Supplementary Material for
"Identifying Common Trend Determinants in Panel Data"

Yoonseok Lee, Peter C. B. Phillips, Suyong Song, and Donggyu Sul

June 2025

Section S.1 contains proofs of the technical lemmas and the corollary; Section S.2 derives limiting distributions when linear trend exists; Section S.3 provides simulation results; Section S.4 includes additional results and tables of the empirical studies on crime rates.

S.1 Proof of Lemmas

Proof of Lemma B1 Recall $e_t = \xi_t + \overline{x}_t^*$, where $\overline{x}_t^* = n^{-1}\sum_{i=1}^n(\mu_i t^{-\kappa_1} + \epsilon_{it} + \varepsilon_{it}t^{-\kappa_2}) = \overline{\mu}t^{-\kappa_1} + \overline{\epsilon}_t + \overline{\varepsilon}_t t^{-\kappa_2}$. We decompose $\widehat{\delta} - \delta$ as

$$\widehat{\delta} - \delta = \left(\sum_{t=1}^T \widetilde{\theta}_t \widetilde{\theta}_t'\right)^{-1} \sum_{t=1}^T \widetilde{\theta}_t e_t = A_{5,T}^{-1}\left(A_{1,T} + A_{2,nT} + A_{3,nT} + A_{4,nT}\right),$$

where

$$A_{1,T} = \sum_{t=1}^T \widetilde{\theta}_t \xi_t, \;\; A_{2,nT} = \overline{\mu}\sum_{t=1}^T \widetilde{\theta}_t t^{-\kappa_1}, \;\; A_{3,nT} = \sum_{t=1}^T \widetilde{\theta}_t \overline{\epsilon}_t, \;\; A_{4,nT} = \sum_{t=1}^T \widetilde{\theta}_t \overline{\varepsilon}_t t^{-\kappa_2}, \;\; A_{5,T} = \sum_{t=1}^T \widetilde{\theta}_t \widetilde{\theta}_t'.$$

When $\xi_t \sim I(0)$, under Assumptions 1 and 2, combining the functional central limit theorem with the continuous mapping theorem (e.g., Park and Phillips (1988)), we have

$$\frac{1}{T}A_{1,T} \Rightarrow \int_0^1 \widetilde{B}_\theta(r)\, dB_\xi(r)$$

as $T \to \infty$. Similarly, by Lemma 3.1 of Chang, Park, and Phillips (2001),

$$\frac{\sqrt{n}}{T^{(3/2)-\kappa_1}}A_{2,nT} = \sqrt{n}\overline{\mu} \cdot \frac{1}{T^{(3/2)-\kappa_1}}\sum_{t=1}^T \widetilde{\theta}_t t^{-\kappa_1} = O_p(1)$$

as $n,T \to \infty$, where $\overline{\mu} = O_p(n^{-1/2})$ and $\int_0^1 r^{-\kappa_1} dr = 1/(1-\kappa_1) < \infty$ as we assume $\kappa_1 \in (0,1/2)$. By Theorem 16 of Phillips and Moon (1999),

$$\frac{\sqrt{n}}{T} A_{3,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\theta_t}{\sqrt{T}} \epsilon_{it} - \frac{1}{T} \sum_{t=1}^{T} \frac{\theta_t}{\sqrt{T}} \cdot \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \epsilon_{it} \right\} = O_p\left(1\right)$$

and similarly

$$\frac{\sqrt{n}}{T^{1-\kappa_2}} A_{4,nT} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\theta_t}{\sqrt{T}} \varepsilon_{it} \left(\frac{t}{T}\right)^{-\kappa_2} - \frac{1}{T} \sum_{t=1}^{T} \frac{\theta_t}{\sqrt{T}} \cdot \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_{it} \left(\frac{t}{T}\right)^{-\kappa_2} \right\} = O_p\left(1\right)$$

for $\kappa_2 \in (0,1/2)$. Noting that

$$\frac{1}{T^2} A_{5,T} \Rightarrow \int_0^1 \widetilde{B}_\theta\left(r\right) \widetilde{B}_\theta\left(r\right)' dr,$$

we thus have

$$T(\widehat{\delta} - \delta) \Rightarrow \left(\int_0^1 \widetilde{B}_\theta\left(r\right) \widetilde{B}_\theta\left(r\right)' dr \right)^{-1} \int_0^1 \widetilde{B}_\theta\left(r\right) dB_\xi\left(r\right)$$

as $n,T \to \infty$ because it is assumed that $n/T \to \infty$. When $\xi_t \sim I(1)$, under Assumptions 1 and 3, $A_{1,T}$ dominates all other terms in the numerator, where

$$\frac{1}{T^2} A_{1,T} \Rightarrow \int_0^1 \widetilde{B}_\theta\left(r\right) B_\xi\left(r\right) dr$$

and hence the desired result follows. □

Proof of Lemma B2 First note that

$$S_{n,t} = \frac{1}{n} \sum_{i=1}^{n} \left(-(\widehat{\delta} - \delta)' \theta_t + \xi_t + x_{it} \right)^2, \tag{S.1}$$

where $x_{it} = \alpha_i + x_{it}^*$. When $\xi_t \sim I(1)$, $\widehat{\delta} - \delta = O_p(1)$ from Lemma B1 and the term $-(\widehat{\delta} - \delta)'\theta_t + \xi_t$ dominates x_{it} in (S.1). It follows that

$$\frac{1}{T^3} Z_{nT}\left(r\right) = \frac{1}{T^3} \sum_{t=1}^{[Tr]} \widetilde{t} \left\{ \left(\xi_t - (\widehat{\delta} - \delta)' \theta_t \right)^2 - \frac{1}{T} \sum_{s=1}^{T} \left(\xi_s - (\widehat{\delta} - \delta)' \theta_s \right)^2 \right\} + o_p\left(1\right).$$

Note that

$$\begin{aligned}
\frac{1}{T^3}\sum_{t=1}^{[Tr]}\widetilde{t}\left(\xi_t-(\widehat{\delta}-\delta)'\theta_t\right)^2 &\Rightarrow \int_0^r \widetilde{s}\left(B_\xi(s)-D_\delta' B_\theta(s)\right)^2 ds \\
\frac{1}{T^3}\sum_{t=1}^{[Tr]}\widetilde{t}\cdot\frac{1}{T}\sum_{s=1}^{T}\left(\xi_s-(\widehat{\delta}-\delta)'\theta_s\right)^2 &\Rightarrow \int_0^r \widetilde{s}ds\int_0^1\left(B_\xi(s)-D_\delta' B_\theta(s)\right)^2 ds,
\end{aligned}$$

where

$$D_\delta=\left(\int_0^1 \widetilde{B}_\theta(\nu)\widetilde{B}_\theta(\nu)'d\nu\right)^{-1}\int_0^1 \widetilde{B}_\theta(\nu)B_\xi(\nu)d\nu$$

from Lemma B1. The desired result follows since

$$\begin{aligned}
& B_\xi(s)-D_\delta' B_\theta(s) \\
=\ & B_\xi(s)-\int_0^1 B_\xi(\nu)\widetilde{B}_\theta(\nu)'d\nu\left(\int_0^1 \widetilde{B}_\theta(\nu)\widetilde{B}_\theta(\nu)'d\nu\right)^{-1}B_\theta(s) \\
=\ & \omega_\xi W_1(s)-\omega_\xi\int_0^1 W_1(\nu)\widetilde{W}_m(\nu)'d\nu\Omega_\theta^{1/2}\left(\Omega_\theta^{1/2}\int_0^1 \widetilde{W}_m(\nu)\widetilde{W}_m(\nu)'d\nu\Omega_\theta^{1/2}\right)^{-1}\Omega_\theta^{1/2}W_m(s) \\
=\ & \omega_\xi V(s)^{1/2}
\end{aligned}$$

from (B.1). □

Proof of Lemma B3 From the supplementary proof below, it can be verified that the leading terms of $Z_{nT}(r)$ are given as

$$\begin{aligned}
Z_{nT}(r) &= \sum_{t=1}^{[Tr]}\widetilde{t}\left(\xi_t^2-\frac{1}{T}\sum_{s=1}^{T}\xi_s^2\right)+\frac{1}{n}\sum_{i=1}^{n}\mu_i^2\sum_{t=1}^{[Tr]}\widetilde{t}\left\{t^{-2\kappa_1}-\frac{1}{T}\sum_{s=1}^{T}s^{-2\kappa_1}\right\} \\
&\quad+\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{\varepsilon_{it}^2 t^{-2\kappa_2}-\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is}^2 s^{-2\kappa_2}\right\}+o_p\left(\max\{T^{3/2},T^{2-2\kappa^*}\}\right) \\
&= C_{1,T}(r)+C_{2,nT}(r)+C_{3,nT}(r)+o_p\left(\max\{T^{3/2},T^{2-2\kappa^*}\}\right), \qquad \text{(S.2)}
\end{aligned}$$

where $\kappa^*=\min\{\kappa_1,\kappa_2\}$. For $C_{1,T}(r)$, with $\mathbb{E}\xi_t^2=\sigma_\xi^2<\infty$, we have

