

# SUPPLEMENTARY MATERIAL for “Endogenous kink threshold regression”

Jianhan Zhang<sup>\*</sup>    Chaoyi Chen<sup>†‡</sup>    Yiguo Sun<sup>§</sup>    Thanasis Stengos<sup>¶</sup>

October 16, 2023

This online supplement is composed of four parts. Section A contains the proof of Theorem 1-Time Series. Section B contains the proof of Theorem 2-panel. Section C shows the proof of Lemma 1. Section D gives the Monte Carlo simulation results for DGP2 and DGP4.

## A Proof of Theorem 1-time series

We first show that  $\sup_{\phi \in B \times \Gamma} |S_n(\phi) - S_n^*(\phi)| \xrightarrow{p} 0$ , which implies that minimizing  $S_n(\phi)$  with respect to  $\phi$  is equivalent to minimizing  $S_n^*(\phi)$ , where  $S_n(\phi)$  and  $S_n^*(\phi)$  are defined in (??) and Assumption T7, respectively. Then, closely following the mathematical proof of Hansen (2017), we can show that the optimization problem of  $S_n^*(\phi)$  with respect to  $\phi$  satisfies all the required conditions in Section 3.2 of Andrews (1994), which completes the proof of this theorem. Hence, below, we only need to prove  $\sup_{\phi \in B \times \Gamma} |S_n(\phi) - S_n^*(\phi)| \xrightarrow{p} 0$ .

Now, we denote  $P_\gamma = X(\gamma)[X'(\gamma)X(\gamma)]^{-1}X'(\gamma)$  and  $P_\gamma^* = X^*(\gamma)[X^{*'}(\gamma)X^*(\gamma)]^{-1}X^{*'}(\gamma)$ , where  $X^*(\gamma)$  is defined in the same form as  $X(\gamma)$  by replacing  $x_t(\gamma)$  with  $x_t^*(\gamma)$ . For any

---

<sup>\*</sup>Department of Economics and Finance, University of Guelph, Guelph, Ontario, N1G 2W1, Canada; Email: [jzhang56@uoguelph.ca](mailto:jzhang56@uoguelph.ca).

<sup>†</sup>Magyar Nemzeti Bank (Central Bank of Hungary), Budapest, 1054, Hungary; Email: [chenc@mnk.hu](mailto:chenc@mnk.hu).

<sup>‡</sup>MNB Institute, John von Neumann University, Kecskemét, 6000, Hungary.

<sup>§</sup>Department of Economics and Finance, University of Guelph, Guelph, Ontario, N1G 2W1, Canada; Email: [yisun@uoguelph.ca](mailto:yisun@uoguelph.ca).

<sup>¶</sup>Department of Economics and Finance, University of Guelph, Guelph, Ontario, N1G 2W1, Canada; Email: [tstengos@uoguelph.ca](mailto:tstengos@uoguelph.ca).

$\phi \in B \times \Gamma$ , we have

$$S_n(\phi) = \frac{1}{n} \sum_{t=1}^n [y_t - x'_t(\gamma)\theta]^2 = \frac{1}{n} y' (I_n - P_\gamma) y. \quad (\text{A.1})$$

Let  $X - \gamma_0 \tau_n$  and  $I_n(\gamma_0)(X - \gamma_0 \tau_n)$  be the matrix form of  $x_t - \gamma_0$  and  $(x_t - \gamma_0)I(x_t \geq \gamma_0)$  respectively, where  $\tau_n$  is an  $n \times 1$  vector of ones,  $I_n(\gamma_0) = \text{diag}\{I(x_1 \geq \gamma_0), \dots, I(x_n \geq \gamma_0)\}$ . Also, let  $\varepsilon$  stack up  $\varepsilon_i$ ,  $v$  stack up  $v_i$ , and  $\hat{v}$  stack up  $\hat{v}_i$  for  $i = 1, \dots, n$ . Below, we decompose  $S_n(\phi)$ . By simple calculation, we have

$$y = X^*(\gamma_0)\theta_0 + \varepsilon = X(\gamma)\theta_0 + [X^*(\gamma) - X(\gamma)]\theta_0 + [X^*(\gamma_0) - X^*(\gamma)]\theta_0 + \varepsilon,$$

where we have

$$\begin{aligned} X^*(\gamma) - X(\gamma) &= [\mathbf{0}_{n \times 1}, \mathbf{0}_{n \times 1}, \mathbf{0}_{n \times \ell}, v - \hat{v}], \\ X^*(\gamma_0) - X^*(\gamma) &= [(\gamma - \gamma_0)\tau_n, I_n(\gamma_0)(X - \gamma_0\tau_n) - I_n(\gamma)(X - \gamma\tau_n), \mathbf{0}_{n \times \ell}, \mathbf{0}_{n \times (1+d_{z1})}]. \end{aligned}$$

