bar{\mathbf{E}}_{-1}, \mathbf{0}_T\right)$, so that

$$\bar{\mathbf{W}}_{i1}' = \mathbf{Q}_{1N}\mathbf{W}_{i1f}' + \bar{\boldsymbol{\Xi}}_1', \text{ with } \mathbf{Q}_{1N} = \begin{pmatrix} 1 & \mathbf{0} & \gamma_{iy}' & \alpha_{iy1} \\ \mathbf{0} & \bar{\boldsymbol{\Gamma}} & \mathbf{0} & \bar{\boldsymbol{\alpha}}_1 \\ \mathbf{0} & \mathbf{0} & \bar{\boldsymbol{\Gamma}} & \bar{\boldsymbol{\alpha}}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}. \tag{S80}$$

Also,

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i}{T} = \frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T} - \frac{1}{T}\left(\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_{i1}}{\sqrt{T}}\right)\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+\left(\frac{\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\right).$$

By Lemma A.1

$$\begin{aligned} \frac{\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i}{\sqrt{T}} &= \mathbf{Q}_{1N}\frac{\mathbf{W}_{i1f}'\boldsymbol{\upsilon}_i}{\sqrt{T}} + \frac{\bar{\boldsymbol{\Xi}}_1'\boldsymbol{\upsilon}_i}{\sqrt{T}} \\ &= \mathbf{Q}_{1N}\frac{\mathbf{W}_{i1f}'\boldsymbol{\eta}_{iy}}{\sigma_{i\eta}\sqrt{T}} + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{\sqrt{T}}{N}\right). \end{aligned}$$

$$\begin{aligned}
\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T} &= \left(\mathbf{Q}_{1N}\mathbf{W}_{i1f}'+\bar{\boldsymbol{\Xi}}_1'\right)\left(\mathbf{W}_{i1f}\mathbf{Q}_{1N}'+\bar{\boldsymbol{\Xi}}_1\right) \\
&= \mathbf{Q}_{1N}\frac{\mathbf{W}_{i1f}'\mathbf{W}_{i1f}}{T}\mathbf{Q}_{1N}'+\mathbf{Q}_{1N}\frac{\mathbf{W}_{i1f}'\bar{\boldsymbol{\Xi}}_1}{T} \\
&\quad +\frac{\bar{\boldsymbol{\Xi}}_1'\mathbf{W}_{i1f}}{T}\mathbf{Q}_{1N}'+\frac{\bar{\boldsymbol{\Xi}}_1'\bar{\boldsymbol{\Xi}}_1}{T} \\
&= \mathbf{Q}_{1N}\frac{\mathbf{W}_{i1f}'\mathbf{W}_{i1f}}{T}\mathbf{Q}_{1N}'+O_p\left(\frac{1}{\sqrt{NT}}\right)+O_p\left(\frac{1}{N}\right).
\end{aligned}$$

As

$$\frac{\mathbf{W}_{i1f}'\mathbf{W}_{i1f}}{T}\overset{T}{\to}\begin{pmatrix}
\frac{\boldsymbol{\zeta}_{iy,-1}'\boldsymbol{\zeta}_{iy,-1}}{T} & \frac{\boldsymbol{\zeta}_{iy,-1}'\mathbf{F}}{T} & \frac{\boldsymbol{\zeta}_{iy,-1}'\mathbf{F}_{-1}}{T} & \frac{\boldsymbol{\zeta}_{iy,-1}'\boldsymbol{\tau}_T}{T} \\
\frac{\mathbf{F}'\boldsymbol{\zeta}_{iy,-1}}{T} & \frac{\mathbf{F}'\mathbf{F}}{T} & \frac{\mathbf{F}'\mathbf{F}_{-1}}{T} & \frac{\mathbf{F}'\boldsymbol{\tau}_T}{T} \\
\frac{\mathbf{F}_{-1}'\boldsymbol{\zeta}_{iy,-1}}{T} & \frac{\mathbf{F}_{-1}'\mathbf{F}}{T} & \frac{\mathbf{F}_{-1}'\mathbf{F}_{-1}}{T} & \frac{\mathbf{F}_{-1}'\boldsymbol{\tau}_T}{T} \\
\frac{\boldsymbol{\tau}_T'\boldsymbol{\zeta}_{iy,-1}}{T} & \frac{\boldsymbol{\tau}_T'\mathbf{F}}{T} & \frac{\boldsymbol{\tau}_T'\mathbf{F}_{-1}}{T} & \frac{\boldsymbol{\tau}_T'\boldsymbol{\tau}_T}{T}
\end{pmatrix}$$

$$= \begin{pmatrix}
\frac{\sigma_{\eta i}^2}{1-\theta^2} & \mathbf{0} & \mathbf{0} & 0 \\
\mathbf{0} & \mathbf{I}_{m^0} & \boldsymbol{\Sigma}_{f1}' & \mathbf{0} \\
\mathbf{0} & \boldsymbol{\Sigma}_{f1} & \mathbf{I}_{m^0} & \mathbf{0} \\
0 & \mathbf{0} & \mathbf{0} & 1
\end{pmatrix},$$

$$\frac{\mathbf{W}_{i1f}'\boldsymbol{\eta}_{iy}}{\sigma_{i\eta}\sqrt{T}}=\begin{pmatrix}
\frac{\boldsymbol{\zeta}_{iy,-1}'\boldsymbol{\eta}_{iy}}{\sigma_{i\eta}\sqrt{T}} \\
\frac{\mathbf{F}'\boldsymbol{\eta}_{iy}}{\sigma_{i\eta}\sqrt{T}} \\
\frac{\mathbf{F}_{-1}'\boldsymbol{\eta}_{iy}}{\sigma_{i\eta}\sqrt{T}} \\
\frac{\boldsymbol{\tau}_T'\boldsymbol{\eta}_{iy}}{\sigma_{i\eta}\sqrt{T}}
\end{pmatrix}\overset{T}{\Longrightarrow}\begin{pmatrix}
\sqrt{\frac{\sigma_{\eta i}^2}{1-\theta^2}}W_i(1) \\
\boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\
\boldsymbol{\Lambda}_f\mathbf{W}_{\mathbf{v},i}(1) \\
W_i(1)
\end{pmatrix},$$

we have $\hat{\sigma}_i^2=\frac{T}{T-2(k+1)-2}\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i}{T}\to 1$ as T and $N\to\infty$. Next, since

$$\begin{aligned}
\frac{\mathbf{H}\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i}{T} &= \frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}-\frac{\mathbf{H}\bar{\mathbf{W}}_{i1}\left(\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}\right)^+\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i}{T} \\
&= \frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}-\frac{\mathbf{H}\bar{\mathbf{W}}_{i1}}{T^{3/2}}\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+\frac{\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i}{\sqrt{T}},
\end{aligned}$$

we have

$$\begin{aligned}
\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_{i1}\mathbf{H}'\mathbf{H}\bar{\mathbf{M}}_{i1}\boldsymbol{\upsilon}_i}{T^2} &= \frac{\boldsymbol{\upsilon}_i'\mathbf{H}'\mathbf{H}\boldsymbol{\upsilon}_i}{T^2}-\frac{\boldsymbol{\upsilon}_i'\mathbf{H}'}{T}\frac{\mathbf{H}\bar{\mathbf{W}}_{i1}}{T^{3/2}}\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+\frac{\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i}{\sqrt{T}} \\
&\quad -\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_{i1}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+\frac{\bar{\mathbf{W}}_{i1}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T} \\
&\quad +\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_{i1}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+\frac{\bar{\mathbf{W}}_{i1}'\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}_{i1}}{T^3}\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+\frac{\bar{\mathbf{W}}_{i1}'\boldsymbol{\upsilon}_i}{\sqrt{T}} \\
&= I-2II+III.
\end{aligned}$$

We look at terms I, II and III in turn. Consider I. Noting that we can write

$$\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T}=\frac{\mathbf{s}_{i\eta}-\left(\bar{\mathbf{S}}-\theta\bar{\mathbf{S}}_{-1}\right)\boldsymbol{\delta}_i}{T\sigma_{i\eta}}=\frac{\mathring{\mathrm{s}}_{i1}}{T}, \tag{S81}$$

where $\bar{\mathbf{S}}=N^{-1}\sum_{i=1}^N\mathbf{S}_i$ with $\mathbf{S}_i=(\mathbf{s}_{i1},...,\mathbf{s}_{i,T})'$, using Lemma A.1 we have

$$\begin{aligned}
I &= \frac{\boldsymbol{\upsilon}_i'\mathbf{H}'\mathbf{H}\boldsymbol{\upsilon}_i}{T^2}=\frac{[\mathbf{s}_{i\eta}-\left(\bar{\mathbf{S}}-\theta\bar{\mathbf{S}}_{-1}\right)\boldsymbol{\delta}_i]'[\mathbf{s}_{i\eta}-\left(\bar{\mathbf{S}}-\theta\bar{\mathbf{S}}_{-1}\right)\boldsymbol{\delta}_i]}{\sigma_{i\eta}^2T^2} \\
&= \frac{\mathbf{s}_{i\eta}'\mathbf{s}_{i\eta}}{\sigma_{i\eta}^2T^2}+O_p(\frac{1}{\sqrt{N}})+O_p(\frac{1}{T\sqrt{N}}).
\end{aligned}$$

As $T\to\infty$, $\frac{\mathbf{s}_{i\eta}'\mathbf{s}_{i\eta}}{\sigma_{i\eta}^2T^2}\overset{T}{\Longrightarrow}\int_0^1W_i^2(r)dr$ and it follows that

$$I\overset{(N,T)_j}{\Longrightarrow}\int_0^1W_i^2(r)dr.$$

Now consider II. Firstly, we can write

$$\mathbf{H}\bar{\mathbf{W}}_{i1} = \mathbf{W}_{i1H,f}\mathbf{Q}_N' + \bar{\boldsymbol{\Xi}}_{1H},$$

where $\mathbf{H}\bar{\mathbf{W}}_{i1} = \left(\mathbf{y}_{i,-1} - y_{i,-1}\boldsymbol{\tau}_T, \bar{\mathbf{Z}} - \boldsymbol{\tau}_T\bar{\mathbf{z}}_0', \bar{\mathbf{Z}}_{-1} - \boldsymbol{\tau}_T\bar{\mathbf{z}}_{-1}', \mathbf{t}_T\right)$, $\mathbf{W}_{i1H,f} = (\mathbf{s}_{i\zeta,-1}, \mathbf{S}_f, \mathbf{S}_{f,-1}, \mathbf{t}_T)$ and $\bar{\boldsymbol{\Xi}}_{1H} = \left(\mathbf{0}_T, \bar{\mathbf{S}}, \bar{\mathbf{S}}_{-1}, \mathbf{0}_T\right)$. So then using (S81) we have

$$\frac{\bar{\mathbf{W}}_{i1}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T} = \frac{\mathbf{Q}_N\mathbf{W}_{i1H,f}'\mathring{\mathbf{s}}_{i1}}{T^{5/2}} + \frac{\bar{\boldsymbol{\Xi}}_{1H}'\mathring{\mathbf{s}}_{i1}}{T^{5/2}}.$$

Using the expression for $\mathring{\mathbf{s}}_{i1}$ given by (S81) and Lemma A.1 together with the results that $\mathbf{t}_T'\bar{\mathbf{E}}/T^{3/2} = O_p(N^{-1/2})$ and $\mathbf{t}_T'\bar{\mathbf{S}}/T^{5/2} = O_p(N^{-1/2})$ which follow from a similar derivation of Lemma A.1, we have that

$$\frac{\mathbf{W}_{i1H,f}'\mathring{\mathbf{s}}_{i1}}{T^{5/2}} = \begin{pmatrix} \frac{\mathbf{s}_{i\zeta,-1}{}'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^{5/2}} + O_p(\frac{1}{\sqrt{NT}}) \\ \frac{\mathbf{S}_f'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^{5/2}} + O_p(\frac{1}{\sqrt{NT}}) \\ \frac{\mathbf{S}_{f,-1}'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^{5/2}} + O_p(\frac{1}{\sqrt{NT}}) \\ \frac{\mathbf{t}_T'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^{5/2}} + O_p(\frac{1}{\sqrt{N}}) \end{pmatrix}, \; \frac{\bar{\boldsymbol{\Xi}}_{1H}'\mathring{\mathbf{s}}_{i1}}{T^{5/2}} = O_p(\frac{1}{\sqrt{NT}}) + O_p(\frac{1}{\sqrt{N^2T}}).$$

Noting that (using proposition 17.3 of Hamilton (1994))

$$\frac{\mathbf{s}_{i\zeta,-1}{}'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^2} \xrightarrow{T} \frac{\sigma_{i\eta}}{1-\theta}\int_0^1 [W_i(r)]^2 dr, \; \frac{\mathbf{S}_f'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^2} \xrightarrow{T} \boldsymbol{\Lambda}_f\int_0^1 [\mathbf{W}_\mathbf{v}(r)]\, W_i(r)dr \;, \frac{\mathbf{t}_T'\mathbf{s}_{i\eta}}{\sigma_{i\eta}T^{5/2}} \xrightarrow{T} \int_0^1 rW_i(r)dr$$

as $T \to \infty$, using Lemma A.1 together with the results in (S65) we have, as $(T, N) \xrightarrow{j} \infty$, that

$$\frac{\bar{\mathbf{W}}_{i1}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}\boldsymbol{\kappa}_{if}, \text{ with } \boldsymbol{\kappa}_{if} = \begin{pmatrix} 0 \\ \mathbf{0} \\ \mathbf{0} \\ \int_0^1 rW_i(r)dr \end{pmatrix}.$$