$$\frac{1}{T^{3/2}}C_{1,T}(r) = \frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{\left(\xi_t^2-\sigma_\xi^2\right)-\frac{1}{T}\sum_{s=1}^{T}\left(\xi_s^2-\sigma_\xi^2\right)\right\}$$

$$
\begin{aligned}
&= \frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}\left(\xi_t^2-\sigma_\xi^2\right)-\frac{1}{T^2}\sum_{t=1}^{[Tr]}\widetilde{t}\cdot\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\left(\xi_s^2-\sigma_\xi^2\right)\\
&\Rightarrow \int_0^r \widetilde{s}dB_{\xi\xi}(s)-\int_0^r \widetilde{s}dsB_{\xi\xi}(1)\\
&= \omega_{\xi\xi}\int_0^r \widetilde{s}d\mathcal{B}(s),
\end{aligned}
\tag{S.3}
$$

where $B_{\xi\xi}(s)=\omega_{\xi\xi}W(s)$ with the standard Brownian motion $W(s)$ and $\mathcal{B}(s)=W(s)-sW(1)$ is the Brownian bridge. For $C_{2,nT}(r)$, note that

$$
\frac{1}{T^{2-2\kappa_1}}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{t^{-2\kappa_1}-\frac{1}{T}\sum_{s=1}^{T}s^{-2\kappa_1}\right\}\rightarrow\int_0^r\left(s-\frac{1}{2}\right)\left(s^{-2\kappa_1}-\frac{1}{1-2\kappa_1}\right)ds=q\left(\kappa_1;r\right)<\infty
$$

for $\kappa_1\in(0,1/2)$, and hence

$$
\frac{1}{T^{2-2\kappa_1}}C_{2,nT}(r)\xrightarrow{p}\sigma_\mu^2 q\left(\kappa_1;r\right),
$$

where $\mathbb{E}\mu_i^2=\sigma_\mu^2<\infty$. For $C_{3,nT}(r)$, as $\mathbb{E}\varepsilon_{it}^2=\sigma_{\varepsilon,i}^2<\infty$ for each i, note that

$$
\begin{aligned}
&\sum_{t=1}^{[Tr]}\widetilde{t}\left\{\varepsilon_{it}^2t^{-2\kappa_2}-\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is}^2s^{-2\kappa_2}\right\}\\
&= \sum_{t=1}^{[Tr]}\widetilde{t}\left\{\left(\varepsilon_{it}^2-\sigma_{\varepsilon,i}^2\right)t^{-2\kappa_2}-\frac{1}{T}\sum_{s=1}^{T}\left(\varepsilon_{is}^2-\sigma_{\varepsilon,i}^2\right)s^{-2\kappa_2}\right\}+\sigma_{\varepsilon,i}^2\sum_{t=1}^{[Tr]}\widetilde{t}\left\{t^{-2\kappa_2}-\frac{1}{T}\sum_{s=1}^{T}s^{-2\kappa_2}\right\}\\
&= O_p\left(T^{(3/2)-2\kappa_2}\right)+O_p\left(T^{2-2\kappa_2}\right)
\end{aligned}
$$

by Lemma 3.1 of Chang, Park, and Phillips (2001). It follows that

$$
\begin{aligned}
\frac{1}{T^{2-2\kappa_2}}C_{3,nT}(r) &= O_p(T^{-1/2})+\frac{1}{n}\sum_{i=1}^{n}\sigma_{\varepsilon,i}^2\cdot\frac{1}{T^{2-2\kappa_2}}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{t^{-2\kappa_2}-\frac{1}{T}\sum_{s=1}^{T}s^{-2\kappa_2}\right\}\\
&= \sigma_\varepsilon^2 q\left(\kappa_2;r\right)+o_p(1),
\end{aligned}
$$

where $\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\sigma_{\varepsilon,i}^2=\sigma_\varepsilon^2<\infty$. Under $n/T\to\infty$, the desired results follow by combining these three terms and verifying the leading terms for each case. □

Proof of (S.2) in Lemma B3 When $\xi_t\sim I(0)$, $\widehat{\delta}-\delta=O_p(T^{-1})$ from Lemma B1 and the term ξ_t+x_{it} dominates $(\widehat{\delta}-\delta)'\theta_t$ in (S.1). Therefore, the leading terms of $Z_{nT}(r)$ are given as

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{(\xi_t + x_{it})^2 - \frac{1}{T}\sum_{s=1}^{T}(\xi_s + x_{is})^2\right\} \\
= & \sum_{t=1}^{[Tr]}\widetilde{t}\left\{\xi_t^2 - \frac{1}{T}\sum_{s=1}^{T}\xi_s^2\right\} + 2\sum_{t=1}^{[Tr]}\widetilde{t}\left\{\xi_t\overline{x}_t - \frac{1}{T}\sum_{s=1}^{T}\xi_s\overline{x}_s\right\} + \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{x_{it}^2 - \frac{1}{T}\sum_{s=1}^{T}x_{is}^2\right\} \\
= & Z_{1,T}(r) + 2Z_{2,nT}(r) + Z_{3,nT}(r),
\end{aligned}$$

where $\overline{x}_t = n^{-1}\sum_{i=1}^{n}x_{it}$. The first term $Z_{1,T}(r)$ is the same as $C_{1,T}(r)$ in (S.3) in the proof of Lemma B3 and hence $Z_{1,T}(r) = O_p(T^{3/2})$.

For the second term $Z_{2,nT}(r)$, since $\overline{x}_t = \overline{\alpha} + \overline{\mu}t^{-\kappa_1} + \overline{\epsilon}_t + \overline{\varepsilon}_t t^{-\kappa_2}$,

$$\begin{aligned}
Z_{2,nT}(r) &= \sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{\xi}_t\overline{\alpha} + \sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t t^{-\kappa_1}})\overline{\mu} + \sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t\overline{\epsilon}_t}) + \sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t\overline{\varepsilon}_t t^{-\kappa_2}}) \\
&= O_p\left(n^{-1/2}T^{3/2} + n^{-1/2}T^{3/2-\kappa_1} + n^{-1/2}T^{3/2} + n^{-1/2}T^{3/2-\kappa_2}\right).
\end{aligned}$$

The first two elements are obtained similarly as $Z_{1,T}(r)$ for

$$\begin{aligned}
\frac{n^{1/2}}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{\xi}_t\overline{\alpha} &= \frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{\xi}_t\cdot\sqrt{n}\overline{\alpha} = O_p(1) \\
\frac{n^{1/2}}{T^{3/2-\kappa_1}}\sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t t^{-\kappa_1}})\overline{\mu} &= \frac{1}{T^{3/2-\kappa_1}}\sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t t^{-\kappa_1}})\cdot\sqrt{n}\overline{\mu} = O_p(1),
\end{aligned}$$

where $\mathbb{E}\alpha_i = \mathbb{E}\mu_i = \mathbb{E}\xi_t = 0$, $\kappa_1 \in (0, 1/2)$. The rest two elements are similarly obtained from Theorem 16 of Phillips and Moon (1999) because

$$\begin{aligned}
\frac{n^{1/2}}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t\overline{\epsilon}_t}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t\epsilon_{it}}) = O_p(1) \\
\frac{n^{1/2}}{T^{3/2-\kappa_2}}\sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t\overline{\varepsilon}_t t^{-\kappa_2}}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{T^{3/2-\kappa_2}}\sum_{t=1}^{[Tr]}\widetilde{t}(\widetilde{\xi_t\varepsilon_{it}t^{-\kappa_2}}) = O_p(1)
\end{aligned}$$

for $\mathbb{E}\xi_t\epsilon_{it} = \mathbb{E}\xi_t\varepsilon_{it} = 0$ and $\kappa_2 \in (0, 1/2)$.

For the third term $Z_{3,nT}(r)$, since $x_{it}^2 = \alpha_i^2 + \mu_i^2 t^{-2\kappa_1} + \epsilon_{it}^2 + \varepsilon_{it}^2 t^{-2\kappa_2} + 2(\alpha_i\mu_i t^{-\kappa_1} + \alpha_i\epsilon_{it} +$