Note that  $X'(\gamma)(I_n - P_\gamma) = \mathbf{0}_{(k+d_{z1}) \times n}$  and  $\hat{\varepsilon} = (v - \hat{v})\beta_{40} + \varepsilon$ , where  $\hat{\varepsilon}$  stacks up  $\hat{\varepsilon}_i$  for  $i = 1, \dots, n$ . Thus, we can show

$$\begin{aligned} S_n(\phi) &= \frac{1}{n} \{ \tau_n(\gamma - \gamma_0)\beta_{10} + [I_n(\gamma_0)(X - \gamma_0\tau_n) - I_n(\gamma)(X - \gamma\tau_n)]\delta_0 + \hat{\varepsilon} \}' (I_n - P_\gamma) \times \\ &\quad \{ \tau_n(\gamma - \gamma_0)\beta_{10} + [I_n(\gamma_0)(X - \gamma_0\tau_n) - I_n(\gamma)(X - \gamma\tau_n)]\delta_0 + \hat{\varepsilon} \} \\ &= \frac{1}{n} \hat{\varepsilon}' (I_n - P_\gamma) \hat{\varepsilon} + \frac{2}{n} \delta_0 [I_n(\gamma)\tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0\tau_n)]' (I_n - P_\gamma) \hat{\varepsilon} \\ &\quad + \frac{1}{n} \delta_0^2 [I_n(\gamma)\tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0\tau_n)]' \\ &\quad \times (I_n - P_\gamma) [I_n(\gamma)\tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0\tau_n)] \end{aligned} \quad (\text{A.2})$$

where  $d(\gamma_0, \gamma) = I_n(\gamma_0) - I_n(\gamma)$ .

Next, we decompose  $S_n^*(\phi)$ . By simple calculation, we have

$$y = X^*(\gamma_0)\theta_0 + \varepsilon = X^*(\gamma)\theta_0 + [X^*(\gamma_0) - X^*(\gamma)]\theta_0 + \varepsilon.$$

As  $X^{*'}(\gamma)(I_n - P_\gamma^*) = \mathbf{0}_{(k+d_{z1}) \times n}$ , we obtain

$$\begin{aligned}
S_n^*(\phi) &= \frac{1}{n} \{ \tau_n(\gamma - \gamma_0) \beta_{10} + [I_n(\gamma_0)(X - \gamma_0 \tau_n) - I_n(\gamma)(X - \gamma \tau_n)] \delta_0 + \varepsilon \}' (I_n - P_\gamma^*) \times \\
&\quad \{ \tau_n(\gamma - \gamma_0) \beta_{10} + [I_n(\gamma_0)(X - \gamma_0 \tau_n) - I_n(\gamma)(X - \gamma \tau_n)] \delta_0 + \varepsilon \} \\
&= \frac{1}{n} \varepsilon' (I_n - P_\gamma^*) \varepsilon + \frac{2}{n} \delta_0' [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' (I_n - P_\gamma^*) \varepsilon \\
&\quad + \frac{1}{n} \delta_0^2 [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' (I_n - P_\gamma^*) \\
&\quad \times [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)].
\end{aligned}$$

Subtracting (A.3) from (A.2), we have

$$\begin{aligned}
&S_n(\phi) - S_n^*(\phi) \\
&= \frac{1}{n} (\hat{\varepsilon}' \hat{\varepsilon} - \varepsilon' \varepsilon) + \frac{2}{n} \delta_0' [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' (\hat{\varepsilon} - \varepsilon) \\
&\quad - \frac{1}{n} \delta_0^2 [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' (P_\gamma - P_\gamma^*) [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)] \\
&\quad - \frac{1}{n} (\hat{\varepsilon}' P_\gamma \hat{\varepsilon} - \varepsilon' P_\gamma^* \varepsilon) + \frac{2}{n} \delta_0' [I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' (P_\gamma^* \varepsilon - P_\gamma \hat{\varepsilon}) \} \\
&= S_1 + 2S_2 - S_{31} - S_{32} + 2S_{33}.
\end{aligned}$$