Now, using (S56), (S57) and (S66) and Lemma A.2 it follows that

$$\begin{aligned} II = \frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_{i1}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}_{i1}'\bar{\mathbf{W}}_{i1}}{T}\right)^+ \frac{\bar{\mathbf{W}}_{i1}'\mathbf{H}'}{T^{3/2}}\frac{\mathbf{H}\boldsymbol{\upsilon}_i}{T} &\overset{(N,T)_j}{\Longrightarrow} \boldsymbol{\vartheta}_{if}'\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^+\mathbf{Q}\boldsymbol{\kappa}_{if} \\ = \boldsymbol{\vartheta}_{if}'\boldsymbol{\kappa}_{if} = W_i(1)\int_0^1 rW_i(r)dr. \end{aligned}$$

Finally consider III. Using (S63) we have

$$\frac{\bar{\mathbf{W}}_{i1}'\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}_{i1}}{T^3} = \frac{\mathbf{Q}_N\mathbf{W}_{i1H,f}'\mathbf{W}_{i1H,f}\mathbf{Q}_N'}{T^3} + \frac{\mathbf{Q}_N\mathbf{W}_{i1H,f}'\bar{\boldsymbol{\Xi}}_{1H}}{T^3} + \frac{\bar{\boldsymbol{\Xi}}_{1H}'\mathbf{W}_{i1H,f}\mathbf{Q}_N'}{T^3} + \frac{\bar{\boldsymbol{\Xi}}_{1H}'\bar{\boldsymbol{\Xi}}_{1H}}{T^3},$$

and by Lemma A.1

$$\frac{\mathbf{W}_{i1H,f}'\bar{\boldsymbol{\Xi}}_{1H}}{T^3} = O_p(\frac{1}{\sqrt{T^2N}}), \; \frac{\bar{\boldsymbol{\Xi}}_{1H}'\bar{\boldsymbol{\Xi}}_{1H}}{T^3} = O_p(\frac{1}{NT}).$$

Noting that as $T \to \infty$

$$\begin{aligned} &\frac{\mathbf{s}_{i\zeta}'\mathbf{s}_{i\zeta}}{T^2} \overset{T}{\Longrightarrow} \left(\frac{\sigma_\eta}{1-\theta}\right)^2 \int_0^1 [W_i(r)]^2 dr, \; \frac{\mathbf{S}_f'\mathbf{S}_f}{T} \overset{T}{\Longrightarrow} \boldsymbol{\Lambda}_f\left(\int_0^1 [\mathbf{W}_\mathbf{v}(r)]\,[\mathbf{W}_\mathbf{v}(r)]'\,dr\right)\boldsymbol{\Lambda}_f', \\ &\frac{\mathbf{S}_f'\mathbf{s}_{i\zeta}}{T^2} \overset{T}{\Longrightarrow} \frac{\sigma_\eta}{1-\theta}\boldsymbol{\Lambda}_f\int_0^1 [\mathbf{W}_\mathbf{v}(r)]\,W_i(r)dr, \; \frac{\mathbf{S}_f'\mathbf{t}_T}{T^{5/2}} \overset{T}{\Longrightarrow} \boldsymbol{\Lambda}_f\int_0^1 r\,[\mathbf{W}_\mathbf{v}(r)]\,dr, \\ &\frac{\mathbf{s}_{i\zeta}'\mathbf{t}_T}{T^{5/2}} \overset{T}{\Longrightarrow} \frac{\sigma_\eta}{1-\theta}\int_0^1 rW_i(r)dr, \; \frac{\mathbf{t}_T'\mathbf{t}_T}{T^3} \xrightarrow{T} 1/3, \end{aligned}$$

we have

$$\begin{aligned}
\frac{\mathbf{W}'_{i1H,f}\mathbf{W}_{i1H,f}}{T^3} &= \begin{pmatrix} \frac{\mathbf{s}'_{i\zeta,-1}\mathbf{s}_{i\zeta,-1}}{T^3} & \frac{\mathbf{s}'_{i\zeta,-1}\mathbf{S}_f}{T^3} & \frac{\mathbf{s}'_{i\zeta,-1}\mathbf{S}_{f,-1}}{T^3} & \frac{\mathbf{s}'_{i\zeta,-1}\mathbf{t}_T}{T^3} \\ \frac{\mathbf{S}'_f\mathbf{s}_{i\zeta,-1}}{T^3} & \frac{\mathbf{S}'_f\mathbf{S}_f}{T^3} & \frac{\mathbf{S}'_f\mathbf{S}_{f,-1}}{T^3} & \frac{\mathbf{S}'_f\mathbf{t}_T}{T^3} \\ \frac{\mathbf{S}'_{f,-1}\mathbf{s}_{i\zeta,-1}}{T^3} & \frac{\mathbf{S}'_{f,-1}\mathbf{S}_f}{T^3} & \frac{\mathbf{S}'_{f,-1}\mathbf{S}_{f,-1}}{T^3} & \frac{\mathbf{S}'_{f,-1}\mathbf{t}_T}{T^3} \\ \frac{\mathbf{t}'_T\mathbf{s}_{i\zeta,-1}}{T^3} & \frac{\mathbf{t}'_T\mathbf{S}_f}{T^3} & \frac{\mathbf{t}'_T\mathbf{S}_{f,-1}}{T^3} & \frac{\mathbf{t}'_T\mathbf{t}_T}{T^3} \end{pmatrix} \\
\overset{T}{\to} \boldsymbol{\Upsilon} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 \end{pmatrix}.
\end{aligned}$$

Using (S68) together with the results in (S69) it follows that

$$\frac{\bar{\mathbf{W}}'_{i1}\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}_{i1}}{T^3} \overset{(N,T)_j}{\Longrightarrow} \mathbf{Q}\boldsymbol{\Upsilon}\mathbf{Q}'.$$

From (S56), (S57) and (S70), together with Lemma A.2 we have

$$\begin{aligned}
III = \frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{W}}_{i1}}{\sqrt{T}}\left(\frac{\bar{\mathbf{W}}'_{i1}\bar{\mathbf{W}}_{i1}}{T}\right)^+ \frac{\bar{\mathbf{W}}'_{i1}\mathbf{H}'\mathbf{H}\bar{\mathbf{W}}_{i1}}{T^3}\left(\frac{\bar{\mathbf{W}}'_{i1}\bar{\mathbf{W}}_{i1}}{T}\right)^+ \frac{\bar{\mathbf{W}}'_{i1}\boldsymbol{\upsilon}_i}{\sqrt{T}} \\
\overset{(N,T)_j}{\Longrightarrow} \boldsymbol{\vartheta}'_{if}\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^+\mathbf{Q}\boldsymbol{\Upsilon}\mathbf{Q}'\left(\mathbf{Q}\mathbf{Q}'\right)^+\mathbf{Q}\boldsymbol{\vartheta}_{if} \\
= \boldsymbol{\vartheta}'_{if}\boldsymbol{\Upsilon}\boldsymbol{\vartheta}_{if} = \frac{1}{3}[W_i(1)]^2.
\end{aligned}$$

Substituting (S62),(S67) and (S71) into (S60), together with (S59), we obtain

$$CSB_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i^2(r)dr + \frac{1}{3}[W_i(1)]^2 - 2W_i(1)\int_0^1 rW_i(r)dr$$

as required. ■

In the intercept only case, using a similar derivation as above it follows that

$$CSB_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \int_0^1 W_i^2(r)dr.$$

S4 Panel Unit Root Test Statistics Considered in the Empirical Application

S4.1 The $P_{\hat{e}}$ Tests of Bai and Ng (2004)

The pooled test statistics proposed by Bai and Ng (2004) are based on PANIC residuals computed using the following transformations of y_{it},

$$\underline{\Delta y}_{it} = \begin{cases} \Delta y_{it}, & \text{for the case with an intercept} \\ \Delta y_{it} - \overline{\Delta y}_i, & \text{for the case with an intercept and a linear trend} \end{cases} \tag{S82}$$

where $\overline{\Delta y}_i = T^{-1}\sum_{t=1}^T \Delta y_{it}$. The principal components of $\underline{\Delta y}_{it}$ are used to estimate $\mathbf{F}$, denoted as $\hat{\mathbf{F}}$, which is $\sqrt{T}$ times the m^0 (assumed number of factors) eigenvectors corresponding to the m^0 largest eigenvalues of the $T \times T$ matrix $\underline{\Delta \mathbf{Y}}\,\underline{\Delta \mathbf{Y}}'$, where $\underline{\Delta \mathbf{Y}} = (\underline{\Delta \mathbf{y}}_1, \underline{\Delta \mathbf{y}}_2, ..., \underline{\Delta \mathbf{y}}_N)$, with $\underline{\Delta \mathbf{y}}_i = (\underline{\Delta y}_{i1}, \underline{\Delta y}_{i2}, ..., \underline{\Delta y}_{iT})'$. Under the normalisation $\hat{\mathbf{F}}'\hat{\mathbf{F}}/T = \mathbf{I}_{m^0}$, the estimates of the factor loadings are given by $\hat{\gamma}_{iy} = \hat{\mathbf{F}}'\underline{\Delta \mathbf{y}}_i/T$, which yield the residuals $\hat{\varepsilon}_{iyt} = \underline{\Delta y}_{it} - \hat{\gamma}'_{iy}\hat{\mathbf{f}}_t$. The PANIC residuals are then computed as

$$\hat{s}_{iyt} = \sum_{s=1}^{t} \hat{\varepsilon}_{iys}. \tag{S83}$$

Theses PANIC residuals are then used to compute the ADF statistic based on the ADF(p) regressions in $\hat{s}_{iyt}$ without deterministics for each cross section unit, i.

The expressions for the $P_{\hat{e}}$ test statistics depending on the panel's deterministics is given by:

With an Intercept:

$$P_{\hat{e}} = \frac{\left(-2\sum_{i=1}^{N}\ln(pv_i^c) - 2N\right)}{\sqrt{4N}},$$

where pv_i^c is the p-values of the ADF statistic for the ADF(p) regressions in $\hat{s}_{iyt}$ without deterministics for each cross section unit. The p-values are obtained using the tables 'adfnc.asc' provided by Serena Ng.

With an Intercept and a Linear Trend:

$$P_{\hat{e}} = \frac{\left(-2\sum_{i=1}^{N}\ln(pv_i^\tau) - 2N\right)}{\sqrt{4N}},$$

where pv_i^τ is the p-values of the ADF statistic for the ADF(p) regressions in $\hat{s}_{iyt}$ without deterministics for each cross section unit. The p-values are obtained using the tables 'lm1.asc' provided by Serena Ng.

These statistics are asymptotically distributed as standard normal so that the null hypothesis is rejected at the 5% level, for example, if $P_{\hat{e}}$ is larger than 1.645.[1]

The variants of $P_{\hat{e}}$ that we consider make use of all the available variables, y_{it} and $\mathbf{x}_{it}$, when computing the principal components. This version is more directly comparable to the test proposed in PSY which makes use of the additional variables, $\mathbf{x}_{it}$. The procedure is similar to that described above with the principal component estimator of $\mathbf{F}$ now computed using $\underline{\Delta\mathbf{z}}_{it} = (\underline{\Delta y}_{it}, \underline{\Delta\mathbf{x}}'_{it})'$, where $\underline{\Delta\mathbf{x}}_{it}$ is constructed from $\Delta\mathbf{x}_{it}$ in a manner similar to that specified by (S82) for $\underline{\Delta y}_{it}$. These variants are denoted by $P_{\hat{e},z}$.

S4.2 The $PMSB$ and P_b Tests of Bai and Ng (2010)

Bai and Ng (2010) propose the $PMSB$ and P_b tests, both of which are briefly described below. The former is the panel version of the modified Sargan-Bhargava test, while the latter is the analog of the t_b^* statistic of Moon and Perron (2004) except that it is based on a different set of residuals and the method of 'defactoring' of the data is different . The $PMSB$ and P_b tests are based on the so called PANIC residuals, which in the context of the notation as set out in Section 2 of PSY, are obtained as follows.

As in Section S4.1, transform Δy_{it} then obtain the PANIC residuals defined by (S83). Following Moon and Perron (2004), the long-run variances are estimated by means of the Andrews–Monahan (Andrews and Monahan, 1992) estimator using the quadratic spectral kernel and pre-whitening.

S4.2.1 P_b Test

The P_b test is then based on a pooled estimate of the autoregressive coefficient ρ in the following regression

$$\hat{s}_{iyt} = \rho\hat{s}_{iy,t-1} + \varepsilon_{iyt}. \tag{S84}$$

where $\hat{s}_{iyt}$ is the PANIC residual defined by (S83). Let

$$\hat{\sigma}_\epsilon^2 = N^{-1}\sum_{i=1}^{N}\hat{\sigma}_{\epsilon i}^2,\ \hat{\omega}_\epsilon^2 = N^{-1}\sum_{i=1}^{N}\hat{\omega}_{\epsilon i}^2,\ \hat{\lambda}_\epsilon = N^{-1}\sum_{i=1}^{N}\hat{\lambda}_{\epsilon i},\ \hat{\phi}_\epsilon^4 = N^{-1}\sum_{i=1}^{N}\hat{\omega}_{\epsilon i}^4 \tag{S85}$$

where $\hat{\sigma}_{\epsilon i}^2$, $\hat{\omega}_{\epsilon i}^2$, and $\hat{\lambda}_{\epsilon i} = \left(\hat{\omega}_{\epsilon i}^2 - \hat{\sigma}_{\epsilon i}^2\right)/2$, are the estimators of the variance, the long-run variance, and the one-sided long-run variance of ε_{iyt}, respectively.

The expression for the test statistic depending on the panel's deterministics is given as follows:

With an Intercept:

$$P_b = \sqrt{N}T(\hat{\rho}^+ - 1)\sqrt{\frac{1}{NT^2}\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iy,t-1}^2\right)\frac{\hat{\omega}_\epsilon^2}{\hat{\phi}_\epsilon^4}},$$

where

$$\hat{\rho}^+ = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iy,t-1}\hat{s}_{iyt} - NT\hat{\lambda}_\epsilon}{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iy,t-1}^2}.$$

[1] Bai and Ng (2010; p.1093) indicate that a two-tailed test is employed for the $P_{\hat{e}}^c$ and $P_{\hat{e}}^\tau$ tests. However, right-tailed tests are appropriate for such pooled tests which are based on the p-values; see Choi (2001), for example.