$\alpha_i\varepsilon_{it}t^{-\kappa_2}+\mu_i\epsilon_{it}t^{-\kappa_1}+\mu_i\varepsilon_{it}t^{-\kappa_1-\kappa_2}+\epsilon_{it}\varepsilon_{it}t^{-\kappa_2})$, a similar derivation yields

$$
\begin{aligned}
Z_{3,nT}(r) &= \sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-2\kappa_1})}\cdot\frac{1}{n}\sum_{i=1}^{n}\mu_i^2+\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(\epsilon_{it}^2-\sigma_{\varepsilon,i}^2)}+\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-2\kappa_2}\varepsilon_{it}^2)} \\
&\quad +2\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-\kappa_1})}\cdot\frac{1}{n}\sum_{i=1}^{n}\alpha_i\mu_i+\frac{2}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{\epsilon}_{it}\alpha_i+\frac{2}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-\kappa_2}\varepsilon_{it})}\alpha_i \\
&\quad +\frac{2}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-\kappa_1}\epsilon_{it})}\mu_i+\frac{2}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-\kappa_1-\kappa_2}\varepsilon_{it})}\mu_i+\frac{2}{n}\sum_{i=1}^{n}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{(t^{-\kappa_2}\epsilon_{it}\varepsilon_{it})} \\
&= O_p\left(T^{2-2\kappa_1}+n^{-1/2}T^{3/2}+T^{2-2\kappa_2}\right) \\
&\quad +O_p\left(n^{-1/2}T^{2-\kappa_1}+n^{-1/2}T^{3/2}+n^{-1/2}T^{3/2-\kappa_2}\right) \\
&\quad +O_p\left(n^{-1/2}T^{3/2-\kappa_1}+n^{-1/2}T^{3/2-\kappa_1-\kappa_2}+n^{-1/2}T^{3/2-\kappa_2}\right)
\end{aligned}
$$

as $\kappa_1,\kappa_2\in(0,1/2)$, $\mathbb{E}\alpha_i\mu_i=\mathbb{E}\epsilon_{it}\varepsilon_{it}=0$, $\mathbb{E}\mu_i^2<\infty$, and $\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\sigma_{\varepsilon,i}^2<\infty$ with $\sigma_{\varepsilon,i}^2=\mathbb{E}\varepsilon_{it}^2$. Note that the first and the third elements of $Z_{3,nT}(r)$ are respectively the same as $C_{2,nT}(r)=O_p\left(T^{2-2\kappa_1}\right)$ and $C_{3,nT}(r)=O_p\left(T^{2-2\kappa_2}\right)$ in the proof of Lemma B3.

The desired result follows because $n/T\to\infty$ yields that the dominating terms are $Z_{1,T}(r)=O_p(T^{3/2})$ and the first $O_p\left(T^{2-2\kappa_1}\right)$ and the third $O_p\left(T^{2-2\kappa_2}\right)$ elements of $Z_{3,nT}(r)$ for any $\kappa_1,\kappa_2\in(0,1/2)$. $\square$

Proof of Lemma B4 Since

$$\widehat{\phi}=\frac{Z_{nT}(1)}{\sum_{t=1}^{T}(\widetilde{t})^2}$$

and $T^{-3}\sum_{t=1}^{T}(\widetilde{t})^2\to1/12$, the results readily follows from Lemma B2 for the case $\xi_t\sim I(1)$. When $\xi_t\sim I(0)$, note that

$$\int_0^1\widetilde{s}d\mathcal{B}(s)=\int_0^1\widetilde{s}dW(s)\sim\mathcal{N}\left(0,\frac{1}{12}\right) \tag{S.4}$$

as $\int_0^1\widetilde{s}ds=0$. The results follow from Lemma B3 since $q(\kappa;1)=-\kappa/\{2(1-\kappa)(1-2\kappa)\}$ and $q(1/4;1)=-1/3$. $\square$

Proof of Lemma B5 Recall $\widehat{u}_t = \widetilde{S}_{nt} - \widetilde{\widehat{\phi} t}$. Lemmas B2 and B4-(i) yield that

$$\begin{aligned}\frac{1}{T^3}\sum_{t=1}^{[Tr]}\widetilde{t}\widehat{u}_t &= \frac{1}{T^3}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{S}_{nt} - \widehat{\phi}\cdot\frac{1}{T^3}\sum_{t=1}^{[Tr]}(\widetilde{t})^2 \\ &\Rightarrow \omega_\xi^2\int_0^r \widetilde{s}\widetilde{V}(s)\,ds - 12\omega_\xi^2\int_0^1 \widetilde{\nu}\widetilde{V}(\nu)\,d\nu\cdot\int_0^r(\widetilde{s})^2\,ds \\ &= \omega_\xi^2\int_0^r \widetilde{s}\left\{\widetilde{V}(s) - \widetilde{s}\left(\int_0^1(\widetilde{\nu})^2\,d\nu\right)^{-1}\int_0^1\widetilde{\nu}\widetilde{V}(\nu)\,d\nu\right\}ds \\ &= \omega_\xi^2\int_0^r \widetilde{s}V^\tau(s)\,ds\end{aligned}$$

for $\int_0^1(\widetilde{\nu})^2\,d\nu = 1/12$ and hence

$$\begin{aligned}T^{-6}\Psi_{nT}(b) &= \frac{1}{T^6}\sum_{t=1}^{T}\sum_{s=1}^{T}K\left(\frac{t-s}{Tb}\right)\left(\widehat{u}_t\widetilde{t}\right)\left(\widehat{u}_s\widetilde{s}\right) \\ &\Rightarrow \omega_\xi^4\int_0^1\int_0^1 K\left(\frac{r-s}{b}\right)\widetilde{r}\widetilde{s}V^\tau(r)\,V^\tau(s)\,drds\end{aligned}$$

using similar steps as Sun (2004). □

Proof of Lemma B6 When $\min\{\kappa_1,\kappa_2\} > 1/4$, Lemmas B3-(i) and B4-(ii) yield that

$$\begin{aligned}\frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}\widehat{u}_t &= \frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]}\widetilde{t}\widetilde{S}_{nt} - T^{3/2}\widehat{\phi}\cdot\frac{1}{T^3}\sum_{t=1}^{[Tr]}(\widetilde{t})^2 \\ &\Rightarrow \omega_{\xi\xi}\int_0^r\widetilde{s}d\mathcal{B}(s) - 12\omega_{\xi\xi}\int_0^1\widetilde{\nu}d\mathcal{B}(\nu)\cdot\int_0^r(\widetilde{s})^2\,ds \\ &= \omega_{\xi\xi}\int_0^r\widetilde{s}\left\{d\mathcal{B}(s) - \widetilde{s}ds\left(\int_0^1(\widetilde{\nu})^2\,d\nu\right)^{-1}\int_0^1\widetilde{\nu}d\mathcal{B}(\nu)\right\} \\ &= \omega_{\xi\xi}\int_0^r\widetilde{s}dW^\tau(s)\end{aligned}$$

where $W^\tau(r)$ is the second-level Brownian bridge (e.g., MacNeill (1978)) defined as $W^\tau(r) = W(r) - rW(1) + 6r(1-r)\{(1/2)W(1) - \int_0^1 W(s)ds\}$, which is the linearly $L_2[0,1]$ demeaned and detrended standard Brownian motion. Note that

$$d\mathcal{B}(s) - \widetilde{s}ds\left(\int_0^1(\widetilde{\nu})^2\,d\nu\right)^{-1}\int_0^1\widetilde{\nu}d\mathcal{B}(\nu)$$

$$
\begin{aligned}
&= dW(s) - dsW(1) - \left(s - \frac{1}{2}\right) ds 12 \int_0^1 \left(\nu - \frac{1}{2}\right) dW(\nu) \\
&= dW(s) - dsW(1) - \left(s - \frac{1}{2}\right) ds 12 \left\{ W(1) - \int_0^1 W(\nu) d\nu - \frac{1}{2} W(1) \right\} \\
&= dW(s) - dsW(1) + 6(1-2s) ds \left\{ \frac{1}{2} W(1) - \int_0^1 W(\nu) d\nu \right\} \\
&= dW^{\tau}(s)
\end{aligned}
$$

from (S.4) and using the integration by parts. Hence,

$$
\begin{aligned}
T^{-3} \Psi_{nT}(b) &= \frac{1}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{Tb}\right) \left(\widehat{u}_t \widetilde{t}\right) \left(\widehat{u}_s \widetilde{s}\right) \\
&\Rightarrow \omega_{\xi\xi}^2 \int_0^1 \int_0^1 K\left(\frac{t-s}{b}\right) \widetilde{r} \widetilde{s} dW^{\tau}(r) dW^{\tau}(s)
\end{aligned}
$$

similarly as Kiefer and Vogelsang (2005).

When $\min\{\kappa_1, \kappa_2\} = 1/4$, we consider the case $\kappa_1 = 1/4 < \kappa_2$; other cases can be obtained by the same derivation. From Lemmas B3-(ii) and B4-(ii), we have

$$
\begin{aligned}
\frac{1}{T^{3/2}} \sum_{t=1}^{[Tr]} \widetilde{t} \widehat{u}_t &\Rightarrow \left\{ \omega_{\xi\xi} \int_0^r \widetilde{s} d\mathcal{B}(s) + \sigma_{\mu}^2 q(1/4; r) \right\} - \left\{ 12 \omega_{\xi\xi} \int_0^1 \widetilde{\nu} d\mathcal{B}(\nu) - 4\sigma_{\mu}^2 \right\} \int_0^r (\widetilde{s})^2 ds \\
&= \omega_{\xi\xi} \int_0^r \widetilde{s} dW^{\tau}(s) + \sigma_{\mu}^2 \int_0^r \widetilde{s} \left(\widetilde{(s^{-1/2})} + 4\widetilde{s} \right) ds \\
&= \omega_{\xi\xi} \int_0^r \widetilde{s} \left\{ dW^{\tau}(s) + \frac{\sigma_{\mu}^2}{\omega_{\xi\xi}} \left(s^{-1/2} + 4s - 4 \right) ds \right\},
\end{aligned}
$$

where $q(1/4; r) = \int_0^r \widetilde{s} \widetilde{(s^{-1/2})} ds$ and $\int_0^1 \nu^{-1/2} d\nu = 2$. It hence yields the desired result as above.