Below, we show that  $S_1$ ,  $S_2$ ,  $S_{31}$ ,  $S_{32}$ , and  $S_{33}$  are all  $o_p(1)$  uniformly over the domain of  $\phi$ .

**S<sub>1</sub>:** Denoting  $\hat{\pi} = [\hat{\pi}'_x, \hat{\pi}'_z]'$ ,  $\pi_0 = [\pi'_{x0}, \pi'_{z0}]'$  and  $\rho_t = [p'_{xt}, p'_{zt}]'$ , we have  $\hat{\pi} - \pi_0 = O_p(n^{-1/2})$  under Assumptions T2(b) and T5(b) and

$$\begin{aligned}
S_1 &= \frac{1}{n} (\hat{\varepsilon}' \hat{\varepsilon} - \varepsilon' \varepsilon) \\
&= \frac{2}{n} \beta'_{40} (v - \hat{v})' \varepsilon + \frac{1}{n} \beta'_{40} (v - \hat{v})' (v - \hat{v}) \beta_{40} \\
&= O_p(n^{-1/2}) + O_p(n^{-1}) = o_p(1),
\end{aligned} \tag{A.3}$$

since  $n^{-1} \varepsilon' \varepsilon = \sigma_\varepsilon^2 + o_p(1)$  under Assumption T5 and

$$\|v - \hat{v}\|^2 = (\hat{\pi} - \pi_0)' \sum_{t=1}^n \rho_t \rho'_t (\hat{\pi} - \pi_0) \leq n \|\hat{\pi} - \pi_0\|^2 \lambda_{\max}(n^{-1} \sum_{t=1}^n \rho_t \rho'_t) = O_p(1) \tag{A.4}$$

under Assumption T2(b), where  $\lambda_{max}(A)$  denotes the largest eigenvalue of a symmetric matrix  $A$ .

**S<sub>2</sub>:** By Assumptions T2 (a), T7 and applying  $\frac{1}{n} \|\hat{\varepsilon} - \varepsilon\| = \frac{1}{n} \|(v - \hat{v})\beta_{40}\| \leq \frac{1}{n} \|\beta_{40}\| \|v - \hat{v}\| = O_p(n^{-1})$ , for all  $\gamma \in \Gamma$ , we can show

$$S_2 = \frac{1}{n} \delta_0(\gamma - \gamma_0) \tau_n' I_n(\gamma) (\hat{\varepsilon} - \varepsilon) + \frac{1}{n} \delta_0(X - \gamma_0 \tau_n)' d(\gamma_0, \gamma) (\hat{\varepsilon} - \varepsilon) = O_p(n^{-1/2}) + O_p(n^{-1/2}) = o_p(1).$$

**S<sub>31</sub>, S<sub>32</sub>, S<sub>33</sub>:** By showing  $S_{31}, S_{32}, S_{33}$  are all  $o_p(1)$  for any  $\phi \in B \times \Gamma$ , we only need to prove:

$$\max_{\gamma \in \Gamma} \frac{1}{n} \|[I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' [X(\gamma) - X^*(\gamma)]\| = o_p(1), \quad (\text{A.5})$$

$$\max_{\gamma \in \Gamma} \frac{1}{n} \|\hat{\varepsilon}' X(\gamma) - \varepsilon' X^*(\gamma)\| = o_p(1), \quad (\text{A.6})$$

$$\max_{\gamma \in \Gamma} \frac{1}{n} \|X'(\gamma) X(\gamma) - X^{*'}(\gamma) X^*(\gamma)\| = o_p(1). \quad (\text{A.7})$$

First, applying (A.4), we have

$$\max_{\gamma \in \Gamma} \frac{1}{n} \|X(\gamma) - X^*(\gamma)\| = \frac{1}{n} \|\hat{v} - v\| = O_p(n^{-1}) = o_p(1). \quad (\text{A.8})$$