With an Intercept and a Linear Trend:

$$P_b = \sqrt{N}T(\hat{\rho}^+ - 1)\sqrt{\frac{1}{NT^2}\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iy,t-1}^2\right)\frac{5}{6}\frac{\hat{\omega}_\epsilon^6}{\hat{\phi}_\epsilon^4\hat{\sigma}_\epsilon^4}},$$

where $\hat{\omega}_{\epsilon i}^6 = \left(\hat{\omega}_{\epsilon i}^2\right)^3$, and

$$\hat{\rho}^+ = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iy,t-1}\hat{s}_{iyt}}{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iy,t-1}^2} + \frac{3}{T}\frac{\hat{\sigma}_\epsilon^2}{\hat{\omega}_\epsilon^2}.$$

Under the null hypothesis these statistics tend to a standard normal distribution as $N, T \to \infty$ with $N/T \to 0$. The null hypothesis is rejected if P_b is smaller than -1.645 (at the 5% level).

S4.2.2 $PMSB$ Test

The expressions for the $PMSB$ statistic depending on the deterministics are as follows:

With an Intercept:

$$PMSB = \frac{\sqrt{N}\left(\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iyt}^2 - \hat{\omega}_\epsilon^2/2\right)}{\sqrt{\hat{\phi}_\epsilon^4/3}}.$$

With an Intercept and a Linear Trend:

$$PMSB = \frac{\sqrt{N}\left(\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{s}_{iyt}^2 - \hat{\omega}_\epsilon^2/6\right)}{\sqrt{\hat{\phi}_\epsilon^4/45}},$$

where $\hat{s}_{iyt}$ is the PANIC residuals defined by (S83), $\hat{\omega}_\epsilon^2$ and $\hat{\phi}_\epsilon^4$ are defined by (S85).

Under the null hypothesis the above statistics tend to a standard normal distribution as $N, T \to \infty$ with $N/T \to 0$. The null hypothesis is rejected if $PMSB$ is less than -1.645 (at the 5% level).

S4.3 The t_b^* Test of Moon and Perron (2004) for the Case of an Intercept Only

The t_b^* test is defined similar to the P_b statistic of Bai and Ng though it is based on defactored panel data, obtained by projecting the panel data onto the space orthogonal to the (estimated) factor loadings.

Keeping in line with the notation in PSY consider the model

$$\begin{aligned} y_{it} &= \alpha_i + y_{it}^0, \\ y_{it}^0 &= \rho_i y_{i,t-1}^0 + u_{it} \\ u_{it} &= \gamma_{iy}'\mathbf{f}_t + \varepsilon_{iyt}. \end{aligned}$$

Consider the residuals from a pooled regression of y_{it} on y_{it-1},

$$\hat{e}_{it} = y_{it} - \hat{\rho}y_{it-1} \text{ with } \hat{\rho} = \sum_{i=1}^{N}\sum_{t=1}^{T}y_{it}y_{it-1}/\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it-1}^2. \tag{S86}$$

Assuming that the second moment of α_i is bounded, since the stochastic trend term y_{it}^0 dominates α_i, for the purpose of estimating ρ the presence of α_i can be ignored (see p.86 of Moon and Perron, 2004). Moon and Perron propose to apply principal components to $\hat{e}_{it}$, in order to extract the factors and their loadings, $\hat{\gamma}_{iy}$. The residuals $\hat{e}_{it}$ are then defactored by projecting them onto the space orthogonal to the estimated factor loadings.

Define a $N \times 1$ residual vector $\hat{\mathbf{e}}_t = (\hat{e}_{1t}, \hat{e}_{2t}, ..., \hat{e}_{Nt})'$ and a $N \times N$ projection matrix $\mathbf{Q}_{\hat{\gamma}} = \mathbf{I}_N - \hat{\gamma}(\hat{\gamma}'\hat{\gamma})^{-1}\hat{\gamma}'$ where $\hat{\gamma}$ is a $N \times m^0$ factor loading matrix $\hat{\gamma} = (\hat{\gamma}_{1y}, \hat{\gamma}_{2y}, ..., \hat{\gamma}_{Ny})'$, so that

$$\tilde{\mathbf{e}}_t = \mathbf{Q}_{\hat{\gamma}}\hat{\mathbf{e}}_t. \tag{S87}$$

The t_b^* test statistic is defined by

$$t_b^* = \sqrt{N}T(\hat{\rho}_{pool}^* - 1)\sqrt{\frac{1}{NT^2}\sum_{t=1}^{T}\tilde{\mathbf{e}}_{t-1}'\tilde{\mathbf{e}}_{t-1}\frac{\hat{\omega}_\epsilon^2}{\hat{\phi}_\epsilon^4}},$$

where

$$\hat{\rho}_{pool}^* = \frac{\sum_{t=1}^{T}\tilde{\mathbf{e}}_{t-1}'\tilde{\mathbf{e}}_t - NT\hat{\lambda}_\epsilon}{\sum_{t=1}^{T}\tilde{\mathbf{e}}_{t-1}'\tilde{\mathbf{e}}_{t-1}},$$

and $\hat{\lambda}_\epsilon$, $\hat{\omega}_\epsilon^2$ and $\hat{\phi}_\epsilon^4$ are the estimators of the long-run variances defined by (S85), but they are based on the residuals $\tilde{\mathbf{e}}_t$ defined by (S87) rather than the PANIC residuals. The null hypothesis is rejected if t_b^* is less than -1.645 (at the 5% level).

S4.4 Constant Point Optimal (CPO) and Ploberger-Phillips (PP) Tests of Moon, Perron and Phillips (2007; MPP)

Initially, in Sections S4.4.1 and S4.4.2, we introduce the CPO and PP tests in the simple case where the errors are cross sectionally independent and serially uncorrelated. These tests are then extended to the case where the errors follow a factor structure and are serially correlated in Sections S4.4.3 and S4.4.4.

S4.4.1 CPO Test of MPP

Following the notations in PSY, the model considered is given by

$$y_{it} = \mathbf{a}_i'\mathbf{d}_t + y_{it}^0,\ t = 0, 1, ..., T,\ i = 1, 2, ..., N$$

where $\mathbf{d}_t = (1, t)'$ and $\mathbf{a}_i = (a_{0i}, a_{1i})'$

$$y_{it}^0 = \rho_i y_{i,t-1}^0 + u_{it}. \tag{S88}$$

Define a homogeneous local alternative $\rho_i = \rho_c$, which depends on the specification of the deterministics (as defined below), such that $\rho_c = 1$ when $c = 0$ and $\rho_c \to 1$ as N or T tends to infinity, so that

$$\mathbf{y}_i = (y_{i0}, y_{i1}, ..., y_{iT})',\ T + 1 \times 1 \tag{S89}$$

$$\mathbf{\Delta}_c = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\rho_c & 1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & 0 & 0 \\ \vdots & & -\rho_c & 1 & 0 \\ 0 & \cdots & 0 & -\rho_c & 1 \end{bmatrix}.$$

and $\mathbf{\Delta}_0 = \mathbf{\Delta}_c(\rho_c = 1)$. $\mathbf{\Delta}_c$ and $\mathbf{\Delta}_0$ are $(T+1) \times (T+1)$ matrices. Similarly define

$$\mathbf{a}_i = \mathbf{a}_{ci} \text{ when } \rho_i = \rho_c \text{ and } \mathbf{a}_i = \mathbf{a}_{0i} \text{ when } \rho_i = 1.$$

The Case With an Intercept only Consider the homogeneous local alternative

$$\rho_c = 1 - \frac{c}{N^{1/2}T}.$$

Define

$$L_c\left(a_{ci}, \sigma_i^2\right) = \sum_{i=1}^{N} -\frac{1}{2\sigma_i^2}\left[\mathbf{\Delta}_c\left(\mathbf{y}_i - \boldsymbol{\tau}_{T+1}a_{ci}\right)\right]'\left[\mathbf{\Delta}_c\left(\mathbf{y}_i - \boldsymbol{\tau}_{T+1}a_{ci}\right)\right],$$

$$L_0\left(a_{0i}, \sigma_i^2\right) = \sum_{i=1}^{N} -\frac{1}{2\sigma_i^2}\left[\mathbf{\Delta}_0\left(\mathbf{y}_i - \boldsymbol{\tau}_{T+1}a_{0i}\right)\right]'\left[\mathbf{\Delta}_0\left(\mathbf{y}_i - \boldsymbol{\tau}_{T+1}a_{0i}\right)\right],$$

where $\boldsymbol{\tau}_{T+1}$ is a $(T+1) \times 1$ vector of ones. The derivative of $L_c\left(a_{ci}, \sigma_i^2\right)$ with respect to a_{ci} is given by

$$\frac{\partial L_c\left(a_{ci}, \sigma_i^2\right)}{\partial a_{ci}} = \sum_{i=1}^{N}\frac{1}{\sigma_i^2}\left(\mathbf{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\left[\mathbf{\Delta}_c\mathbf{y}_i - \mathbf{\Delta}_c\boldsymbol{\tau}_{T+1}a_{ci}\right],$$

so that the first order condition for the i^{th} unit solves

$$\frac{1}{\sigma_i^2}\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\left[\boldsymbol{\Delta}_c\mathbf{y}_i-\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\hat{a}_{ci}\right]=\mathbf{0}$$

$$\frac{1}{\sigma_i^2}\left[\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\left(\boldsymbol{\Delta}_c\mathbf{y}_i\right)-\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\hat{a}_{ci}\right]=\mathbf{0}.$$

It follows that

$$\hat{a}_{ci}=\left[\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right]^{-1}\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\left(\boldsymbol{\Delta}_c\mathbf{y}_i\right).$$

Noting that

$$\boldsymbol{\Delta}_c\mathbf{y}_i=\begin{pmatrix} y_{i0} \\ y_{i1}-\rho_c y_{i0} \\ y_{i2}-\rho_c y_{i1} \\ \vdots \\ y_{iT}-\rho_c y_{i,T-1}\end{pmatrix},\ \boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}=\begin{pmatrix} 1 \\ 1-\rho_c \\ 1-\rho_c \\ \vdots \\ 1-\rho_c\end{pmatrix},$$

we have

$$\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\left(\boldsymbol{\Delta}_c\mathbf{y}_i\right)=y_{i0}+\left(1-\rho_c\right)\sum_{t=1}^{T}\left(y_{it}-\rho_c y_{it-1}\right)$$

and

$$\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}=1+T\left(1-\rho_c\right)^2.$$

Therefore

$$\begin{aligned}\hat{a}_{ci}&=\left[\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right]^{-1}\left(\boldsymbol{\Delta}_c\boldsymbol{\tau}_{T+1}\right)'\left(\boldsymbol{\Delta}_c\mathbf{y}_i\right)\\&=\frac{y_{i0}+\left(1-\rho_c\right)\sum_{t=1}^{T}\left(y_{it}-\rho_c y_{it-1}\right)}{1+T\left(1-\rho_c\right)^2}.\end{aligned}$$

Note that

$$\hat{a}_{0i}=y_{i0}, \tag{S90a}$$

since under the null $\rho_c=1$. Therefore it is easily seen that

$$\hat{\sigma}_{0i}^2=\hat{\sigma}_i^2=\left[\boldsymbol{\Delta}_0\left(\mathbf{y}_i-\boldsymbol{\tau}_{T+1}\hat{a}_{0i}\right)\right]'\left[\boldsymbol{\Delta}_0\left(\mathbf{y}_i-\boldsymbol{\tau}_{T+1}\hat{a}_{0i}\right)\right] \tag{S91}$$

$$=\frac{\sum_{t=1}^{T}\left(y_{it}-y_{i,t-1}\right)^2}{T}. \tag{S92}$$

The scaled feasible likelihood ratio test statistic is given by (c.f. the bottom of p.424 of MPP 2007)

$$CPO_2=\frac{1}{\sqrt{2c^2}}\left\{-2\left[L_c\left(\hat{a}_{ci},\hat{\sigma}_{0i}^2\right)-L_0\left(\hat{a}_{0i},\hat{\sigma}_{0i}^2\right)\right]-\frac{1}{2}c^2\right\}. \tag{S93}$$

Note that $\min_b L_c\left(a_{ci},\sigma_i^2\right)$ and $\min_b L_0\left(a_{0i},\sigma_i^2\right)$ at the bottom of p.424 of MPP 2007 are replaced by $L_c\left(\hat{a}_{ci},\hat{\sigma}_{0i}^2\right)$ and $L_0\left(\hat{a}_{0i},\hat{\sigma}_{0i}^2\right)$, respectively.

It is shown that, under the null hypothesis, as $N,T\to\infty$ with $N/T\to 0$,

$$CPO_2\to N\left(0,1\right).$$

The null hypothesis is rejected if CPO_2 is smaller than -1.645 (at the 5% level). In the experiment in PSY, the value of c is set to 1.