Finally, when $\min\{\kappa_1, \kappa_2\} < 1/4$, we also consider the case when $\kappa_1 < \kappa_2$. From Lemmas B3-(iii) and B4-(ii), we have

$$
\begin{aligned}
\frac{1}{T^{2-2\kappa_1}} \sum_{t=1}^{[Tr]} \widetilde{t} \widehat{u}_t &= \frac{1}{T^{2-2\kappa_1}} \sum_{t=1}^{[Tr]} \widetilde{t} \widetilde{S}_{nt} - T^{1+2\kappa_1} \widehat{\phi} \cdot \frac{1}{T^3} \sum_{t=1}^{[Tr]} (\widetilde{t})^2 \\
&\xrightarrow{p} \sigma_{\mu}^2 \int_0^r \widetilde{s} \widetilde{(s^{-2\kappa_1})} ds + \frac{6\kappa_1 \sigma_{\mu}^2}{(1-\kappa_1)(1-2\kappa_1)} \int_0^r (\widetilde{s})^2 ds
\end{aligned}
$$

$$= \sigma_\mu^2 \int_0^r \widetilde{s} \left\{ s^{-2\kappa_1} + \frac{6\kappa_1 s - (1+2\kappa_1)}{(1-\kappa_1)(1-2\kappa_1)} \right\} ds.$$

The desired result follows similarly as above. □

Proof of Corollary 1 Let $Z_{iT}(r) = \sum_{t=1}^{[Tr]} \widetilde{t}\widetilde{\Delta}_{it}$. When $\xi_t \sim I(0)$, similarly as in the proof of Lemma B3, the dominant terms of $Z_{iT}(r)$ can be obtained as

$$\begin{aligned} Z_{iT}(r) &= \sum_{t=1}^{[Tr]} \widetilde{t}\widetilde{(\xi_t + \epsilon_{it})^2} + 2\alpha_i \sum_{t=1}^{[Tr]} \widetilde{t}\widetilde{(\xi_t + \epsilon_{it})} + \mu_i^2 \sum_{t=1}^{[Tr]} \widetilde{t}\widetilde{(t^{-2\kappa_1})} \\ &\quad + 2\alpha_i\mu_i \sum_{t=1}^{[Tr]} \widetilde{t}\widetilde{(t^{-\kappa_1})} + o_p(\min\{T^{3/2}, T^{2-2\kappa_1}, T^{2-\kappa_1}\}) \\ &= Z_{iT,1}(r) + Z_{iT,2}(r) + Z_{iT,3}(r) + Z_{iT,4}(r) + o_p(\min\{T^{3/2}, T^{2-2\kappa_1}, T^{2-\kappa_1}\}) \\ &= O_p\left(T^{3/2}\right) + 2\alpha_i O_p\left(T^{3/2}\right) + \mu_i^2 O_p\left(T^{2-2\kappa_1}\right) + 2\alpha_i\mu_i O_p\left(T^{2-\kappa_1}\right). \end{aligned} \tag{S.5}$$

If $\alpha_i\mu_i \neq 0$ and $|\alpha_i\mu_i| < \infty$ a.s., the dominant term of $Z_{iT}(r)$ becomes $Z_{iT,4}(r)$, from which

$$T^{1+\kappa_1}\widehat{\varphi}_i = \left(\frac{1}{12}\right)^{-1} 2\alpha_i\mu_i \int_0^1 \widetilde{\nu}\widetilde{(\nu^{-\kappa_1})} d\nu + o_p(1) = \frac{-12\kappa_1\alpha_i\mu_i}{(1-\kappa_1)(2-\kappa_1)} + o_p(1)$$

and

$$\begin{aligned} \frac{1}{T^{2-\kappa_1}} \sum_{t=1}^{[Tr]} \widetilde{t}\widehat{u}_{it} &= 2\alpha_i\mu_i \left\{ \int_0^r \widetilde{s}\widetilde{(s^{-\kappa_1})} ds - 12 \int_0^1 \widetilde{\nu}\widetilde{(\nu^{-\kappa_1})} d\nu \int_0^r (\widetilde{\nu})^2 d\nu \right\} + o_p(1) \\ &= \frac{\alpha_i\mu_i f(r, \kappa_1)}{(1-\kappa_1)(2-\kappa_1)} + o_p(1), \end{aligned}$$

where $f(r;\kappa) = \kappa(4r^3 - 6r^2 + 3r) - (2-\kappa)(r^2 - r + r^{1-\kappa}) + 2(1-\kappa)r^{2-\kappa}$. Therefore,

$$\begin{aligned} &\frac{1}{T^{2(2-\kappa_1)}} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{Tb}\right) \left(\widetilde{t}\widehat{u}_{it}\right) \left(\widetilde{s}\widehat{u}_{is}\right) \\ &= \left(\frac{\alpha_i\mu_i}{(1-\kappa_1)(2-\kappa_1)}\right)^2 \int_0^1 \int_0^1 K\left(\frac{r-s}{b}\right) f(r;\kappa_1) f(s;\kappa_1) drds + o_p(1) \end{aligned}$$

and hence

$$
\begin{aligned}
\mathcal{T}_{\varphi_i}(b) &= \frac{T^{1+\kappa_1}\widehat{\varphi}_i}{\left\{\left(\frac{1}{T^3}\sum_{t=1}^{T}\left(\widetilde{t}\right)^2\right)^2 \frac{1}{T^{2(2-\kappa_1)}}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{Tb}\right)\left(\widetilde{t}\widehat{u}_{it}\right)\left(\widetilde{s}\widehat{u}_{is}\right)\right\}^{1/2}} \\
&= \frac{-\kappa_1 \operatorname{sgn}\left(\alpha_i\mu_i\right)}{\left\{\int_0^1\int_0^1 K\left(\frac{r-s}{b}\right) f\left(r;\kappa_1\right) f\left(s;\kappa_1\right) dr ds\right\}^{1/2}} + o_p\left(1\right),
\end{aligned}
\tag{S.6}
$$

where $\operatorname{sgn}(c) = 1\{c > 0\} - 1\{c < 0\}$.

If $\alpha_i\mu_i = 0$ a.s., the dominant terms of $Z_{iT}(r)$ are different depending on the values of α_i and μ_i. First, if $\alpha_i = 0$ and $\mu_i \neq 0$ a.s., $Z_{iT,1}(r)$ and $Z_{iT,3}(r)$ terms in (S.5) are dominant for $\kappa_1, \kappa_2 \in (0, 1/2)$. In this case, conditional on μ_i, we can derive

$$
\begin{cases}
T^{3/2}\widehat{\varphi}_i \Rightarrow \mathcal{N}\left(0, 12\omega_i^2\right) & \text{if } \kappa_1 < 1/4 \\
T^{3/2}\widehat{\varphi}_i \Rightarrow \mathcal{N}\left(-4\mu_i^2, 12\omega_i^2\right) & \text{if } \kappa_1 = 1/4 \\
T^{1+2\kappa_1}\widehat{\varphi}_i \xrightarrow{p} \frac{-6\kappa_1}{(1-\kappa_1)(1-2\kappa_1)}\mu_i^2 & \text{if } \kappa_1 > 1/4
\end{cases}
$$

similarly as Lemma B4, where ω_i^2 is the long-run variance of $(\xi_t + \epsilon_{it})^2$, and the limit of $\sum_{t=1}^{[Tr]} \widetilde{t}\widehat{u}_{it}$ is obtained as in Lemma B6 with replacing $\lambda(r)$ and $\lambda^*(r)$ by $\left(4r + r^{-1/2} - 4\right)\mu_i^2$ and $c(\kappa_1; r)\mu_i^2$, respectively. From these two results, we can derive the limit of $\mathcal{T}_{\varphi_i}(b)$ as in Theorems 1 and 2, which yields a very similar form as that of $\mathcal{T}_{\phi}(b)$: $F_0(b)$, negatively-shifted $F_0(b)$, or a negative degenerating point that only depends on κ_1. Second, if $\alpha_i \neq 0$ and $\mu_i = 0$ a.s., $Z_{iT,1}(r) + Z_{iT,2}(r)$ are dominant for any $\kappa_1, \kappa_2 \in (0, 1/2)$, and we can derive that $\mathcal{T}_{\varphi_i}(b) \Rightarrow F_0(b)$ as $T \to \infty$, whether Δ_{it} is negatively associated with t or unassociated with t. Finally, if $\alpha_i = \mu_i = 0$ a.s., $Z_{iT,1}(r)$ term is dominant for any $\kappa_1, \kappa_2 \in (0, 1/2)$, and we can derive the identical results as the second case.