Next, by Assumptions T2(a) and T7, we have

$$\begin{aligned} & \max_{\gamma \in \Gamma} \frac{1}{n} \|[I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)]' [X(\gamma) - X^*(\gamma)]\| \\ & \leq \max_{\gamma \in \Gamma} \frac{1}{n} \|I_n(\gamma) \tau_n(\gamma - \gamma_0) + d(\gamma_0, \gamma)(X - \gamma_0 \tau_n)\| \|X(\gamma) - X^*(\gamma)\| = O_p(n^{-1/2}) = o_p(1), \end{aligned}$$

which verifies (A.5).

Next, we show (A.6). Applying triangle inequality gives

$$\begin{aligned} & \max_{\gamma \in \Gamma} \frac{1}{n} \|\hat{\varepsilon}' X(\gamma) - \varepsilon' X^*(\gamma)\| \\ & \leq \max_{\gamma \in \Gamma} \frac{1}{n} \|\varepsilon' [X(\gamma) - X^*(\gamma)]\| + \max_{\gamma \in \Gamma} \frac{1}{n} \|\beta_{40}' (v - \hat{v})' X(\gamma)\| \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} & = \frac{1}{n} \|\varepsilon' (\hat{v} - v)\| + \max_{\gamma \in \Gamma} \frac{1}{n} \|\beta_{40}' (v - \hat{v})' X(\gamma)\| \\ & = O_p(n^{-1/2}) + O_p(n^{-1/2}) = o_p(1), \end{aligned} \quad (\text{A.10})$$

where  $\frac{1}{n} \|\varepsilon'(\hat{v} - v)\| = O_p(n^{-1/2})$  by (A.3), and  $\max_{\gamma \in \Gamma} \frac{1}{n} \|(v - \hat{v})'X(\gamma)\| \leq \frac{1}{n} \|v - \hat{v}\| \|X(\gamma)\| = O_p(n^{-1/2})$  by (A.4) and under Assumption T2.

Finally, by Assumption P2 and (A.8), we can show

$$\begin{aligned} & \max_{\gamma \in \Gamma} \frac{1}{n} \|X'(\gamma)X(\gamma) - X^{*'}(\gamma)X^*(\gamma)\| \\ &= \max_{\gamma \in \Gamma} \frac{2}{n} \|(X - \gamma\tau_n)'(\hat{v} - v)\| + \max_{\gamma \in \Gamma} \frac{2}{n} \|[I_n(\gamma)(X - \gamma\tau_n)]'(\hat{v} - v)\| + \frac{2}{n} \|z'(\hat{v} - v)\| + \frac{1}{n} \|\hat{v}'\hat{v} - v'v\| \\ &= O_p(n^{-1/2}) + O_p(n^{-1/2}) + O_p(n^{-1/2}) + \frac{1}{n} \|\hat{v}\hat{v}' - vv'\|, \end{aligned} \quad (\text{A.11})$$

where  $z$  is an  $n \times \ell$  matrix, and  $z = [z_1, z_2, \dots, z_n]'$ . Then, we only left to show  $\frac{1}{n} \|\hat{v}'\hat{v} - v'v\| = o_p(1)$ . Note that,

$$\begin{aligned} \frac{1}{n} \|\hat{v}'\hat{v} - v'v\| &= \frac{1}{n} \|(\hat{v} - v)'(\hat{v} - v)\| + \frac{2}{n} \|(\hat{v} - v)'v\| \\ &= O_p(n^{-1}) + A, \end{aligned} \quad (\text{A.12})$$

where, under Assumption T2, for any bounded  $v_t$ , we have

$$A = \frac{1}{n} \|(\hat{v} - v)'v\| \leq \frac{1}{n} \|(\hat{v} - v)\| \|v\| = O_p(n^{-1/2}). \quad (\text{A.13})$$

Thus, combining (A.11),(A.12) with (A.13), we obtain

$$\frac{1}{n} \max_{\gamma \in \Gamma} \|X'(\gamma)X(\gamma) - X^{*'}(\gamma)X^*(\gamma)\| = o_p(1).$$

To sum up, we have  $\sup_{\phi \in B \times \Gamma} |S_n(\phi) - S_n^*(\phi)| \xrightarrow{p} 0$ , which completes the proof of Theorem 1-time series.