The Case With an Intercept and Trend Consider the homogeneous local alternative

$$\rho_c=1-\frac{c}{N^{1/4}T}.$$

Define

$$\mathbf{D}=\left(\boldsymbol{\tau}_{T+1},\mathrm{t}_{T+1}\right),\ \mathrm{t}_{T+1}=\left(0,1,2,...,T\right)', \tag{S94}$$

so that

$$L_c\left(\mathbf{a}_{ci},\sigma_i^2\right)=\sum_{i=1}^{N}-\frac{1}{2\sigma_i^2}\left[\boldsymbol{\Delta}_c\left(\mathbf{y}_i-\mathbf{D}\mathbf{a}_{ci}\right)\right]'\left[\boldsymbol{\Delta}_c\left(\mathbf{y}_i-\mathbf{D}\mathbf{a}_{ci}\right)\right].$$

$$L_0\left(\mathbf{a}_{0i},\sigma_i^2\right)=\sum_{i=1}^{N}-\frac{1}{2\sigma_i^2}\left[\mathbf{\Delta}_0\left(\mathbf{y}_i-\mathbf{D}\mathbf{a}_{0i}\right)\right]'\left[\mathbf{\Delta}_c\left(\mathbf{y}_i-\mathbf{D}\mathbf{a}_{0i}\right)\right].$$

Following the same line of derivation as in the intercept only case it follows that

$$\frac{\partial L_c\left(\mathbf{a}_{ci},\sigma_i^2\right)}{\partial \mathbf{a}_{ci}}=\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\left(\mathbf{\Delta}_c\mathbf{D}\right)'\left[\mathbf{\Delta}_c\left(\mathbf{y}_i-\mathbf{D}\mathbf{a}_{ci}\right)\right],$$

and so the first order condition for the i^{th} unit solves

$$\frac{1}{\sigma_i^2}\left(\mathbf{\Delta}_c\mathbf{D}\right)'\left[\mathbf{\Delta}_c\left(\mathbf{y}_i-\mathbf{D}\hat{\mathbf{a}}_{ci}\right)\right]=\mathbf{0},$$

$$\frac{1}{\sigma_i^2}\left[\left(\mathbf{\Delta}_c\mathbf{D}\right)'\left(\mathbf{\Delta}_c\mathbf{y}_i\right)-\left(\mathbf{\Delta}_c\mathbf{D}\right)'\mathbf{\Delta}_c\mathbf{D}\hat{\mathbf{a}}_{ci}\right]=\mathbf{0},$$

and thus

$$\hat{\mathbf{a}}_{ci}=\left[\left(\mathbf{\Delta}_c\mathbf{D}\right)'\mathbf{\Delta}_c\mathbf{D}\right]^{-1}\left(\mathbf{\Delta}_c\mathbf{D}\right)'\left(\mathbf{\Delta}_c\mathbf{y}_i\right).$$

But

$$\begin{aligned}\mathbf{\Delta}_c\mathbf{D}&=\begin{bmatrix}1 & 0 & \cdots & 0 & 0\\ -\rho_c & 1 & \ddots & & \vdots\\ 0 & \ddots & \ddots & 0 & 0\\ \vdots & & -\rho_c & 1 & 0\\ 0 & \cdots & 0 & -\rho_c & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\\ 1 & 2\\ \vdots & \vdots\\ 1 & T\end{bmatrix}\\ &=\begin{bmatrix}1 & 0\\ 1-\rho_c & 1\\ 1-\rho_c & 2-\rho_c\\ \vdots & \vdots\\ 1-\rho_c & T-(T-1)\rho_c\end{bmatrix},\end{aligned}$$

so

$$\left(\mathbf{\Delta}_c\mathbf{D}\right)'\mathbf{\Delta}_c\mathbf{D}=\begin{bmatrix}1+T\left(1-\rho_c\right)^2 & \left(1-\rho_c\right)\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]\\ \left(1-\rho_c\right)\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right] & \sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]^2\end{bmatrix}$$

and

$$\left(\mathbf{\Delta}_c\mathbf{D}\right)'\mathbf{\Delta}_c\mathbf{y}_i=\begin{bmatrix}y_{i0}+\left(1-\rho_c\right)\sum_{t=1}^{T}\left(y_{it}-\rho_c y_{it-1}\right)\\ \sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]\left(y_{it}-\rho_c y_{it-1}\right)\end{bmatrix}.$$

Therefore

$$\begin{aligned}\hat{\mathbf{a}}_{ci}&=\left[\left(\mathbf{\Delta}_c\mathbf{D}\right)'\mathbf{\Delta}_c\mathbf{D}\right]^{-1}\left(\mathbf{\Delta}_c\mathbf{D}\right)'\left(\mathbf{\Delta}_c\mathbf{y}_i\right)\\ &=\begin{bmatrix}\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]^2 & -\left(1-\rho_c\right)\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]\\ -\left(1-\rho_c\right)\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right] & 1+T\left(1-\rho_c\right)^2\end{bmatrix}\times\\ &\begin{bmatrix}y_{i0}+\left(1-\rho_c\right)\sum_{t=1}^{T}y_{it}-\rho_c y_{it-1}\\ \sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]\left(y_{it}-\rho_c y_{it-1}\right)\end{bmatrix}\times\frac{1}{q_c}\\ &=\frac{1}{q_c}\begin{pmatrix}h_{c1i}\\ h_{c2i}\end{pmatrix}\end{aligned}$$

with

$$\begin{aligned}q_c=&\left\{\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]^2\right\}\left[1+T\left(1-\rho_c\right)^2\right]\\ &-\left(1-\rho_c\right)^2\left\{\sum_{t=1}^{T}\left[t-\rho_c(t-1)\right]\right\}^2,\end{aligned}$$

and

$$h_{c1i} = \left\{\sum_{t=1}^{T}[t-\rho_c(t-1)]^2\right\}\left[y_{i0}+(1-\rho_c)\sum_{s=1}^{T}(y_{is}-\rho_c y_{is-1})\right]$$
$$-\left\{(1-\rho_c)\sum_{t=1}^{T}[t-\rho_c(t-1)]\right\}\left\{\sum_{s=1}^{T}[s-\rho_c(s-1)](y_{is}-\rho_c y_{is-1})\right\}$$

$$h_{c2i} = \left[1+T(1-\rho_c)^2\right]\left\{\sum_{t=1}^{T}[t-\rho_c(t-1)](y_{it}-\rho_c y_{it-1})\right\}$$
$$-\left\{(1-\rho_c)\sum_{t=1}^{T}[t-\rho_c(t-1)]\right\}\left[y_{i0}+(1-\rho_c)\sum_{s=1}^{T}(y_{is}-\rho_c y_{is-1})\right].$$

When $\rho_c = 1$, noting that $q_{ci} = T$

$$\hat{\mathbf{a}}_{0i} = \begin{pmatrix} y_{i0} \\ T^{-1}\sum_{t=1}^{T}(y_{it}-y_{it-1}) \end{pmatrix} = \begin{pmatrix} y_{i0} \\ T^{-1}(y_{iT}-y_{i0}) \end{pmatrix} \tag{S95}$$

which coincides with the first equation of Section 5.2 in MPP. To compute the feasible statistics, firstly σ_i^2 is replaced by

$$\hat{\sigma}_{0i}^2 = T^{-1}\left[\mathbf{\Delta}_0(\mathbf{y}_i-\mathbf{D}\hat{\mathbf{a}}_{0i})\right]'\left[\mathbf{\Delta}_0(\mathbf{y}_i-\mathbf{D}\hat{\mathbf{a}}_{0i})\right]. \tag{S96}$$

The scaled feasible likelihood ratio test statistic (c.f. p.427 of MPP 2007) is given by

$$CPO_3 = \frac{1}{\sqrt{\frac{c^4}{45}}}\left\{-2\left[L_c\left(\hat{\mathbf{a}}_{ci},\hat{\sigma}_{0i}^2\right)-L_0\left(\hat{\mathbf{a}}_{0i},\hat{\sigma}_{0i}^2\right)\right]+w\right\} \tag{S97}$$
$$w = \frac{Nc}{N^{1/4}}+\omega_{p2T}\frac{Nc^2}{N^{1/2}}+\omega_{p4T}\frac{Nc^4}{N}$$

with

$$\omega_{p2T} = -\frac{1}{T}\sum_{t=1}^{T}\frac{t-1}{T}+\frac{2}{T}\sum_{t=1}^{T}\frac{t}{T}\left(\frac{t-1}{T}\right)-\frac{1}{3}$$
$$\omega_{p4T} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{t-1}{T}\frac{s-1}{T}\min\left(\frac{t-1}{T},\frac{s-1}{T}\right)$$
$$-\frac{2}{3}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{t-1}{T}\right)^2+\frac{1}{9}.$$

Note that $\min_{\mathbf{a}} L_c\left(\mathbf{a}_{ci},\sigma_i^2\right)$ and $\min_{\mathbf{a}} L_0\left(\mathbf{a}_{0i},\sigma_i^2\right)$ at the bottom of p.427 of MPP 2007 are replaced by $L_c\left(\hat{\mathbf{a}}_{ci},\hat{\sigma}_{0i}^2\right)$ and $L_0\left(\hat{\mathbf{a}}_{0i},\hat{\sigma}_{0i}^2\right)$, respectively.

It is shown that, under the null hypothesis, as $N,T\to\infty$ with $N/T\to 0$,

$$CPO_3 \to N(0,1).$$

The null hypothesis is rejected if CPO_3 is smaller than -1.645 (at the 5% level). In the experiment in PSY, the value of c is set to 1.

S4.4.2 Ploberger Phillips (PP) Test

This test is used in the case of a linear trend:

$$PP = \sqrt{45N}\left[\frac{1}{NT^2}\sum_{i=1}^{N}(\mathbf{y}_i-\mathbf{D}\hat{\mathbf{a}}_{0i})'(\mathbf{y}_i-\mathbf{D}\hat{\mathbf{a}}_{0i})/\hat{\sigma}_{0i}^2-\omega_{1T}\right]$$
$$\omega_{1T} = T^{-1}\sum_{t=1}^{T}\frac{t}{T}\left(1-\frac{t}{T}\right),$$

where $\mathbf{D}$, $\hat{\mathbf{a}}_{0i}$ and $\hat{\sigma}_{0i}^2$ are defined as in equations (S94), (S95) and (S96). It is shown that $PP\to N(0,1)$.

The null hypothesis is rejected if PP is smaller than -1.645 (at the 5% level).

S4.4.3 CPO and PP Tests Under an Error Factor Structure

When u_{it} in (S88) contains a factor structure, namely

$$u_{it} = \sum_{\ell=1}^{m^0} \gamma_{\ell iy} f_{\ell t} + \varepsilon_{it} = \gamma_{iy}' \mathbf{f}_t + \varepsilon_{iyt},$$

we follow the procedure set out in Section 6.3 of MPP:

1. Compute $\hat{\mathbf{y}}_i^0 = \mathbf{y}_i - \boldsymbol{\tau}_{T+1}\hat{a}_{0i}$ (with an intercept) or $\hat{\mathbf{y}}_i^0 = \mathbf{y}_i - \mathbf{D}\hat{a}_{0i}$ (with a trend), where $\mathbf{y}_i$, $\hat{a}_{0i}$, $\mathbf{D}$ and $\hat{a}_{0i}$ are defined by (S89), (S90a), (S94) and (S95). Define $\hat{\mathbf{y}}_i^* = \left(\hat{y}_{i1}^0, \hat{y}_{i2}^0, ..., \hat{y}_{iT}^0\right)'$ and $\hat{\mathbf{y}}_{i,-1}^* = \left(\hat{y}_{i0}^0, \hat{y}_{i1}^0, ..., \hat{y}_{iT-1}^0\right)'$.
2. Following MPP, ρ is estimated by the pooled OLS estimator (see Moon et al., 2007; p.422-3):

   (a) For the case of an intercept (p.425)

   $$\hat{\rho}_{pool}^{+} = \left(\sum_{i=1}^{N} \frac{\hat{\mathbf{y}}_{i,-1}^{*\prime}\hat{\mathbf{y}}_{i,-1}^{*}}{\hat{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{N} \frac{\hat{\mathbf{y}}_{i,-1}^{*\prime}\hat{\mathbf{y}}_{i}^{*}}{\hat{\sigma}_i^2} + \frac{3}{T}$$

   (b) For the case of a trend (p.432)

   $$\hat{\rho}_{pool}^{+} = \left(\sum_{i=1}^{N} \frac{\hat{\mathbf{y}}_{i,-1}^{*\prime}\hat{\mathbf{y}}_{i,-1}^{*}}{\hat{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{N} \frac{\hat{\mathbf{y}}_{i,-1}^{*\prime}\hat{\mathbf{y}}_{i}^{*}}{\hat{\sigma}_i^2} + \frac{7.5}{T}$$

   with $\hat{\sigma}_i^2 = T^{-1}\left(\boldsymbol{\Delta}_0\hat{\mathbf{y}}_i^0\right)'\left(\boldsymbol{\Delta}_0\hat{\mathbf{y}}_i^0\right)$, following the definitions of $\hat{\sigma}_{2,iT}^2$ and $\hat{\sigma}_{3,iT}^2$ in Moon et al. (2007; p.422 & 428), assuming no error serial correlation.[2]

3. Given the number of factors, m^0, m^0 principal components and associated factor loadings are extracted from $\hat{u}_{it} = \hat{y}_{it}^0 - \hat{\rho}_{pool}^{+}\hat{y}_{i,t-1}^0$; see (S88).
4. Use the CPO tests (as well as the PP tests) on orthgonalised $\mathbf{y}_i$, namely, on $\mathbf{Y}'\mathbf{Q}_{\tilde{\gamma}}$, where $\mathbf{Y}' = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N)$ and $\mathbf{Q}_{\tilde{\gamma}} = \mathbf{I}_N - \tilde{\gamma}(\tilde{\gamma}'\tilde{\gamma})^{-1}\tilde{\gamma}'$ where $\tilde{\gamma}$ is a $N \times m^0$ factor loading matrix $\tilde{\gamma} = (\tilde{\gamma}_{1y}, \tilde{\gamma}_{2y}, ..., \tilde{\gamma}_{Ny})'$.

S4.4.4 CPO and PP Test Under Error Serial Correlation

In the case of error serial correlation, following Section 6.4 of MPP, the estimators of σ_i^2 above are replaced by their long-run variance counterparts. Following Moon and Perron (2004), the long-run variances are estimated based on the Andrews and Monahan (1992) method using the quadratic spectral kernel and pre-whitening. For further details, see Moon and Perron (2004).