When $\xi_t \sim I(1)$, the dominant terms of $Z_{iT}(r)$ are the same as those in Lemma B2, and hence the limiting distribution remains the same. $\square$

This corollary shows that, when $\xi_t \sim I(0)$ and $\alpha_i\mu_i = 0$ a.s., $\mathcal{T}_{\varphi_i}(b)$ converges to $F_0(b)$, negatively-shifted $F_0(b)$, or a negative point given in (23). Unlike $\mathcal{T}_{\phi}(b)$, however, it can converge to $F_0(b)$ for any $\kappa_1 \in (0, 1/2)$ if $\mu_i = 0$. Hence, $\mathcal{T}_{\varphi_i}(b)$ cannot fully distinguish the case when Δ_{it} is negatively associated with t from the case when Δ_{it} is unassociated with t. On the other hand, when $\alpha_i\mu_i > 0$ a.s., a direct calculation yields that the limit of $\mathcal{T}_{\varphi_i}(b)$

in (S.6) ranges $[-25.91, -15.03]$ over $0 < \kappa_1 < 1/2$ for $b = 0.1$, $[-19.50, -10.96]$ for $b = 0.2$, $[-16.37, -9.25]$ for $b = 0.3$, $[-14.57, -8.26]$ for $b = 0.4$, which are far below the degenerating point given in (23). Based on these findings and from the fact that we consider the one-sided test, we use the fist-stage test critical value $c_1 = -1.2$ for our empirical analysis, which is near the 10% percentile of $F_1(b)$, whether $\alpha_i\mu_i > 0$ or $\alpha_i\mu_i < 0$ a.s.

S.2 Limiting Distributions when Linear Trends Exist

Nonstationary variables often exhibit linear trends, which can be modeled using a random walk process with a drift term. As we discussed in Remark 3, we suppose the nonstationary common trend τ_t satisfies

$$\tau_t = \tau_{t-1} + c_\tau + \varepsilon_{\tau,t} = c_\tau t + \zeta_{\tau,t} \tag{S.7}$$

for some $c_\tau \neq 0$, where $\zeta_{\tau,t} = \sum_{s=1}^{t} \varepsilon_{\tau,t}$ with $\varepsilon_{\tau,t} \sim I(0)$ and $\tau_0 = 0$. θ_t is generated as either of the followings:

$$\begin{cases} \text{(random walk)} & \theta_t = \theta_{t-1} + \varepsilon_{\theta,t} = \zeta_{\theta,t} \\ \text{(random walk with drift)} & \theta_t = \theta_{t-1} + c_\theta + \varepsilon_{\theta,t} = c_\theta t + \zeta_{\theta,t} \quad \text{for some } c_\theta \neq 0, \end{cases}$$

where $\zeta_{\theta,t} = \sum_{s=1}^{t} \varepsilon_{\theta,t}$ with $\varepsilon_{\theta,t} \sim I(0)$ and $\theta_0 = 0$. For simplicity, we suppose $\theta_t \in \mathbb{R}^1$. We can consider the following feasible cases.

Case 1: θ_t is random walk with drift and ξ_t is stationary without trend Suppose $\theta_t = c_\theta t + \zeta_{\theta,t}$ and $\xi_t = \tau_t - \delta\theta_t = (c_\tau - \delta c_\theta)t + (\zeta_{\tau,t} - \delta\zeta_{\theta,t})$, where $(c_\tau - \delta c_\theta) = 0$ and $(\zeta_{\tau,t} - \delta\zeta_{\theta,t}) \sim I(0)$. This is the case when a proper θ_t was chosen so that it shares both the deterministic and the stochastic trends with τ_t in (S.7). In other words, the linear combination $\tau_t - \delta\theta_t$ with δ eliminates both the linear trend and the stochastic trend, and hence the cointegration error ξ_t is a mean-zero stationary process without a linear trend. In this case, similarly as in the proof of Lemma B1, we can show that

$$T^{3/2}(\widehat{\delta} - \delta) = \frac{\sum_{t=1}^{T} \widetilde{\theta}_t e_t}{\sum_{t=1}^{T} (\widetilde{\theta}_t)^2} = \frac{T^{-3/2} \sum_{t=1}^{T} (c_\theta \widetilde{t} + \widetilde{\zeta}_{\theta,t})\xi_t + o_p(1)}{T^{-3} \sum_{t=1}^{T} (c_\theta \widetilde{t} + \widetilde{\zeta}_{\theta,t})^2} \Rightarrow \frac{c_\theta \omega_\xi \int_0^1 \widetilde{r} dW_1(r)}{c_\theta^2 \int_0^1 (\widetilde{r})^2 dr}.$$

For S_{nt} in (S.1), since $\widehat{\delta}-\delta=O_p(T^{-3/2})$, $(\widehat{\delta}-\delta)\theta_t$ term is still dominated by the other terms ξ_t+x_{it} as in the original case without linear trends. Therefore, all the lemmas for the $\xi\sim I(0)$ do not change and the limiting distribution of $\mathcal{T}_\phi(b)$ remains the same as in Theorems 1 and 2. $\square$

Case 2: θ_t is random walk with drift and ξ_t is random walk Suppose $\theta_t=c_\theta t+\zeta_{\theta,t}$ and $\xi_t=\tau_t-\delta\theta_t=(c_\tau-\delta c_\theta)t+(\zeta_{\tau,t}-\delta\zeta_{\theta,t})$, where $(c_\tau-\delta c_\theta)=0$ but $(\zeta_{\tau,t}-\delta\zeta_{\theta,t})\sim I(1)$. This is the case when θ_t only shares the linear trend of τ_t in (S.7), but the detrended processes do not have a cointegrating relation. In other words, the linear combination $\tau_t-\delta\theta_t$ with δ eliminates the linear trend but not the stochastic trend. In this case, similarly as in the proof of Lemma B1, we can show that

$$T^{1/2}(\widehat{\delta}-\delta)=\frac{T^{-5/2}\sum_{t=1}^T(c_\theta\widetilde{t}+\widetilde{\zeta}_{\theta,t})\xi_t+o_p(1)}{T^{-3}\sum_{t=1}^T(c_\theta\widetilde{t}+\widetilde{\zeta}_{\theta,t})^2}\Rightarrow\frac{c_\theta\omega_\xi\int_0^1\widetilde{r}W_1(r)dr}{c_\theta^2\int_0^1(\widetilde{r})^2dr}=D_{\delta,A}$$

and thus, for S_{nt} in (S.1), the term $\xi_t-(\widehat{\delta}-\delta)\theta_t$ still dominates x_{it}. It follows that

$$\begin{aligned}
T^{-3}Z_{nT}(r) &= \frac{1}{T^3}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{\left(\xi_t-(\widehat{\delta}-\delta)(c_\theta t+\zeta_{\theta,t})\right)^2-\frac{1}{T}\sum_{s=1}^{T}\left(\xi_s-(\widehat{\delta}-\delta)(c_\theta s+\zeta_{\theta,s})\right)^2\right\}+o_p(1)\\
&\Rightarrow \int_0^r\widetilde{s}\left(B_\xi(s)-D_{\delta,A}c_\theta s\right)^2ds-\int_0^r\widetilde{s}ds\int_0^1\left(B_\xi(s)-D_{\delta,A}c_\theta s\right)^2ds\\
&= \omega_\xi^2\int_0^r\widetilde{s}\widetilde{V_A}(s)\,ds,
\end{aligned}$$

where

$$\begin{aligned}
\left(B_\xi(s)-D_{\delta,A}c_\theta s\right)^2 &= \left\{B_\xi(s)-c_\theta\omega_\xi\int_0^1W_1(\nu)\widetilde{\nu}d\nu\left(c_\theta^2\int_0^1\widetilde{\nu}^2d\nu\right)^{-1}c_\theta s\right\}^2\\
&= \omega_\xi^2\left\{W_1(s)-\int_0^1W_1(\nu)\widetilde{\nu}d\nu\left(\int_0^1(\widetilde{\nu})^2d\nu\right)^{-1}s\right\}^2\\
&= \omega_\xi^2V_A(s)
\end{aligned}$$

and ω_ξ^2 is the long-run variance of ξ_t. Therefore, following the proofs of Lemmas B4 and B5, the limiting distribution is obtained as

$$\mathcal{T}_\phi(b) = \frac{T^{-3}Z_{nT}(1)\left(T^{-3}M_T\right)^{-1}}{\left\{\left(T^{-3}M_T\right)^{-1}T^{-6}\Psi_{nT}(b)\left(T^{-3}M_T\right)^{-1}\right\}^{1/2}} \Rightarrow \frac{\int_0^1 \widetilde{r}\widetilde{V_A}\left(r\right)dr}{\left\{\int_0^1\int_0^1 K\left(\frac{r-s}{b}\right)\widetilde{r}\widetilde{s}V_A^\tau\left(r\right)V_A^\tau\left(s\right)drds\right\}^{1/2}}$$

where $M_T = \sum_{t=1}^T(\widetilde{t})^2$ and

$$V_A^\tau\left(r\right) = \widetilde{V_A}\left(r\right) - \widetilde{r}\left(\int_0^1 (\widetilde{\nu})^2 d\nu\right)^{-1}\int_0^1 \widetilde{\nu}\widetilde{V_A}\left(\nu\right)d\nu. \tag{S.8}$$