## B Proof of Theorem 2-panel

Similar to the proof of Theorem 1-Time Series, we first prove  $\sup_{\phi \in B \times \Gamma} |S_{NT}(\phi) - S_{NT}^*(\phi)| \xrightarrow{p} 0$ , which implies that the minimizer of  $S_{NT}(\phi)$  is also the minimizer of  $S_{NT}^*(\phi)$ , where  $\phi = (\theta', \gamma)'$ , and the definition of  $S_{NT}(\phi)$  and  $S_{NT}^*(\phi)$  are given by (??) and Assumption P7,

respectively. Then, by showing that, for  $\phi \in B \times \Gamma$ , the optimization problem of  $S_{NT}^*(\phi)$  satisfies all the four required conditions in Section 3.2 of Andrews (1994), we verify Theorem 1-panel.

We first show that  $\sup_{\phi \in B \times \Gamma} |S_{NT}(\phi) - S_{NT}^*(\phi)| \xrightarrow{p} 0$ .

Define  $P(\gamma) = \Delta x(\gamma)[\Delta x(\gamma)' \Delta x(\gamma)]^{-1} \Delta x(\gamma)'$  and  $P^*(\gamma)$  is in the same form of  $P(\gamma)$  with  $\Delta x(\gamma)$  being replaced by  $\Delta x^*(\gamma)$ , where  $\Delta x(\gamma) = [\Delta x_{12}(\gamma), \dots, \Delta x_{1T}(\gamma), \dots, \Delta x_{N2}, \dots, \Delta x_{NT}(\gamma)]'$  is an  $[N(T-1)] \times (k + d_{z1})$  matrix, and  $\Delta x^*(\gamma)$  equals  $\Delta x(\gamma)$  with  $\Delta x_{it}(\gamma)$  being replaced by  $\Delta x_{it}^*(\gamma)$ . Then, for any  $\phi \in B \times \Gamma$ ,  $S_{NT}(\phi)$  can be expressed as

$$S_{NT}(\phi) = \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T (\Delta y_{it} - \Delta x_{it}(\gamma)' \theta)^2 = \frac{1}{N(T-1)} \Delta y' P(\gamma) \Delta y,$$

where  $\Delta y = [\Delta y_{12}, \dots, \Delta y_{1T}, \dots, \Delta y_{N2}, \dots, \Delta y_{NT}]'$  is an  $[N(T-1)] \times 1$  vector. Note that, we can rewrite  $\Delta y$  as

$$\Delta y = \Delta x^*(\gamma_0) \theta_0 + \Delta \varepsilon = [\Delta x^*(\gamma_0) - \Delta x(\gamma)] \theta_0 + \Delta x(\gamma) \theta_0 + \Delta \varepsilon,$$

where  $\Delta \varepsilon = [\Delta \varepsilon_{12}, \dots, \Delta \varepsilon_{1T}, \dots, \Delta \varepsilon_{N2}, \dots, \Delta \varepsilon_{NT}]'$  an  $[N(T-1)] \times 1$  vector.

As  $\Delta x(\gamma)' [I_{NT} - P(\gamma)] = \mathbf{0}_{(1+d_{z1}) \times N(T-1)}$ , we have

$$\begin{aligned} S_{NT}(\phi) &= \frac{1}{N(T-1)} \{ [\Delta x^*(\gamma_0) - \Delta x(\gamma)] \theta_0 + \Delta \varepsilon \}' [I_N - P(\gamma)] \\ &\quad \times \{ [\Delta x^*(\gamma_0) - \Delta x(\gamma)] \theta_0 + \Delta \varepsilon \}. \end{aligned}$$

where  $I_{NT}$  is an identity matrix with dimension  $N(T-1)$ .