S5 Point Optimal Panel Unit Root Test with Serially Correlated Errors of Moon, Perron and Phillips (2011)

For the generalised point optimal panel unit root test of Moon et al. (2011), denoted by $\widetilde{CPO}$ in PSY, first the y_{it} series of interest is defactored as described in (S4.4.3), and the $\widetilde{CPO}$ test is then applied to the defactored data. The $\widetilde{CPO}$ test is computed in the same way as the CPO test described in (S4.4.1) where the estimators of σ_i^2 are replaced by their long-run variance counterparts, and in addition the centering of (S93) and (S97) is adjusted to accommodate the second-order bias induced by the correlation between the error and lagged values of the dependent variable as suggested by Moon et al. (2011).

[2] When there is error serial correlation, these variances are to be replaced by the long-run variance estimators; see Section S4.4.4.

S6 Small Sample Performance: Monte Carlo Evidence

In what follows we investigate by means of Monte Carlo simulations the small sample properties of the $CIPS$ and CSB tests defined in PSY, and compare their performance to the tests proposed in the literature described above. Specifically, we consider the pooled test statistic $P_{\hat{e}}$ of Bai and Ng (2004) based on the PANIC residuals, a panel version of the modified Sargan–Bhargava test (denoted by $PMSB$) and a PANIC residual-based Moon and Perron (2004) type test (denoted by P_b), both of which are proposed by Bai and Ng (2010), the t_b^* statistic of Moon and Perron (2004) for the case of an intercept only,[3] the PP statistic which is a defactored version of the optimal invariant test of Ploberger and Phillips (2002) for the case of an intercept and a linear trend, and the CPO test, that is the defactored version of the common point optimal test of Moon, Perron and Phillips (2007). The theory of the CPO test is developed by Moon et al. for the serially uncorrleated case, but it is claimed (see Section 6.4 in Moon et al. (2007, p. 436)), that replacing variances in their CPO statistic with long-run variances should result in a test with a correct size under quite general short memory error autocorrelations. However, our preliminary experiments suggested that this claim might not be valid. Upon communicating these results to the authors, Moon, Perron and Phillips provided us with another modification of the CPO test that appropriately allows for residual serial correlation (see Moon, Perron and Phillips, 2011). In addition to replacing the variance of the errors by the long run variance, in this recent paper Moon et al. also adjust the centering of the statistic to accommodate for the second-order bias induced by the correlation between the error and lagged values of the dependent variable. In the Monte Carlo simulations reported below we only include the modified CPO test, denoted by $\widetilde{CPO}$.

The $P_{\hat{e}}$ test is defined in Section 2.4 of Bai and Ng (2004, p.1140), the t_b^* test in Section 2.2.2 of Moon and Perron (2004, p.91), the P_b and $PMSB$ tests in Section 3, p.1094, eq. (9) and Section 3.1, p.1095, eq.(11), respectively of Bai and Ng (2010), the CPO and PP tests in Section 4.1, p.424; Section 5.1, p.427; and Section 5.3.1, p.429, eq. (20), respectively, in Moon et al. (2007), and the $\widetilde{CPO}$ test in Section 2.2, p.4; Section 2.3, p.5, of Moon et al. (2011). In computing the CPO and $\widetilde{CPO}$ test statistics we set the constant term (the 'c' term in Moon et al.) to unity. Also, following Moon and Perron (2004), the long-run variances for the $PMSB$, P_b, t_b^*, PP, CPO and $\widetilde{CPO}$ test statistics are estimated by means of the Andrews and Monahan (1992) method using the quadratic spectral kernel and prewhitening. See Moon and Perron (2004) for further details.

The details of the computation of the critical values for the $CIPS$ and CSB tests are set out in Section 4.2. Both the $CIPS$ and CSB tests reject the null when the value of the statistic is smaller than the relevant critical value, at the chosen level of significance. We do not report size adjusted results, since such results are likely to have limited value in empirical applications. See, for example, Horowitz and Savin (2000).

S6.1 Monte Carlo Design

In their Monte Carlo experiments Bai and Ng (2010, Section 5) set $m^0 = 1$ and do not allow for serial correlation in the idiosyncratic errors. Here we consider a more general set up and allow for two factors ($m^0 = 2$), and also consider experiments where the idiosyncratic errors are serially correlated. Following Bailey, Kapetanios and Pesaran (2012) we generate one of the factors in the y_{it} equations as strong and the second factor as semi-strong. Accordingly, the data generating process (DGP) for the $\{y_{it}\}$ is given by

$$y_{it} = d_{iyt} + \rho_i y_{i,t-1} + \gamma_{iy1} f_{1t} + \gamma_{iy2} f_{2t} + \varepsilon_{iyt}, i = 1, 2, ..., N; t = -49, ..., T, \tag{S98}$$

with $y_{i,-50} = 0$, where $\gamma_{iy1} \sim iidU[0, 2]$, for $i = 1, 2, ..., N$; $\gamma_{iy2} \sim iidU[0, 1]$ for $i = 1, ..., [N^\alpha]$, and $\gamma_{iy2} = 0$ for $i = [N^\alpha] + 1, [N^\alpha] + 2, ..., N$ (where $[\cdot]$ denotes the integer part); $f_{\ell t} \sim iidN(0, 1)$ for $\ell = 1, 2$, $\varepsilon_{iyt} \sim iidN(0, \sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5, 1.5]$. The exponent of cross-sectional dependence of the first (strong) factor is 1, and for the second (semi-strong) factor, it is set to 0.75, guided by the empirical results reported in Bailey et al. (2012). See, also Chudik et al. (2011).

At the stage of implementing the tests, we assume that $m_{\max} = 2$, and hence set $k = m_{\max} - 1 = 1$. The additional regressor, x_{it}, is generated as

$$\Delta x_{it} = d_{ix} + \gamma_{ix1} f_{1t} + \varepsilon_{ixt}, \tag{S99}$$

where

$$\varepsilon_{ixt} = \rho_{ix} \varepsilon_{ixt-1} + \varpi_{ixt}, , \varpi_{ixt} \sim iidN(0, 1 - \rho_{ix}^2), \tag{S100}$$

[3]The t_a^* test of Moon and Perron (2004) is not included since they summarise the experimental results saying "in almost all cases, the test based on the t_b^* statistic has better size properties." Similarly, the P_a test of Bai and Ng (2010) is not included.

$i = 1, 2, ..., N; t = -49, ..., T$, with $\varepsilon_{ix,-50} = 0$, and $\rho_{ix} \sim iidU[0.2, 0.4]$. The factor loadings in (S99) are generated as $\gamma_{ix1} \sim iidU[0, 2]$, so that

$$E(\mathbf{\Gamma}_i) = \begin{pmatrix} 1 & \frac{1}{2}N^{-0.25} \\ 1 & 0 \end{pmatrix}, \tag{S101}$$

and hence the rank condition is satisfied when N is finite, but fails when $N \to \infty$. In this way we also check the robustness of the $CIPS$ and CSB tests to failure of the rank condition for sufficiently large N.

We considered two specifications for the deterministics in y_{it} and x_{it}. For the case of an intercept only, $d_{iyt} = (1 - \rho_i)\alpha_{iy}$ with $\alpha_{iy} \sim iidN(1, 1)$ and $d_{ix} = 0$; for the case of an intercept and a linear trend, $d_{iyt} = \mu_{iy} + (1 - \rho_i)\delta_i t$ with $\mu_{iy} \sim iidU[0.0, 0.02]$ and $\delta_i \sim iidU[0.0, 0.02]$, and $d_{ix} = \delta_{ix}$ with $\delta_{ix} \sim iidU[0.0, 0.02]$.

To examine the impact of the residual serial correlation on the proposed tests we consider the DGP in which the idiosyncratic errors ε_{iyt} are generated as

$$\varepsilon_{iyt} = \rho_{iy\varepsilon}\varepsilon_{iyt-1} + (1 - \rho_{iy\varepsilon}^2)^{1/2}\eta_{iyt}, \text{ for } t = -49, -48, ..., 0, 1, ..., T, \tag{S102}$$

with $\varepsilon_{iy,-50} = 0$, where $\eta_{iyt} \sim iidN(0, \sigma_i^2)$, and $\sigma_i^2 \sim iidU[0.5, 1.5]$. We considered a positively serially correlated case, $\rho_{iy\varepsilon} \sim iidU[0.2, 0.4]$, as well as a negatively serially correlated case, $\rho_{iy\varepsilon} \sim iidU[-0.4, -0.2]$. The first 50 observations are discarded.

The parameters α_{iy},δ_i,μ_{iy}, δ_{ix}, $\rho_{iy\varepsilon}$, γ_{iy1},γ_{iy2}, ρ_i, γ_{ix1}, ρ_{ix}, and σ_i are redrawn over each replication. The DGP is given by (S98) with $\rho_i = \rho = 1$ for size, and $\rho_i \sim iidU[0.90, 0.99]$ for power. All tests are conducted at the 5% significance level. All combinations of $N, T = 20, 30, 50, 70, 100, 200$ are considered, and all experiments are based on 2,000 replications each.

In the case where the errors of y_{it} are serially correlated, lag augmentation is required for the asymptotic validity of the $CIPS$ and CSB tests as well as the pooled tests of Bai and Ng (2004). For these tests, in the Monte Carlo results that follow, lag augmentation is selected according to $\hat{p} = \left[4(T/100)^{1/4}\right]$ (where $[\cdot]$ denotes the integer part). For the other tests, the statistics are adjusted using a non-parametric estimator of the long run variance. In our Monte Carlo results we use the long run variance of Andrews and Monahan (1992). Also note that the asymptotic normality of the $PMSB$, $P_{\hat{e}}$, P_b, t_b^*, PP, CPO and $\widetilde{CPO}$ test statistics require $N/T \to 0$ as N and T go to infinity, while the asymptotic validity of the $CIPS$ and CSB tests only requires that $\sqrt{T}/N \to 0$, which allows N and T to expand at the same rate.

S6.2 Results

Size and power of the tests are summarised in Tables S1 to S6. We do not report size adjusted results, since such results are likely to have limited value in empirical applications. See, for example, Horowitz and Savin (2000). Table S1 provides the results for the panel with an intercept only, and with serially uncorrelated idiosyncratic errors. The size properties of the $P_{\hat{e}}$, t_b^*, and P_b tests are very similar: they tend to over-reject the null moderately across combinations of N and T, with the extent of over-rejection rising as N increases. These results are consistent with those reported in Gengenbach, Palm and Urbain (2009) and Bai and Ng (2010). The $\widetilde{CPO}$ test, has good size properties when T is larger than N, but these tests begin to show serious size distortions as N increases relative to T, which is in line with the condition $N/T \to 0$ that underlies the theory of these tests. The $PMSB$ test of Bai and Ng (2010) tends to under-reject the null when T and N are small, which is in accordance with the results reported in Bai and Ng (2010, Table 1). For example, when $T = N = 20$, the estimated size is 0.65% at the 5% nominal level. In contrast, the $CIPS$ and CSB tests have the correct size for all combinations of sample sizes, even when T is small relative to N. In terms of power, the CSB test has satisfactory power which is almost consistently higher than that of $CIPS$, though most of the other tests do tend to display higher power (which could partly be due to the over-sized nature of the other tests). An exception is the $PMSB$ test for small values of T and N, which exhibits lower power than the CSB test.

The results for the case with a linear trend are summarised in Table S2. The tendency of the over-rejection of $P_{\hat{e}}$ for small T is more serious than for the case with an intercept only. For example, even when $T = 200$ and $N = 100$, the size of $P_{\hat{e}}$ is 8.4%. The size of the defactored version of the Ploberger and Phillips test, the PP test, which is only considered for the case with an intercept and a linear trend, is close to the nominal level only when T is much larger than N. The size distortion of the P_b test is similar to that for the case of an intercept only case, though somewhat less pronounced. The over-rejection tendency of the $\widetilde{CPO}$ test is now even more pronounced as compared to the intercept only case. The $PMSB$ test is now even more under-sized. When $T = N = 20$, the size of the $PMSB$ test is 0.20%, and even when $N = T = 100$, the size of the $PMSB$ is 1.85% at the 5% nominal level. Again, the $CIPS$ and CSB tests have the correct size for all combinations of sample sizes and their power rises in N and T, as to be expected. Power discrepancies between the CSB and $CIPS$ tests are less pronounced in this case, with the former still showing higher power than the latter. The other tests have higher power than these two tests, but given their size distortions a straightforward power comparison would be problematic. The $PMSB$ test continues to be an exception for smaller values of T, where now the power of this test is almost negligible for $T = 20$, and for $T = 30$ the power ranges from 0.85 to 2.75 across different values of N. Even when $T = 70$, the CSB test has greater power than the $PMSB$ test, for small N.

Tables S3 and S4 present the results for the case where ε_{iyt} are positively serially correlated for the intercept only and linear trend cases, respectively. The results for the case where ε_{iyt} are negatively serially correlated are summarised in Tables S5 and S6. The effect of allowing for residual serial correlation on the $P_{\hat{e}}$, P_b, PP and $\widetilde{CPO}$ tests is to accentuate the tendency of these tests to over-reject the null. Positive serial correlation in ε_{iyt} seems to be more problematic for the size of these tests as compared to negative serial correlation. The $\widetilde{CPO}$ test has good size properties for values of $T > N$, although it continues to show significant size distortions when $N > T$. The $PMSB$ test, in the case of positive serial correlation, shows some tendency to over-reject for small T and large N. By contrast, the effect of negative serial correlation on the $PMSB$ test is relatively minor, but as in the serially uncorrelated case reported in Tables S1 and S2, the $PMSB$ test tends to under-reject. The size and power of the $CIPS$ and CSB tests are not much affected by residual serial correlation once the underlying regressions are augmented with lagged changes. As the results in Tables S1-S6 show, the $CIPS$ and CSB tests do not display any size distortions for all values of N and T, irrespective of whether the idiosyncratic errors are serially correlated or not.

Overall, the $CIPS$ and CSB tests perform well in most cases, always having the correct size. The evidence on power is mixed, with no one test dominating, and the outcomes difficult to compare due to the size distortion of some of the tests, and the fact that the power of the tests are differently affected by the number of factors and the choice of factor loadings.