□

Case 3: θ_t is random walk and ξ_t contains a linear trend Suppose $\theta_t = \zeta_{\theta,t}$ and $\xi_t = \tau_t - \delta\theta_t = c_\tau t + (\zeta_{\tau,t} - \delta\zeta_{\theta,t})$. This is the case that one incorrectly chose θ_t that does not contain a linear trend. Thus, the uncontrolled linear trend of τ_t, $c_\tau t$, dominates $\zeta_{\xi,t} = \zeta_{\tau,t} - \delta\zeta_{\theta,t}$ in the regression error ξ_t. In this case, whether $\zeta_{\xi,t}$ is $I(0)$ or $I(1)$, we can show that $\widehat{\delta} \to \infty$ because

$$T^{-1/2}(\widehat{\delta} - \delta) = \frac{T^{-5/2}\sum_{t=1}^T \widetilde{\theta}_t\left(c_\tau t + \zeta_{\xi,t}\right) + o_p\left(1\right)}{T^{-2}\sum_{t=1}^T(\widetilde{\theta}_t)^2} \Rightarrow \frac{c_\tau\int_0^1 r\widetilde{B}_\theta\left(r\right)dr}{\int_0^1 \widetilde{B}_\theta^2\left(r\right)dr} = D_{\delta,B}$$

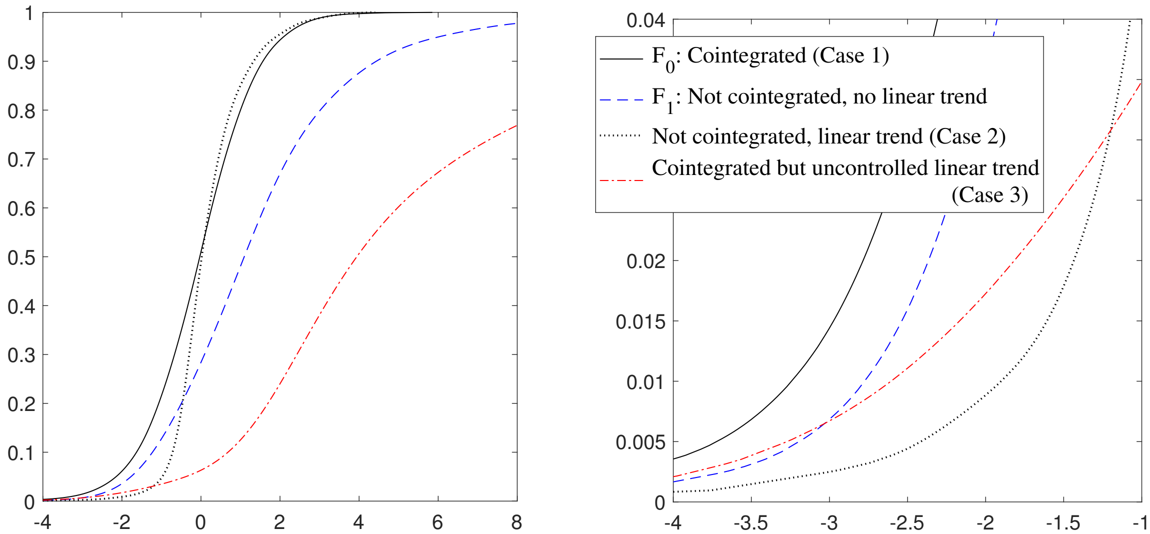
and hence, for S_{nt} in (S.1), the term $\xi_t - (\widehat{\delta} - \delta)\theta_t$ dominates x_{it} like Case 2 above. Following the same steps, we can derive that

$$\begin{aligned}
T^{-4}Z_{nT}(r) &= \frac{1}{T^4}\sum_{t=1}^{[Tr]}\widetilde{t}\left\{\left((c_\tau t + \zeta_{\xi,t}) - (\widehat{\delta} - \delta)\theta_t\right)^2 - \frac{1}{T}\sum_{s=1}^{T}\left((c_\tau s + \zeta_{\xi,s}) - (\widehat{\delta} - \delta)\theta_t\right)^2\right\} + o_p\left(1\right) \\
&\Rightarrow \int_0^r \widetilde{s}\left(c_\tau s - D_{\delta,B}B_\theta\left(s\right)\right)^2 ds - \int_0^r \widetilde{s}ds\int_0^1\left(c_\tau s - D_{\delta,B}B_\theta\left(s\right)\right)^2 ds \\
&= c_\tau^2\int_0^r \widetilde{s}\widetilde{V_B}\left(s\right)ds,
\end{aligned}$$

where

$$\left(c_\tau s - D_{\delta,B}B_\theta\left(s\right)\right)^2 = \left\{c_\tau s - c_\tau\int_0^1 \nu\widetilde{B}_\theta\left(\nu\right)d\nu\left(\int_0^1 \widetilde{B}_\theta^2\left(\nu\right)d\nu\right)^{-1}B_\theta\left(s\right)\right\}^2$$

Figure S1: Limiting distributions of $\mathcal{T}_\phi(b)$ with linear trends



(a) CDF's (b) Left-tail zoom of CDF's

Note: The figure on the left depicts the four CDF's of the limiting distributions. The black solid line is $F_0(b)$, which is also for Case 1 when $\min\{\kappa_1, \kappa_2\} > 1/4$ as in Theorem 1; the black dotted line is for Case 2; and the red dash-dotted line is for Case 3. The blue dashed line is $F_1(b)$ with absence of trend as given in Figure 3. All the distribution functions are simulated with $b = 0.1$ and $T = 5,000$ over 10,000 replications. The figure on the right zooms in the left-side tails of the CDF's.

$$\begin{aligned}
&= c_\tau^2 \left\{ s - \omega_\theta \int_0^1 \nu \widetilde{W}_m(\nu) d\nu \left(\omega_\theta^2 \int_0^1 \widetilde{W}_m^2(\nu) d\nu \right)^{-1} \omega_\theta W_m(s) \right\}^2 \\
&= c_\tau^2 \left\{ s - \int_0^1 \nu \widetilde{W}_m(\nu) d\nu \left(\int_0^1 \widetilde{W}_m^2(\nu) d\nu \right)^{-1} W_m(s) \right\}^2 \\
&= c_\tau^2 V_B(s)
\end{aligned}$$

and ω_θ^2 is the long-run variance of θ_t. Therefore, the limiting distribution is obtained as

$$\mathcal{T}_\phi(b) = \frac{T^{-4} Z_{nT}(1) \left(T^{-3} M_T\right)^{-1}}{\left\{ \left(T^{-3} M_T\right)^{-1} T^{-8} \Psi_{nT}(b) \left(T^{-3} M_T\right)^{-1} \right\}^{1/2}} \Rightarrow \frac{\int_0^1 \widetilde{r} \widetilde{V}_B(r) \, dr}{\left\{ \int_0^1 \int_0^1 K\left(\frac{r-s}{b}\right) \widetilde{r} \widetilde{s} V_B^\tau(r) V_B^\tau(s) \, dr ds \right\}^{1/2}}$$

where $V_B^\tau(r)$ is defined as in (S.8) with $\widetilde{V}_B$. $\square$

When ξ_t is stationary without trend (hence τ_t and θ_t are cointegrated), Case 1 shows that the limiting distribution of $\mathcal{T}_\phi(b)$ remains the same as those in Theorems **1** and **2**, even when τ_t and θ_t have deterministic trends. If τ_t and θ_t are not cointegrated (Case 2), or when the linear trend is not properly controlled for and hence ξ_t is trend-stationary (Case 3), the limiting distributions of $\mathcal{T}_\phi(b)$ are pivotal and behaves in a similar way as the case of $\xi_t \sim I(1)$ in Theorem **1**. Under the latter cases 2 and 3, $S_{n,t}$ would be positively associated with t, and the test $\mathcal{T}_\phi(b)$ should not reject $\phi \geq 0$. Figure S1 depicts the the CDF's of the limiting distributions of those cases with linear trends and their left-side tails. It shows that the the left tails of the cases 2 and 3 are much thinner than that of $F_1(b)$. Therefore, the test $\mathcal{T}_\phi(b)$ can be applied whether or not the nonstationary latent trend τ_t contain deterministic trends over the stochastic trends, using the same critical values from $F_0(b)$ as given in Tables 7 and 8.

S.3 Simulations

We suppose the following data generating process:

$$
\begin{aligned}
y_{it} &= a_i + \tau_t + x_{it}^* \\
x_{it}^* &= \mu_i t^{-\kappa_1} + \epsilon_{it} + \varepsilon_{it} t^{-\kappa_2} \\
\tau_t &= 2\theta_t + \xi_t \\
\theta_t &= \theta_{t-1} + \varepsilon_{\theta,t}
\end{aligned}
$$

where a_i, ϵ_{it}, and $\varepsilon_{\theta,t}$ are $iid\mathcal{N}(0,1)$; $\mu_i \sim iid\mathcal{N}\left(0,\sigma_\mu^2\right)$ and $\varepsilon_{it} \sim iid\mathcal{N}(0,\sigma_\varepsilon^2)$. We let the variances of μ_i and ε_{it} are the same $\sigma_\mu^2 = \sigma_\varepsilon^2 = \sigma^2$, and the decaying rates are the same $\kappa_1 = \kappa_2 = \kappa$. When τ_t and θ_t are cointegrated, we let $\xi_t \sim iid\mathcal{N}(0,1)$; when they are not cointegrated, we let $\Delta\xi_t \sim iid\mathcal{N}(0,1)$. We simulate 5,000 times to obtain the rejection probabilities of $\mathcal{T}_\phi^0(b)$ with the critical value -1.961 of 5% significance level from Table 1, where we use the Bartlett kernel and $b = 0.1$ in HAR estimation. The same set of simulation was done with $\mathcal{T}_\phi(b)$, but the results are very similar and hence omitted.