Similarly, we can also decompose  $\Delta y$  as

$$\Delta y = \Delta x^*(\gamma_0) \theta_0 + \Delta \varepsilon = [\Delta x^*(\gamma_0) - \Delta x^*(\gamma)] \theta_0 + \Delta x^*(\gamma) \theta_0 + \Delta \varepsilon.$$

Since  $\Delta x^{*'}(\gamma) [I_{NT} - P^*(\gamma)] = \mathbf{0}_{(k+d_{z1}) \times N(T-1)}$ , we have

$$\begin{aligned} S_{NT}^*(\phi) &= \frac{1}{N(T-1)} \{ [\Delta x^*(\gamma_0) - \Delta x^*(\gamma)] \theta_0 + \Delta \varepsilon \}' [I_{NT} - P^*(\gamma)] \\ &\quad \times \{ [\Delta x^*(\gamma_0) - \Delta x^*(\gamma)] \theta_0 + \Delta \varepsilon \}. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
& S_{NT}(\phi) - S_{NT}^*(\phi) \\
&= \frac{1}{N(T-1)} \{[\Delta x^*(\gamma) - \Delta x(\gamma)]\theta_0\}' [I_{NT} - P(\gamma)] \{[\Delta x^*(\gamma) - \Delta x(\gamma)]\theta_0\} \\
&\quad + \frac{1}{N(T-1)} \{[\Delta x^*(\gamma_0) - \Delta x^*(\gamma)]\theta_0 + \Delta\varepsilon\}' [P^*(\gamma) - P(\gamma)] \{[\Delta x^*(\gamma_0) - \Delta x^*(\gamma)]\theta_0 + \Delta\varepsilon\} \\
&\quad + \frac{2}{N(T-1)} \{[\Delta x^*(\gamma) - \Delta x(\gamma)]\theta_0\}' [I_{NT} - P(\gamma)] \{[\Delta x^*(\gamma_0) - \Delta x^*(\gamma)]\theta_0 + \Delta\varepsilon\} \\
&= \mathcal{S}_1 + \mathcal{S}_2 + 2\mathcal{S}_3,
\end{aligned}$$

where denoting  $\Delta w = [\Delta w'_{12}, \dots, \Delta w'_{1T}, \dots, \Delta w'_{N2}, \dots, \Delta w'_{NT}]'$  for  $w = v$  and  $\hat{v}$  and  $\chi(\gamma) = [(X_{12} - \gamma\tau_2)' \mathbf{I}_{12}(\gamma), \dots, (X_{1T} - \gamma\tau_2)' \mathbf{I}_{1T}(\gamma), \dots, (X_{N2} - \gamma\tau_2)' \mathbf{I}_{N2}(\gamma), \dots, (X_{NT} - \gamma\tau_2)' \mathbf{I}_{NT}(\gamma)]'$ , we have

$$\begin{aligned}
\Delta x^*(\gamma) - \Delta x(\gamma) &= [\mathbf{0}_{N(T-1) \times 1}, \mathbf{0}_{N(T-1) \times 1}, \mathbf{0}_{N(T-1) \times 1}, \Delta v - \Delta \hat{v}], \\
\Delta x^*(\gamma_0) - \Delta x^*(\gamma) &= [\mathbf{0}_{N(T-1) \times 1}, \chi(\gamma_0) - \chi(\gamma), \mathbf{0}_{N(T-1) \times 1}, \mathbf{0}_{N(T-1) \times (1+d_{z1})}].
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
\mathcal{S}_1 &= \frac{1}{N(T-1)} \beta'_{40} (\Delta v - \Delta \hat{v})' (I_{NT} - P(\gamma)) (\Delta v - \Delta \hat{v}) \beta_{40} \\
&\leq \frac{1}{N(T-1)} \lambda_{\max}(I_{NT} - P(\gamma)) \beta'_{40} (\Delta v - \Delta \hat{v})' (\Delta v - \Delta \hat{v}) \beta_{40} \\
&= O_p([N(T-1)]^{-1})
\end{aligned} \tag{B.1}$$

since  $I_{NT} - P(\gamma)$  is an idempotent matrix with  $\lambda_{\max}(I_{NT} - P(\gamma)) = 1$  and denoting  $\hat{\Pi} = [\hat{\Pi}'_x, \hat{\Pi}'_z]'$ ,  $\Pi_0 = [\Pi'_{x0}, \Pi'_{z0}]'$ , and  $\mathcal{P}_{it} = [p'_{x,it}, p'_{z,it}]'$ , we have  $\hat{\Pi} - \Pi_0 = O_p([N(T-1)]^{-1/2})$  under Assumptions P2(b) and P5(b) and