Table S1: Size and Power of Alternative Panel Unit Root Tests with Two Factors, Factors and Idiosyncratic Errors are Serially Uncorrelated, $m^0 = 2$ Known, With an Intercept Only

	Size: $\rho_i = \rho = 1$						**Power**: $\rho_i \sim iidU[0.90, 0.99]$					
(T, N)	20	30	50	70	100	200	20	30	50	70	100	200
$CIPS(\hat{p}, k = 1)$												
20	5.75	6.40	5.10	5.50	5.50	6.10	7.80	10.70	10.85	13.15	11.95	14.85
30	5.40	6.60	5.35	5.70	5.85	6.15	11.40	13.65	17.10	17.10	18.55	21.85
50	5.00	5.60	5.90	6.10	4.80	5.90	17.35	22.10	27.10	27.50	32.05	38.40
70	5.45	4.85	4.60	5.70	5.35	5.25	27.95	33.40	40.75	47.45	50.00	56.35
100	5.65	7.05	6.10	4.95	5.75	5.45	44.65	54.45	67.10	68.20	78.60	82.15
200	4.95	4.55	5.60	5.65	4.85	4.80	97.40	99.50	99.95	99.95	100.00	100.00
$CSB(\hat{p}, k = 1)$												
20	6.35	6.10	5.60	4.95	5.80	6.10	14.25	15.80	18.50	23.45	24.80	31.20
30	5.70	5.85	5.20	5.60	5.55	4.10	20.50	24.80	31.70	36.80	40.50	46.95
50	6.35	6.00	5.80	5.85	5.55	5.55	39.20	47.75	62.20	70.30	77.25	87.70
70	5.70	5.80	6.35	6.15	5.75	5.60	61.40	75.40	89.55	94.30	98.00	99.50
100	4.55	5.20	5.95	6.10	5.40	6.60	79.05	89.65	97.95	98.70	99.60	99.95
200	6.50	4.75	6.15	5.15	6.20	5.85	94.85	97.80	99.45	99.90	99.95	100.00
$P_{\hat{e}}(\hat{p})$												
20	10.50	10.15	13.40	13.05	14.15	19.65	23.45	28.05	35.60	42.30	53.40	74.60
30	9.40	8.40	9.05	8.35	7.45	11.00	30.45	39.30	52.10	64.75	76.90	93.85
50	8.65	8.45	9.25	9.25	10.40	10.35	59.10	70.60	88.30	94.35	97.50	99.50
70	6.65	7.55	7.85	7.90	8.05	8.65	77.00	89.60	97.50	98.70	99.75	100.00
100	7.20	7.10	6.95	6.20	6.10	6.70	90.80	97.70	99.65	99.90	99.95	100.00
200	7.25	6.60	6.75	5.85	5.75	6.50	99.80	100.00	100.00	100.00	100.00	100.00
$PMSB$												
20	0.65	1.10	1.35	1.10	2.00	3.55	3.95	6.25	11.25	16.20	23.25	46.25
30	1.15	1.25	1.45	1.60	1.60	2.00	10.50	20.20	35.55	50.95	68.05	89.60
50	1.45	1.85	1.90	2.35	2.05	2.35	41.25	61.30	84.95	92.30	96.60	98.90
70	1.85	2.40	2.55	2.40	2.25	1.85	68.05	85.25	96.25	98.00	99.25	99.75
100	2.10	3.10	3.50	2.60	3.10	2.65	88.20	97.40	99.25	99.85	99.90	100.00
200	3.05	2.90	3.15	3.80	3.60	2.85	99.50	99.95	100.00	100.00	100.00	100.00
P_b												
20	8.65	8.65	9.50	9.40	11.65	19.35	28.95	35.45	51.60	63.00	76.30	93.20
30	7.35	7.70	7.55	8.10	8.60	12.70	47.80	60.95	78.55	86.65	94.30	98.70
50	7.55	6.95	7.60	6.05	7.80	8.95	77.90	88.55	96.05	98.00	98.85	99.60
70	7.05	7.50	6.95	7.00	7.25	5.95	90.45	95.75	99.20	99.20	99.85	100.00
100	7.25	6.60	7.15	6.70	6.00	7.20	96.80	99.45	99.80	99.95	100.00	100.00
200	8.30	6.75	6.45	6.15	5.55	5.65	99.95	100.00	100.00	100.00	100.00	100.00
t_b^*												
20	10.45	10.05	13.10	13.75	18.00	20.50	82.75	91.30	97.00	97.80	98.45	99.55
30	10.35	9.65	10.80	10.50	13.55	16.65	93.25	96.55	99.05	99.05	99.80	99.75
50	7.65	9.05	7.95	7.95	9.95	11.35	98.05	99.40	99.70	99.95	100.00	100.00
70	8.10	7.85	7.80	8.20	9.10	10.05	99.30	99.80	99.90	99.90	100.00	100.00
100	7.95	7.50	7.70	7.35	7.85	7.70	99.90	100.00	100.00	100.00	100.00	100.00
200	8.20	6.65	6.55	7.05	6.25	6.85	100.00	100.00	100.00	100.00	100.00	100.00
$\widetilde{CPO}$												
20	7.80	10.15	14.45	18.60	23.50	39.80	32.50	46.45	65.00	75.30	83.95	94.35
30	7.85	8.10	11.65	13.60	16.65	26.95	48.20	64.20	82.00	88.70	94.75	97.30
50	6.45	5.35	8.25	9.45	11.85	16.45	74.65	87.30	95.75	97.65	99.05	99.15
70	6.10	6.20	7.85	8.55	10.35	13.05	87.80	95.45	98.65	98.95	99.60	99.95
100	5.45	6.10	7.50	7.80	9.15	12.90	95.95	99.15	99.75	99.90	99.85	100.00
200	5.90	5.25	6.70	6.40	7.20	10.75	99.85	99.95	100.00	100.00	100.00	100.00

Notes: y_{it} is generated as $y_{it} = d_{iyt} + \rho_i y_{i,t-1} + \gamma_{iy1} f_{1t} + \gamma_{iy2} f_{2t} + \varepsilon_{iyt}, i = 1, 2, ..., N; t = -49, 48, ...0, 1, ..., T$, with $y_{i,-50} = 0$, where $\gamma_{iy1} \sim iidU[0, 2]$, for $i = 1, 2, ..., N$; $\gamma_{iy2} \sim iidU[0, 1]$ for $i = 1, ..., [N^\alpha]$ and $\gamma_{iy2} = 0$ for $i = [N^\alpha] + 1, [N^\alpha] + 2, ..., N$ (where $[\cdot]$ denotes the integer part); $f_{\ell t} \sim iidN(0, 1)$ for $\ell = 1, 2$, $\varepsilon_{iyt} \sim iidN(0, \sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5, 1.5]$; $\Delta x_{it} = d_{ix} + \gamma_{ix1} f_{1t} + \varepsilon_{ixt}$, where, $d_{ix} = 0$, $\varepsilon_{ixt} = \rho_{ix}\varepsilon_{ixt-1} + \varpi_{ixt},$, $\varpi_{ixt} \sim iidN(0, 1 - \rho_{ix}^2)$, $i = 1, 2, ..., N; t = -49, 48, ...0, 1, ..., T$, with $\varepsilon_{ix,-50} = 0$, and $\rho_{ix} \sim iidU[0.2, 0.4]$. The factor loadings in (S99) are generated as $\gamma_{ix1} \sim iidU[0, 2]$; $d_{iyt} = (1 - \rho_i)\alpha_{iy}$ with $\alpha_{iy} \sim iidN(1, 1)$. The parameters α_{iy}, $\rho_{iy\varepsilon}$, γ_{iy1},γ_{iy2}, ρ_i, γ_{ix1}, ρ_{ix}, and σ_i are redrawn over each replication. The first 50 observations are discarded. The $CIPS(\hat{p})$ and $CSB(\hat{p})$ tests are the proposed panel unit root tests, defined by (28) and (34), respectively, based on cross section augmentation using y_{it} and x_{it} with lag-augmentation order selected according to $\hat{p} = \lceil 4(T/100)^{1/4} \rceil$. $P_{\hat{e}}(\hat{p})$ is the test of Bai and Ng (2004) with lag-augmentation order $\hat{p} = \lceil 4(T/100)^{1/4} \rceil$, and $PMSB$ and P_b are the pooled tests of Bai and Ng (2010), all of which are based on two extracted factors from y_{it},. The t_b^* test is the Moon and Perron (2004) test, and the $\widetilde{CPO}$ is the defactored point optimal test with serially correlated errors of Moon, Perron and Phillips (2011), based on two extracted factors from y_{it}. The $PMSB$, P_b, t_b^*, $\widetilde{CPO}$ tests use the automatic lag-order selection for the estimation of the long-run variances following Andrews and Monahan (1992). All tests are conducted at the 5% significance level, and the $CIPS(\hat{p})$ and $CSB(\hat{p})$ tests are based on the critical values for the corresponding $\hat{p} = \left[4(T/100)^{1/4}\right]$ and the number of additional regressors, k. All experiments are based on 2000 replications.

Table S2: Size and Power of Alternative Panel Unit Root Tests with Two Factors, Factors and Idiosyncratic Errors are Serially Uncorrelated, $m^0 = 2$ Known, With an Intercept and a Linear Trend

	Size: $\rho_i = \rho = 1$						**Power**: $\rho_i \sim iidU[0.90, 0.99]$					
(T, N)	20	30	50	70	100	200	20	30	50	70	100	200
$CIPS(\hat{p}, k = 1)$												
20	6.45	5.20	6.30	6.30	5.45	5.50	7.25	6.55	7.85	7.85	5.80	8.05
30	5.30	5.40	5.90	6.80	5.85	5.45	6.85	8.15	9.00	10.45	11.95	11.75
50	6.35	5.45	5.65	6.10	5.85	5.35	10.00	10.40	13.00	14.00	17.90	20.75
70	5.55	5.50	5.60	5.20	4.65	4.65	14.70	17.40	22.15	25.75	26.65	31.35
100	5.20	5.90	6.30	5.25	5.00	5.10	23.45	29.60	37.85	39.40	46.45	52.10
200	5.60	5.70	5.65	5.30	6.15	3.75	83.80	91.25	97.85	99.25	99.80	99.95
$CSB(\hat{p}, k = 1)$												
20	6.35	5.40	5.80	5.15	5.20	5.65	8.60	8.85	11.55	12.10	13.35	19.25
30	6.80	6.15	5.80	5.95	5.85	5.70	10.65	12.10	14.45	18.45	20.65	25.80
50	5.95	5.80	5.20	5.60	4.50	5.80	15.50	19.15	23.50	29.65	33.55	41.75
70	6.05	4.95	5.90	5.70	5.85	5.25	25.50	33.60	46.45	54.70	65.75	80.40
100	4.65	5.55	5.80	6.35	5.45	5.00	44.15	58.25	75.85	84.95	91.95	97.90
200	5.40	5.10	5.10	6.20	6.15	5.75	87.20	94.85	98.75	99.60	99.85	100.00
$P_{\hat{e}}(\hat{p})$												
20	15.25	18.00	21.45	21.65	29.05	36.30	17.40	19.25	25.10	26.35	32.30	43.50
30	12.25	11.95	12.65	14.75	14.80	19.90	15.75	17.25	19.50	24.00	25.85	40.25
50	10.80	10.95	12.75	10.95	13.40	17.70	20.95	25.50	34.55	39.05	47.80	71.90
70	8.85	9.20	10.35	11.40	12.70	12.95	30.00	39.35	52.50	64.80	75.65	92.85
100	7.60	7.45	8.00	7.75	7.35	6.50	45.75	58.55	76.50	85.70	91.40	98.70
200	8.40	7.45	7.25	8.20	8.40	7.75	94.20	98.45	99.80	99.90	100.00	100.00
$PMSB$												
20	0.20	0.25	0.25	0.45	0.30	0.75	0.40	0.20	0.15	0.55	0.35	0.75
30	0.35	0.50	0.35	0.75	0.95	0.55	0.85	1.40	1.70	2.10	2.80	2.75
50	1.45	1.30	1.35	1.00	0.85	0.90	7.05	9.10	14.65	19.20	26.20	48.00
70	1.55	1.55	1.25	1.40	1.65	0.90	16.20	24.85	42.00	54.10	68.70	88.20
100	2.30	2.60	2.55	2.30	1.85	1.65	41.20	58.55	80.10	89.30	92.10	97.95
200	3.45	2.90	2.35	3.10	3.20	2.60	90.50	96.60	98.90	99.45	99.80	99.90
P_b												
20	5.80	5.65	6.20	6.25	8.35	9.55	7.80	7.35	9.05	9.30	10.00	15.50
30	6.05	6.35	5.90	6.50	5.90	7.25	10.00	10.90	13.20	16.65	19.50	29.70
50	7.45	5.25	6.25	4.85	5.60	6.65	23.35	28.55	37.45	44.30	54.80	77.65
70	7.65	5.90	6.20	5.20	5.05	4.95	37.30	48.60	63.30	72.95	82.70	94.60
100	7.70	6.60	5.90	6.00	5.05	4.80	63.10	75.70	89.10	94.45	95.30	98.70
200	7.60	5.80	5.65	5.65	5.10	5.35	95.15	97.80	99.35	99.55	100.00	99.95
PP												
20	0.65	0.35	1.55	0.85	1.45	2.60	1.25	0.75	2.35	1.65	2.75	5.10
30	1.00	1.00	1.25	1.65	2.10	2.45	2.00	3.50	4.55	5.95	8.15	13.55
50	2.20	2.25	2.60	1.40	2.15	2.95	11.10	14.75	23.65	28.75	39.85	60.75
70	2.45	2.30	1.85	2.75	3.30	3.55	22.65	33.95	48.80	61.00	74.05	87.55
100	3.20	3.05	3.05	3.60	3.90	4.00	47.70	66.10	84.35	90.50	92.30	96.95
200	3.75	3.30	3.60	4.90	4.30	6.05	92.20	97.10	99.05	99.55	99.90	99.90
$\widetilde{CPO}$												
20	12.80	18.40	32.60	41.45	51.20	74.00	15.15	24.65	41.55	51.80	62.65	82.35
30	8.15	11.80	16.45	24.65	32.60	52.80	14.15	21.80	33.05	46.50	58.25	79.20
50	5.65	7.45	11.50	12.75	16.85	29.45	22.90	32.10	50.15	62.25	74.50	87.55
70	4.40	5.80	7.00	9.50	13.15	20.75	32.65	49.65	67.65	78.65	88.40	94.55
100	4.45	4.65	6.75	8.10	9.85	15.40	54.50	73.20	89.40	94.45	95.10	98.30
200	3.85	3.75	5.05	7.05	6.85	10.25	92.35	97.40	99.20	99.75	99.90	99.95

Notes: y_{it} is generated as described in the note to Table S1, but $d_{iyt} = \mu_{iy} + (1 - \rho_i)\delta_i t$ with $\mu_{iy} \sim iidU[0.0, 0.02]$ and $\delta_i \sim iidU[0.0, 0.02]$, and $d_{ixt} = \delta_{ix}$ with and $\delta_{ix} \sim iidU[0.0, 0.02]$. The PP test is a defactored version of the optimal invariant test of Ploberger and Phillips (2002), based on two extracted factors from y_{it}. See also the notes to Table S1 for the specification of the rest of the parameters and the test statistics.