Table S1 presents the rejection probabilities of the test statistic $\mathcal{T}_\phi^0(0.1)$ under the null case (i.e., Theorem **1**) that $S_{n,t}$ is not negatively associated with t. The first two panels

Table S1: Rejection Probabilities under Null Cases

		T				
ξ_t	n	25	50	100	200	400
$I(0)$	25	0.052	0.056	0.045	0.041	0.047
	50	0.053	0.051	0.051	0.050	0.043
	100	0.052	0.052	0.049	0.049	0.045
	200	0.051	0.054	0.051	0.051	0.053
	400	0.053	0.054	0.047	0.050	0.050
$I(1)$	25	0.033	0.029	0.032	0.030	0.025
	50	0.032	0.030	0.028	0.027	0.021
	100	0.032	0.030	0.028	0.027	0.030
	200	0.033	0.028	0.031	0.024	0.030
	400	0.029	0.031	0.033	0.032	0.026
$I(1)$ & *trend*-θ_t	25	0.018	0.008	0.006	0.006	0.005
	50	0.019	0.012	0.007	0.007	0.002
	100	0.020	0.009	0.005	0.006	0.004
	200	0.017	0.013	0.005	0.005	0.006
	400	0.015	0.008	0.006	0.007	0.004
$I(0)$+*trend*	25	0.015	0.012	0.010	0.010	0.010
	50	0.012	0.012	0.010	0.009	0.013
	100	0.014	0.012	0.012	0.010	0.010
	200	0.011	0.013	0.012	0.013	0.010
	400	0.011	0.012	0.013	0.012	0.009

Note: $\xi_t \sim I(0)$ corresponds to $F_0(b)$; $\xi_t \sim I(1)$ corresponds to $F_1(b)$; $\xi_t \sim I(0)$ & trend-θ_t is Case 2 discussed in Section S.2; $\xi_t \sim I(0)$+*trend* is Case 3 discussed in Section S.2.

Table S2: Rejection Probabilities under Alternative Cases

		$\kappa = 0.20$					$\kappa = 0.25$				
				T					T		
σ^2	n	25	50	100	200	400	25	50	100	200	400
1	25	0.218	0.255	0.283	0.338	0.378	0.233	0.258	0.266	0.294	0.304
	50	0.245	0.291	0.338	0.358	0.401	0.258	0.289	0.316	0.311	0.321
	100	0.269	0.301	0.338	0.403	0.445	0.285	0.298	0.315	0.346	0.352
	200	0.286	0.314	0.355	0.387	0.453	0.299	0.310	0.330	0.330	0.351
	400	0.295	0.320	0.366	0.397	0.444	0.309	0.317	0.336	0.340	0.348
3	25	0.629	0.747	0.832	0.886	0.924	0.674	0.749	0.803	0.825	0.845
	50	0.716	0.808	0.875	0.933	0.953	0.752	0.804	0.841	0.886	0.887
	100	0.761	0.856	0.921	0.943	0.974	0.785	0.846	0.887	0.895	0.921
	200	0.799	0.873	0.927	0.949	0.976	0.814	0.861	0.898	0.906	0.926
	400	0.813	0.883	0.924	0.965	0.977	0.827	0.873	0.896	0.923	0.925
5	25	0.840	0.921	0.971	0.990	0.996	0.878	0.929	0.964	0.975	0.979
	50	0.907	0.968	0.988	0.997	0.999	0.926	0.967	0.979	0.991	0.992
	100	0.944	0.978	0.993	0.998	1.000	0.951	0.975	0.989	0.993	0.996
	200	0.958	0.986	0.996	0.999	1.000	0.963	0.983	0.992	0.996	0.998
	400	0.963	0.989	0.997	0.999	1.000	0.967	0.987	0.993	0.996	0.996
10	25	0.967	0.996	1.000	1.000	1.000	0.985	0.998	1.000	1.000	1.000
	50	0.994	0.999	1.000	1.000	1.000	0.996	0.999	1.000	1.000	1.000
	100	0.996	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000
	200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	400	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

consider two scenarios. (i) $\xi_t \sim I(0)$ case: τ_t and θ_t are cointegrated, but $\gamma(\xi_t^2, t)$ dominates $\gamma(R_t, t)$ as in (16). (ii) $\xi_t \sim I(1)$ case: τ_t and θ_t are not cointegrated as in (15). For these scenarios, we set $x_{it}^* = \epsilon_{it}$ with $\sigma^2 = 1$, where κ is very close to zero and hence the dominance of ξ_t's variance makes the other decaying terms irrelevant. The last two panels consider the cases with trends discussed in Section S.2: (iii) Case 2, where $\xi_t \sim I(1)$ and θ_t imposes a linear trend; and (iv) Case 3, where ξ_t is trend-stationary but θ_t does not have a linear trend. We let $c_\tau = c_\theta = 0.5$. We find that the size is well controlled even in small samples. Furthermore, as discussed in Section S.2, the cases with linear trends well belong to the null case.

These results are also consistent with the shapes of the limiting distributions given in Figure S1, where both $F_1(b)$ (blue dashed line) and the distribution of Case 3 (red dash-dotted line) stochastically dominate $F_0(b)$ (black solid line). Furthermore, while the distribution under Case 2 does not stochastically dominate $F_0(b)$, it also has a positive mode like $F_1(b)$ and a very thin left tail, and thus does not affect our one-sided test.

Table S2 summarizes the the rejection probabilities of $\mathcal{T}_\phi^0(0.1)$ under the alternative case (i.e., Theorem **2**) that $S_{n,t}$ is negatively associated with t. We consider $\kappa \in \{0.20, 0.25\}$ and change the values of $\sigma^2 \in \{1, 3, 5, 10\}$. As well predicted from Theorem **2**, the rejection probability (i.e., power of the test) improves as σ^2 gets large, which is because the variance ratio ω_*^2 between x_{it}^* and ξ_t increases. (Recall the variance of ξ_t is fixed as unity and hence $\omega_*^2 = \sigma^2$ here.) The power improves with n, T as well, with the effect of T being more pronounced.

S.4 Supplementary Results on Crime Rate Example

Table S3 provides the summary statistics, the source of each data, and the data details used in the crime rate applications of the paper. All values in the table are presented in levels (i.e., before log-transformation). Tables S4 and S5 report the results in Tables 3 and 6, respectively, where the trend determinant variables for θ_t are lagged by two periods, instead of one.

Remark As noted in the footnote 10, we can enrich the subgroup estimate by examining any potentially missing convergent members in $\widehat{\mathcal{G}}(\theta)^c$ using the following automated procedure. First, for all $i \in \widehat{\mathcal{G}}(\theta)^c$, we sort individuals by the distance from y_{it} toward the common trend (i.e., $d_{it} = y_{it} - \widehat{\delta}'\theta_t$) during the most recent sampling periods $t = T^\epsilon, \ldots, T$, where $T^\epsilon = T - [\epsilon T]$ for some small $\epsilon > 0$. This is based on the observation that divergent series often exhibit signs of divergence towards the end of the sample period (e.g., Phillips and Sul (2007)). The order statistics of such vectors $d_i^\epsilon = (d_{iT^\epsilon}, \ldots, d_{iT})'$ can be obtained using the forecast depth by Lee and Sul (2023a). For instance, we can use the Mahalanobis forecast depth given as (see Lee and Sul (2023b) for other depths)

$$\mathcal{D}_i = \frac{1}{1 + d_i^{\epsilon\prime} V_\epsilon^{-1} d_i^\epsilon} \quad \text{with} \quad V_\epsilon = \frac{1}{|\widehat{\mathcal{G}}(\theta)^c|} \sum\nolimits_{i \in \widehat{\mathcal{G}}(\theta)^c} d_i^\epsilon d_i^{\epsilon\prime}.$$

Second, we sort individuals by their forecast depths in descending order, which shows the degree of proximity of y_{it} to the common trend $\widehat{\delta}'\theta_t$ over $t = T^\epsilon, \ldots, T$.

Third, we sequentially add individuals with the largest depth to the subgroup estimate, who are not originally included in $\widehat{\mathcal{G}}(\theta)$, until the t-test $\mathcal{T}_\phi^0(b)$ using the extended subgroup exceeds the critical value (e.g., -1.96 for $b = 0.1$ and 5% significance level). This yields the enriched subgroup estimate. In our property crime analysis, this procedure indeed improves the convergent subgroup size for the 'Prison' variable.

Table S3: Data Description and Summary Statistics

	Mean	Std. dev.	Min	Max	Period	Source
Violent Crimes						
All Violent Crime	431.29	212.88	56.85	1244.33	1987-2021	FBI UCR
Assault	277.60	143.95	34.09	785.72	1987-2021	FBI UCR
Robbery	110.20	81.50	6.40	624.66	1987-2021	FBI UCR
Homicide	5.40	3.06	0.16	20.35	1987-2021	FBI UCR
Property Crimes						
All Property Crime	3424.69	1210.91	964.70	7819.90	1987-2021	FBI UCR
Burglary	735.78	350.37	73.73	2294.26	1987-2021	FBI UCR
Larceny	2358.46	769.40	711.91	5106.13	1987-2021	FBI UCR
Motor Vehicle Theft	330.45	197.94	29.52	1157.66	1987-2021	FBI UCR
Trend Variables						
population (age 10-19)	13.92	0.10	12.76	14.69	1986-2020	Census
population (age 20-29)	14.44	0.20	13.36	17.76	1986-2020	Census
population (age 30-39)	14.77	0.27	12.87	17.06	1986-2020	Census
population (age 40-49)	13.68	0.21	10.94	15.38	1986-2020	Census
police officer	2.21	0.02	2.03	2.41	1986-2020	FBI UCR
incarceration	2.09	0.07	1.14	2.59	1986-2020	BJS
real GDP	50016.86	1312.64	36698	62606	1986-2020	FRED

Notes: (i) FBI UCR is the FBI Uniform Crime Report; Census is the U.S. Census Bureau; BJS is the Bureau of Justice Statistics; FRED is the Federal Reserve Economic Data at the St. Louis Fed. (ii) All the 'Crimes' are defined as the number of crimes per 100,000 population; 'population' is percentage of population in each specific age group (Demog = log(population)); 'police officer' is the number of non-civilian police officers per 1,000 population (Police = log(police officer)); 'incarceration' is the incarceration count per 1,000 population (Prison = log(incarceration)); 'real GDP' is real GDP per capita in 2017 dollars (RGDP = log(real GDP)).

Table S4: Long-Run Trend Determinants for Violent Crimes (with 2-year lagged variables)

Crime	θ_t	$\hat{\delta}$	$se(\hat{\delta})$	$\mathcal{T}_\phi(0.1)$	$\mathcal{T}_\phi^0(0.1)$
Violent	Demog	3.684*	0.323	-7.858*	-9.346*
	Police	-0.924	1.285	-3.591*	-6.313*
	Prison	-0.885*	0.208	0.380	0.541
	RGDP	-1.458*	0.142	66.116	7.435
Assault	Demog	3.127*	0.366	-4.124*	-9.247*
	Police	-0.802	1.100	-1.798	-4.364*
	Prison	-0.778*	0.197	0.500	0.919
	RGDP	-1.221*	0.163	35.386	5.968
Homicide	Demog	3.124*	0.531	-6.514*	-3.538*
	Police	-2.056*	1.180	-0.306	-3.660*
	Prison	-0.973	0.182	0.055	1.294
	RGDP	-1.189*	0.267	11.161	2.606
Robbery	Demog	5.780*	0.587	-20.075*	-2.582*
	Police	-0.532	2.072	-2.439*	-8.898*
	Prison	-1.184*	0.323	-0.668	-1.287
	RGDP	-2.367*	0.209	84.106	2.806

Note: (i) $\hat{\delta}$ is the least squares estimate from (10) and $se(\hat{\delta})$ is its standard error from Phillips and Park (1988). $\mathcal{T}_\phi(0.1)$ and $\mathcal{T}_\phi^0(0.1)$ are respectively the t-ratios defined in (12) and (14) with $b = 0.1$ and the Bartlett kernel. (ii) 'Demog' is the log of the fraction of young adult population between age 10 and 39, 'Police' is the log of the number of non-civilian police officers per capita, 'Prison' is the log of the local incarceration per capita, and 'RGDP' is the log of the Real GDP per capita. (iii) From Definition 1, θ_t becomes a long-run trend determinant if $\widehat{\delta}$ is significantly different from zero and $\mathcal{T}_\phi(0.1) < -2.04$ or $\mathcal{T}_\phi^0(0.1) < -1.96$, where the 5% critical values are from Table 1; only 'Demog' satisfies these two conditions. (* indicates significance at 5%.)

Table S5: Long-run Trend Determinants for Property Crimes (with 2-year lagged variables)

Crime	θ_t	$\hat{\delta}$	$se(\hat{\delta})$	$se^0(\hat{\delta})$	$\mathcal{T}_\phi(0.1)$	$\mathcal{T}_\phi^0(0.1)$	Group
Property	Demog	n.a.	n.a.	n.a.	n.a.	n.a.	0
	Police	0.280	1.500	1.760	-21.630*	-26.331*	50
	Prison	-0.910*	0.220	0.360	-3.054*	-3.461*	40
	RGDP	n.a.	n.a.	n.a.	n.a.	n.a.	0
Burglary	Demog	3.370*	0.500	0.550	-11.708*	-12.086*	1
	Police	0.570	2.150	2.400	-12.385*	-10.976*	49
	Prison	-1.100*	0.340	–	-3.094*	–	42
		-1.090*	–	0.500	–	-3.245*	43
	RGDP	-3.900*	0.320	–	-1.260	–	1
		n.a.	–	n.a.	–	n.a.	0
Larceny	Demog	2.080*	0.450	0.400	-1.623	-1.688	1
	Police	0.310	1.320	1.590	-17.378*	-22.701*	50
	Prison	-0.830*	0.200	0.320	-2.948*	-3.488*	40
	RGDP	-2.270*	0.160	–	-1.238	–	1
		n.a.	–	n.a.	–	n.a.	0
Motor Vehicle Theft	Demog	5.862*	0.420	0.491	-17.238*	-2.387*	
	Police	-0.682	2.261	2.377	0.302	-6.521*	
	Prison	-1.389*	0.302	0.385	-1.097	-1.046	
	RGDP	-2.278*	0.189	0.281	53.495	1.571	

Note: (i) 'Group' is the number of states selected in the subgroup estimate $\widehat{\mathcal{G}}(\theta)$ using $\mathcal{T}_{\varphi_i}(0.1)$ and $\mathcal{T}_{\varphi_i}^0(0.1)$. When they are different, each of $\hat{\delta}$ and $\mathcal{T}_\phi(0.1)$ are reported in separate lines. (ii) $\hat{\delta}$ is the least squares estimate from (10), and $se(\hat{\delta})$ and $se^0(\hat{\delta})$ are the standard error from Phillips and Park (1988) using $\mathcal{T}_{\varphi_i}(0.1)$ and $\mathcal{T}_{\varphi_i}^0(0.1)$, respectively. $\mathcal{T}_\phi(0.1)$ and $\mathcal{T}_\phi^0(0.1)$ are the t-ratios defined in (12) and (14) with $b = 0.1$ and the Bartlett kernel. When $\widehat{\mathcal{G}}(\theta)$ is empty, $\hat{\delta}$ and $\mathcal{T}_\phi(b)$ cannot be obtained and marked as 'n.a.'. (iii) 'Demog' is the log of the fraction of young adult population between age 10 and 39, 'Police' is the log of the number of non-civilian police officers per capita, 'Prison' is the log of the local incarceration per capita, and 'RGDP' is the log of the Real GDP per capita. (iv) From Definition 1, θ_t becomes a long-run trend determinant if $\widehat{\delta}$ is significantly different from zero and $\mathcal{T}_\phi(0.1) < -2.04$ or $\mathcal{T}_\phi^0(0.1) < -1.96$. Furthermore, the size of the subgroup estimate $\widehat{\mathcal{G}}(\theta)$ should be large enough to include most of the states. For 'Property', 'Burglary', and 'Larceny', only 'Prison' satisfies all these three conditions. (v) 'Motor Vehicle Theft' satisfies the weak σ-convergence and hence does not require to get the subgroup, so no group size is given. For this case, 'Demog' is identified as a long-run trend determinant, which is defined as the population fraction of age from 10 to 49. (* indicates significance at 5%.)

References

Chang, Y., J. Y. Park, and P. C. B. Phillips (2001): "Nonlinear econometric models with cointegrated and deterministically trending regressors," *The Econometrics Journal*, 4(1), 1–36.

Lee, Y., and D. Sul (2023a): "Depth-weighted Forecast Combination: Application to COVID-19 Cases," *Advances in Econometrics*, 45B, 235–260.

Lee, Y., and D. Sul (2023b): "Depth-weighted means of noisy data: An application to estimating the average effect in heterogeneous panels," *Journal of Multivariate Analysis*, 196, 105–165.

MacNeill, I. B. (1978): "Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times," *The Annals of Statistics*, 6(2), 422–433.

Park, J. Y. and P. C. B. Phillips (1988): "Statistical inference in regressions with integrated processes: Part 1," *Econometric Theory*, 4(3), 468–497.

Phillips, P. C. B. and H. R. Moon (1999): "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67(5), 1057–1111.

Phillips, P. C. B., and J. Y. Park (1988): "Asymptotic Equivalence of Ordinary Least Squares and Generalized Least Squares in Regressions with Integrated Regressors," *Journal of the American Statistical Association*, 83(401), 111–115.

Phillips, P. C. B. and D. Sul (2007): "Transition modeling and econometric convergence tests," *Econometrica*, 75(6), 1771–1855.

Sun, Y. (2004): "A Convergent t-Statistic in Spurious Regressions," *Econometric Theory*, 20(5), 943-962.