Table S3: Size and Power of Alternative Panel Unit Root Tests with Two Factors, Factors are Serially Uncorrelated but Idiosyncratic Errors are Positively Serially Correlated, $m^0 = 2$ Known, With an Intercept Only

	Size: $\rho_i = \rho = 1$						**Power**: $\rho_i \sim iidU[0.90, 0.99]$					
(T, N)	20	30	50	70	100	200	20	30	50	70	100	200
$CIPS(\hat{p}, k=1)$												
20	5.00	5.65	4.05	4.30	3.80	4.15	7.40	8.65	8.50	10.70	9.25	11.65
30	4.40	5.45	3.85	4.20	4.15	4.55	9.65	11.45	14.55	14.80	16.00	18.75
50	4.30	5.30	5.25	4.70	3.90	5.15	16.55	20.40	24.05	24.65	28.75	34.60
70	4.90	5.00	4.45	5.00	4.35	4.30	26.10	30.55	37.55	44.40	45.15	51.50
100	5.45	6.20	5.60	4.10	5.55	4.95	41.95	51.10	62.85	62.65	74.60	78.25
200	4.75	4.45	5.05	5.55	4.65	4.55	96.45	99.10	99.80	100.00	100.00	100.00
$CSB(\hat{p}, k=1)$												
20	6.85	6.40	6.15	6.30	6.75	6.20	13.40	15.30	17.70	23.25	23.65	30.55
30	5.40	6.25	5.65	5.65	6.00	4.45	18.65	22.95	28.75	33.85	37.70	43.90
50	5.90	5.60	5.95	5.65	5.65	6.10	36.65	43.55	58.60	67.40	74.25	86.10
70	5.15	6.15	5.50	5.85	5.20	5.80	60.45	74.85	90.65	95.70	99.05	99.85
100	4.35	4.80	5.75	5.75	5.15	6.30	80.35	90.90	98.60	99.50	99.90	100.00
200	6.35	4.40	5.40	5.10	5.65	5.45	97.25	99.10	99.75	100.00	100.00	100.00
$P_{\hat{e}}(\hat{p})$												
20	12.00	15.40	18.15	18.00	21.65	31.00	22.40	24.10	30.30	37.05	44.20	63.60
30	10.05	9.90	11.55	12.05	10.60	15.60	30.00	35.75	48.60	61.20	75.45	93.45
50	8.55	9.10	9.35	9.45	10.60	11.90	58.45	71.60	89.40	95.35	98.60	99.80
70	7.35	6.95	8.20	7.70	8.55	9.80	78.05	91.15	98.55	99.45	99.90	100.00
100	7.20	7.35	6.25	6.40	5.70	6.95	93.05	98.50	100.00	100.00	100.00	100.00
200	7.50	6.45	5.65	6.70	5.60	6.65	99.95	100.00	100.00	100.00	100.00	100.00
$PMSB$												
20	1.70	3.80	8.45	7.60	12.50	21.50	3.40	5.65	10.40	13.50	21.85	42.55
30	1.35	2.90	3.85	5.15	6.55	10.05	8.80	17.30	32.80	49.25	67.65	92.50
50	1.80	2.10	2.15	2.75	3.25	5.75	40.10	61.20	87.40	94.90	98.85	99.90
70	1.70	2.35	2.65	2.60	2.95	3.50	68.70	88.25	98.20	99.45	100.00	100.00
100	2.00	2.95	3.55	3.00	3.80	3.45	90.60	98.60	99.90	100.00	100.00	100.00
200	2.85	2.50	3.25	3.85	3.55	3.15	99.90	100.00	100.00	100.00	100.00	100.00
P_b												
20	9.65	12.60	16.65	17.80	21.95	32.85	21.00	24.60	39.15	46.30	60.05	83.95
30	7.15	8.35	10.80	11.45	13.30	19.85	38.95	52.40	72.90	83.55	94.60	99.45
50	7.00	6.30	7.20	6.95	8.25	11.40	76.70	88.10	97.55	99.10	99.65	99.90
70	6.25	6.50	7.05	7.05	7.75	6.85	90.95	97.25	99.80	99.85	99.95	100.00
100	6.65	6.75	6.55	5.95	5.70	7.50	97.95	99.85	100.00	100.00	100.00	100.00
200	8.65	6.85	6.00	5.80	5.60	5.75	99.95	100.00	100.00	100.00	100.00	100.00
t_b^*												
20	8.70	8.20	11.35	12.05	15.70	20.65	80.45	90.70	97.95	98.95	99.70	99.95
30	7.85	7.50	8.95	9.00	10.60	13.30	92.40	96.95	99.45	99.65	99.95	100.00
50	6.25	6.85	6.45	6.65	7.70	8.45	98.55	99.80	100.00	99.95	100.00	100.00
70	6.70	6.30	5.85	5.85	7.00	6.75	99.75	99.95	100.00	100.00	100.00	100.00
100	6.85	6.70	6.55	5.75	5.55	6.00	100.00	100.00	100.00	100.00	100.00	100.00
200	7.40	6.10	6.30	6.55	5.45	6.05	100.00	100.00	100.00	100.00	100.00	100.00
$\widetilde{CPO}$												
20	8.00	11.30	16.35	23.85	29.45	51.20	32.80	46.20	68.15	81.35	89.20	97.60
30	6.75	8.25	12.20	13.90	18.55	33.55	48.85	67.15	85.25	93.15	97.75	99.45
50	5.60	5.20	7.80	8.55	11.70	16.70	78.05	90.70	98.05	99.05	99.75	99.90
70	5.25	6.15	7.35	8.30	9.85	11.30	89.80	97.60	99.60	99.75	99.95	100.00
100	4.95	5.70	5.85	6.80	7.50	9.90	97.05	99.90	100.00	100.00	100.00	100.00
200	5.85	5.05	5.15	5.40	5.60	7.60	99.95	100.00	100.00	100.00	100.00	100.00

Notes: y_{it} is generated as described in the notes to Table S1, except that $\varepsilon_{iyt} = \rho_{iy\varepsilon}\varepsilon_{iyt-1} + (1-\rho_{iy\varepsilon}^2)^{1/2}\eta_{iyt}, \eta_{iyt} \sim iidN(0, \sigma_i^2)$, $\varepsilon_{iy,-50} = 0$, $\sigma_i^2 \sim iidU[0.5, 1.5]$, $\rho_{iy\varepsilon} \sim iidU[0.2, 0.4]$. See also the notes to Table S1 for the specification of the rest of the parameters and the test statistics.

Table S4: Size and Power of Alternative Panel Unit Root Tests with Two Factors, Factors are Serially Uncorrelated but Idiosyncratic Errors are Positively Serially Correlated, $m^0 = 2$ Known, With an Intercept and a Linear Trend

	Size: $\rho_i = \rho = 1$						**Power**: $\rho_i \sim iidU[0.90, 0.99]$					
(T, N)	20	30	50	70	100	200	20	30	50	70	100	200
$CIPS(\hat{p}, k = 1)$												
20	5.05	3.90	3.85	4.15	3.50	2.75	6.15	4.55	5.50	5.65	4.10	5.05
30	4.15	5.05	4.35	4.95	3.75	3.45	6.00	6.30	6.40	7.95	9.35	9.25
50	5.80	4.50	4.65	5.05	4.95	4.30	8.95	9.55	11.05	12.15	16.30	19.05
70	5.10	4.65	4.45	4.65	3.95	4.00	13.85	15.60	18.95	23.30	24.55	28.70
100	5.25	5.50	5.30	4.70	4.05	4.50	21.70	27.60	33.65	36.60	43.00	47.15
200	5.60	4.85	5.75	4.85	5.75	3.35	79.35	89.95	96.55	98.40	99.35	99.90
$CSB(\hat{p}, k = 1)$												
20	6.50	5.70	6.05	4.65	5.50	5.30	8.55	8.65	10.80	11.15	13.00	18.20
30	5.65	5.05	5.20	5.25	5.00	4.85	8.70	9.70	12.90	15.35	18.30	22.40
50	4.80	5.25	4.25	4.45	4.20	4.65	12.40	15.70	19.35	24.45	28.15	36.10
70	4.90	3.65	4.70	4.20	4.35	3.80	21.65	28.75	39.10	47.85	58.50	73.10
100	4.15	4.30	5.10	4.50	4.45	4.35	40.20	55.15	72.60	82.25	90.60	98.10
200	4.45	3.95	4.10	4.70	5.10	4.80	90.40	96.90	99.50	99.95	100.00	100.00
$P_{\hat{e}}(\hat{p})$												
20	20.60	24.95	30.60	35.50	41.70	54.60	21.60	24.95	32.00	35.55	42.00	55.35
30	15.35	16.20	18.35	22.85	23.65	32.60	17.95	20.15	24.00	30.10	32.50	48.65
50	12.90	13.70	14.80	14.65	17.50	23.95	23.00	27.40	38.85	45.45	54.80	78.95
70	9.70	10.75	11.85	13.10	15.40	16.95	33.20	41.60	56.75	69.80	81.65	96.60
100	7.95	8.65	8.75	8.05	8.85	8.50	48.00	62.65	82.15	89.95	95.85	99.85
200	8.45	8.30	7.95	8.65	9.15	8.60	96.25	99.55	100.00	100.00	100.00	100.00
$PMSB$												
20	0.75	1.45	4.05	5.90	9.30	15.05	0.70	1.10	3.00	4.80	6.35	11.75
30	0.70	1.50	2.35	4.05	3.50	6.70	1.25	1.75	2.75	4.85	5.65	9.55
50	1.55	1.50	1.75	1.60	2.10	3.60	6.30	9.75	16.70	22.50	32.30	58.90
70	1.75	1.65	2.15	1.75	2.55	2.45	17.25	26.45	46.40	61.05	75.75	95.05
100	2.15	2.95	2.60	2.80	2.30	2.55	43.50	63.25	85.20	94.45	96.85	99.85
200	3.05	2.65	2.70	3.25	3.60	3.10	93.05	98.40	99.90	99.95	100.00	100.00
P_b												
20	8.70	11.25	16.10	17.80	24.05	31.65	9.10	9.70	14.90	15.35	21.70	29.90
30	6.45	6.95	9.40	11.40	12.75	18.15	9.15	10.10	15.15	18.25	21.85	34.55
50	6.90	5.15	6.55	6.00	7.00	10.60	20.90	26.40	38.20	45.45	56.40	80.50
70	6.80	5.60	6.55	5.45	6.20	7.35	36.60	47.70	63.80	76.60	88.00	97.55
100	6.65	6.00	6.10	6.15	5.60	5.50	63.85	77.35	92.35	97.35	98.35	99.95
200	7.35	6.15	5.65	5.60	5.40	5.55	97.20	99.40	99.95	99.95	100.00	100.00
PP												
20	1.40	2.00	3.80	4.35	6.00	12.20	2.60	2.75	6.40	6.65	10.50	21.50
30	1.65	2.45	3.05	3.60	5.70	9.05	3.90	5.75	8.70	13.80	20.20	35.70
50	2.65	3.00	3.75	3.15	4.10	6.75	13.45	19.60	32.40	40.40	55.10	77.95
70	2.85	2.60	2.85	3.60	4.10	5.55	25.30	38.85	59.10	72.25	85.70	96.20
100	3.25	3.45	3.70	3.10	3.80	3.95	52.40	70.65	90.50	96.35	97.65	99.75
200	3.40	3.20	3.35	4.10	3.75	4.40	95.45	99.05	99.90	99.90	100.00	99.95
$\widetilde{CPO}$												
20	22.00	34.10	55.45	66.65	78.60	92.90	25.95	42.85	64.55	76.55	85.90	96.15
30	11.90	16.80	29.25	42.40	54.50	80.25	20.25	32.80	50.20	65.50	80.30	93.80
50	7.40	9.30	16.00	19.85	26.40	45.60	26.75	39.55	62.05	75.00	87.50	97.05
70	5.55	6.55	9.30	12.95	17.55	31.15	37.20	55.90	77.80	87.90	94.90	98.65
100	4.55	5.35	6.65	9.10	11.55	18.20	59.80	79.00	94.45	97.95	99.10	99.90
200	3.50	3.85	4.65	5.90	6.25	8.85	95.50	99.15	99.95	99.90	100.00	99.95

Notes: y_{it} is generated as described in Table S3, but $d_{iyt} = \mu_{iy} + (1 - \rho_i)\delta_i t$ with $\mu_{iy} \sim iidU[0.0, 0.02]$ and $\delta_i \sim iidU[0.0, 0.02]$, and $d_{it} = \delta_{ix}$ with $\delta_{ix} \sim iidU[0.0, 0.02]$. See also the notes to Tables S1 and S3 for the specification of the rest of the parameters and the test statistics.

Table S5: Size and Power of Alternative Panel Unit Root Tests with Two Factors, Factors are Serially Uncorrelated but Idiosyncratic Errors are Negatively Serially Correlated, $m^0 = 2$ Known, With an Intercept Only

	Size: $\rho_i = \rho = 1$						**Power**: $\rho_i \sim iidU[0.90, 0.99]$					
(T, N)	20	30	50	70	100	200	20	30	50	70	100	200
$CIPS\ (\hat{p}, k = 1)$												
20	6.65	7.20	6.35	6.60	7.40	6.85	8.45	11.55	11.60	13.95	13.20	15.60
30	6.60	7.10	6.45	6.70	7.40	7.45	12.10	14.90	17.05	18.00	19.15	22.95
50	5.10	6.20	6.40	6.90	5.70	6.05	17.60	21.90	26.95	28.30	31.90	38.50
70	5.95	5.50	5.35	6.00	6.30	5.65	28.80	33.45	41.55	47.65	50.95	56.85
100	6.30	7.40	7.00	5.40	6.15	5.95	46.40	56.25	68.20	70.10	79.35	83.35
200	6.20	5.15	6.15	5.65	5.45	5.20	97.75	99.50	99.90	99.90	100.00	100.00
$CSB(\hat{p}, k = 1)$												
20	6.60	5.40	5.25	4.65	6.10	5.65	14.45	16.40	19.30	24.60	25.80	31.30
30	6.05	5.85	5.25	5.55	5.80	4.30	21.55	25.65	32.15	36.90	41.45	47.25
50	6.55	5.95	6.35	6.50	5.40	5.80	38.70	46.65	60.95	67.50	75.50	85.15
70	6.35	6.25	6.75	6.25	6.45	5.15	58.65	71.85	84.45	90.25	94.35	97.80
100	5.60	6.00	6.20	6.40	5.60	6.50	74.25	85.80	94.30	95.60	98.30	99.70
200	7.05	5.90	6.75	6.10	7.00	6.80	91.40	94.10	97.90	98.90	99.50	100.00
$P_{\hat{e}}(\hat{p})$												
20	9.15	10.35	11.15	12.45	12.40	15.35	23.20	27.05	35.15	41.70	50.70	71.05
30	9.65	9.40	8.80	8.35	8.65	10.35	30.20	37.05	49.35	60.10	71.10	87.35
50	9.25	8.45	8.45	9.90	10.45	11.45	55.75	66.15	82.40	90.15	94.45	97.60
70	6.90	8.05	8.20	8.05	7.95	8.85	71.50	84.30	93.15	96.20	98.60	99.60
100	7.25	6.90	7.40	6.55	6.95	6.05	86.65	95.45	99.15	99.35	99.55	99.95
200	7.35	6.95	6.65	6.15	6.00	6.50	99.45	99.95	99.95	100.00	100.00	100.00
$PMSB$												
20	0.75	0.90	1.10	1.05	1.00	1.65	5.00	7.90	15.05	20.75	27.60	51.10
30	1.30	1.40	1.55	1.50	1.20	0.95	12.25	22.30	36.30	49.45	62.45	81.35
50	1.70	1.70	2.05	2.10	1.95	1.35	38.10	56.65	77.50	86.40	91.05	94.55
70	2.20	2.45	2.90	2.20	2.20	1.60	62.95	78.40	90.90	94.30	97.25	98.55
100	2.20	3.10	3.50	2.60	3.05	2.35	81.30	92.60	96.65	98.00	98.70	99.55
200	3.30	3.15	3.25	3.70	3.70	2.55	98.05	99.45	99.95	99.90	100.00	100.00
P_b												
20	11.55	11.70	14.80	17.00	17.05	28.95	37.25	47.15	62.75	72.95	81.30	92.20
30	11.05	11.05	11.15	12.65	13.05	17.60	52.90	64.95	80.45	86.25	91.50	95.85
50	9.70	9.15	9.45	8.85	10.40	12.55	76.35	85.00	92.70	95.40	97.15	97.30
70	9.20	9.10	9.10	9.15	9.65	8.30	87.15	92.45	97.00	96.95	98.60	99.45
100	8.65	8.30	8.80	7.70	7.30	9.35	93.70	97.95	99.00	99.30	99.40	99.85
200	8.75	7.00	7.15	7.10	6.45	6.40	99.50	99.65	99.95	100.00	100.00	100.00
t_b^*												
20	11.75	12.25	15.40	18.15	21.20	26.35	79.90	88.15	94.00	95.05	96.55	97.70
30	12.10	12.55	13.65	15.20	17.05	22.30	90.55	93.15	96.55	97.65	98.60	98.95
50	9.10	9.20	10.70	10.80	14.05	17.35	95.15	98.15	98.40	99.35	99.50	99.95
70	9.10	9.60	9.00	10.30	10.95	15.00	98.45	99.25	99.50	99.50	99.85	100.00
100	9.40	8.65	9.20	8.75	10.45	10.95	99.15	99.90	99.90	99.95	100.00	100.00
200	8.00	6.75	7.65	7.90	7.80	9.35	100.00	100.00	100.00	100.00	100.00	100.00
$\widetilde{CPO}$												
20	7.25	10.15	13.50	18.40	22.60	36.15	30.20	41.15	56.95	66.50	72.55	84.10
30	7.95	9.65	13.10	15.45	19.45	27.40	44.25	58.50	73.85	79.95	87.20	92.00
50	6.70	6.85	9.40	10.90	14.65	20.85	67.75	80.70	90.20	93.35	95.10	96.50
70	6.75	8.05	9.40	10.75	12.95	19.00	82.70	90.55	95.25	96.40	97.85	99.25
100	6.50	7.55	9.65	10.95	12.75	19.65	91.35	96.80	98.60	98.90	99.20	99.70
200	6.40	6.45	8.90	9.30	10.85	18.05	99.05	99.45	99.95	100.00	100.00	100.00

Notes: y_{it} is generated as described in the notes to Table S1, except that $\varepsilon_{iyt} = \rho_{iy\varepsilon}\varepsilon_{iyt-1} + (1 - \rho_{iy\varepsilon}^2)^{1/2}\eta_{iyt}, \eta_{iyt} \sim iidN(0, \sigma_i^2)$, $\varepsilon_{iy,-50} = 0$, $\sigma_i^2 \sim iidU[0.5, 1.5]$, $\rho_{iy\varepsilon} \sim iidU[-0.4, -0.2]$. See also the notes to Table S1 for the specification of the rest of the parameters and the test statistics.

Table S6: Size and Power of Alternative Panel Unit Root Tests with Two Factors,
Factors are Serially Uncorrelated but Idiosyncratic Errors are Negatively Serially Correlated, $m^0 = 2$ Known,
With an Intercept and a Linear Trend

	Size: $\rho_i = \rho = 1$						**Power**: $\rho_i \sim iidU[0.90, 0.99]$					
(T, N)	20	30	50	70	100	200	20	30	50	70	100	200
$CIPS(\hat{p}, k=1)$												
20	6.65	7.15	7.95	7.60	7.35	6.15	6.75	7.05	9.70	9.50	8.90	8.00
30	6.00	6.40	6.70	7.35	7.40	7.10	8.00	8.40	9.65	11.05	12.70	13.00
50	6.85	5.85	7.25	6.75	6.85	5.90	11.30	11.65	13.50	15.80	17.85	19.80
70	5.90	6.15	6.40	6.60	5.10	6.00	15.30	17.85	23.55	25.75	28.55	30.80
100	6.15	6.70	5.90	6.30	5.85	5.75	24.75	29.35	36.00	42.05	45.05	55.30
200	7.20	5.80	6.85	4.50	6.00	4.65	84.20	92.35	98.20	98.95	99.80	99.95
$CSB(\hat{p}, k=1)$												
20	6.85	6.35	6.05	5.50	6.20	5.25	8.85	8.70	11.95	11.90	15.25	18.55
30	6.75	5.65	6.90	6.65	5.95	6.60	11.35	14.45	17.45	18.15	22.55	28.70
50	5.75	7.05	6.00	6.75	5.05	6.00	16.25	21.15	28.70	30.55	35.90	44.20
70	6.75	6.40	6.60	5.95	7.45	5.90	26.60	34.95	48.35	55.45	66.35	79.10
100	6.00	6.70	6.30	8.10	7.45	5.95	41.65	54.85	74.30	80.10	89.10	96.90
200	6.75	6.80	6.95	6.95	6.95	8.15	82.50	90.55	95.80	97.95	99.15	99.85
$P_{\hat{e}}(\hat{p})$												
20	12.10	12.95	13.75	16.05	16.75	18.90	13.65	13.75	15.90	17.70	19.45	24.70
30	9.10	8.85	8.90	9.20	10.90	11.05	12.10	13.25	13.85	16.10	17.90	22.45
50	8.35	9.45	9.85	10.15	9.70	11.35	18.40	20.05	26.50	30.55	39.15	56.30
70	8.70	8.15	8.30	8.60	9.00	9.35	25.20	32.25	40.05	50.65	58.75	80.50
100	7.15	6.95	5.25	5.60	6.10	5.55	39.15	48.90	64.35	72.50	83.35	92.30
200	6.95	7.90	6.15	6.65	6.65	6.75	89.25	95.20	99.15	99.40	99.90	100.00
$PMSB$												
20	0.20	0.35	0.25	0.40	0.55	0.30	0.65	0.95	0.60	0.95	0.70	0.70
30	1.25	1.15	1.15	0.75	0.60	0.20	2.05	2.30	2.45	2.90	3.60	5.90
50	1.90	1.50	1.70	1.55	1.65	0.80	6.65	11.00	17.45	22.45	30.60	51.55
70	2.20	2.20	1.90	2.15	2.60	1.75	16.75	24.10	40.75	49.45	59.90	79.40
100	3.00	3.20	3.00	3.25	2.80	2.50	37.95	51.80	70.65	78.85	85.25	91.60
200	3.10	3.85	3.60	3.35	4.10	3.30	84.00	91.10	96.00	96.85	97.70	98.90
P_b												
20	11.95	12.10	12.05	13.95	16.00	23.85	14.40	15.30	17.80	20.45	25.50	39.35
30	11.15	11.10	11.40	12.20	14.85	18.75	16.25	20.90	22.30	28.15	34.85	52.20
50	8.85	10.35	10.00	11.30	11.85	13.35	27.25	33.85	45.10	50.65	62.70	78.05
70	9.80	8.85	9.25	9.05	10.00	11.80	39.15	47.55	61.90	69.90	77.60	88.40
100	9.95	9.65	8.70	8.75	8.90	10.40	60.00	69.25	81.85	87.30	90.45	93.95
200	8.55	8.80	7.40	6.80	7.55	7.25	90.15	94.70	97.55	97.30	98.35	99.00
PP												
20	0.25	0.15	0.50	0.15	0.40	0.40	0.30	0.30	0.40	0.30	0.70	0.85
30	0.55	0.75	0.45	0.60	1.20	1.20	1.05	1.45	1.70	2.25	3.65	4.70
50	1.20	1.75	1.35	1.70	1.40	2.15	6.10	8.30	13.70	18.60	22.75	35.20
70	1.65	2.00	1.45	1.80	2.85	3.15	15.75	22.00	36.40	43.80	50.60	68.85
100	2.85	3.65	3.15	3.45	4.00	5.90	40.80	50.65	67.40	75.70	79.85	88.30
200	3.30	3.90	4.25	4.65	6.25	9.45	84.60	91.90	95.75	97.10	97.55	98.80
$\widetilde{CPO}$												
20	6.35	9.15	13.05	19.15	24.00	37.90	7.65	12.15	17.15	24.75	31.00	47.50
30	4.20	6.35	9.15	11.75	16.25	24.40	7.65	12.65	17.70	24.00	32.70	46.10
50	3.80	5.60	7.30	9.40	10.80	17.70	14.80	21.90	32.95	41.70	51.15	66.60
70	3.35	4.50	5.50	7.40	9.80	16.40	23.35	33.35	52.90	61.00	67.70	81.15
100	4.20	6.25	5.90	7.35	9.10	15.30	45.40	58.30	75.75	81.85	85.00	91.35
200	3.50	4.60	5.40	6.40	8.50	14.15	84.90	92.60	96.20	97.40	97.85	99.00

Notes: y_{it} is generated as described in Table S5, but $d_{iyt} = \mu_{iy} + (1 - \rho_i)\delta_i t$ with $\mu_{iy} \sim iidU[0.0, 0.02]$ and $\delta_i \sim iidU[0.0, 0.02]$, and $d_{ix} = \delta_{ix}$ with $\delta_{ix} \sim iidU[0.0, 0.02]$. See also the notes to Tables S1 and S5 for the specification of the rest of the parameters and the test statistics.

References

Andrews, D. and C. Monahan (1992). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 60, 953-966.

Bai, J. and S. Ng (2004). A panic on unit root tests and cointegration. *Econometrica* 72, 1127-1177.

Bai, J. and S. Ng (2010). Panel unit root tests with cross section dependence: A further investigation. *Econometric Theory*, 26,1088-1114

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance* 20, 249-272.

Horowitz, J.L. and N.E. Savin (2000). Empirically relevant critical values for hypothesis tests: A bootstrap approach. *Journal of Econometrics* 95, 375-389.

Moon, H.R. and B. Perron (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics* 122, 81-126.

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

Moon, H.R., Perron, B. and P.C.B. Phillips (2011). Point optimal panel unit root tests with serially correlated errors. Unpublished Manuscript.

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

Pesaran, M.H., Smith, L.V. and T. Yagamata (2010). A panel unit root test in the presence of a multifactor error structure. *Mimeo.*

Ploberger, W. and P.C.B. Phillips (2002). Optimal testing for unit roots in panel data. *Mimeo.*