$$\begin{aligned}
\|\Delta v - \Delta \hat{v}\|^2 &= (\hat{\Pi} - \Pi_0)' \sum_{i=1}^N \sum_{t=2}^T \Delta \mathcal{P}_{it} \Delta \mathcal{P}'_{it} (\hat{\Pi} - \Pi_0) \\
&\leq N(T-1) \left\| \hat{\Pi} - \Pi \right\|^2 \lambda_{\max} \left( \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T \Delta \mathcal{P}_{it} \Delta \mathcal{P}'_{it} \right) \\
&= O_p(1)
\end{aligned} \tag{B.2}$$

under Assumption P2(b).

Next, we consider

$$\begin{aligned}
\mathcal{S}_2 &= \frac{\delta_0^2}{N(T-1)} [\chi(\gamma_0) - \chi(\gamma)]' [P^*(\gamma) - P(\gamma)] [\chi(\gamma_0) - \chi(\gamma)] + \frac{1}{N(T-1)} \Delta \varepsilon' [P^*(\gamma) - P(\gamma)] \Delta \varepsilon \\
&\quad + \frac{2\delta_0}{N(T-1)} [\chi(\gamma_0) - \chi(\gamma)]' [P^*(\gamma) - P(\gamma)] \Delta \varepsilon, \\
&= S_{21} + S_{22} + 2S_{23}
\end{aligned} \tag{B.3}$$

where  $\Delta \varepsilon = [\Delta \varepsilon_{12}, \dots, \Delta \varepsilon_{1T}, \dots, \Delta \varepsilon_{N2}, \dots, \Delta \varepsilon_{NT}]'$ . First, we establish  $S_{21}$ . Under Assumption P2(a), we have  $\frac{1}{N(T-1)} \|\chi(\gamma_0) - \chi(\gamma)\|^2 = O_p(1)$ . Then, by Lemma-1 and Assumption P2, we have

$$\begin{aligned}
S_{21} &= \frac{\delta_0^2}{N(T-1)} \|[\chi(\gamma_0) - \chi(\gamma)]' [P^*(\gamma) - P(\gamma)] [\chi(\gamma_0) - \chi(\gamma)]\|_{sp} \\
&\leq \frac{\delta_0^2}{N(T-1)} \|(\chi(\gamma_0) - \chi(\gamma))\|_{sp}^2 \|P^*(\gamma) - P(\gamma)\|_{sp} \\
&= o_p(1)
\end{aligned} \tag{B.4}$$

where  $\|A\|_{sp}$  is the spectral norm of a square matrix  $A$  and  $\|A\|_{sp} = \lambda_{max}^{1/2}(A'A)$ . Then by Lemma-1 and closely following the proof of (B.4), we can show that  $\mathcal{S}_{22}$  and  $\mathcal{S}_{23}$  are also  $o_p(1)$ .

Last, by simple calculation, we can express  $\mathcal{S}_3$  as

$$\begin{aligned}
\mathcal{S}_3 &= \frac{2\beta_{10}\delta_0}{N(T-1)} (\Delta v - \Delta \hat{v})' [I_{NT} - P(\gamma)] [\chi(\gamma_0) - \chi(\gamma) + \Delta \varepsilon] \\
&= \frac{2\beta_{10}\delta_0}{N(T-1)} (\Delta v - \Delta \hat{v})' [\chi(\gamma_0) - \chi(\gamma) + \Delta \varepsilon] - \frac{2\beta_{10}\delta_0}{N(T-1)} (\Delta v - \Delta \hat{v})' P(\gamma) [\chi(\gamma_0) - \chi(\gamma) + \Delta \varepsilon] \\
&= \mathcal{S}_{31} - 2\mathcal{S}_{32}
\end{aligned} \tag{B.5}$$



where  $\mathcal{S}_{31} = o_p(1)$  under Assumption P2 and

$$\begin{aligned}
\mathcal{S}_{32} &= \frac{2\beta_{10}\delta_0}{N(T-1)}(\Delta v - \Delta \hat{v})'P(\gamma)[\chi(\gamma_0) - \chi(\gamma) + \Delta \varepsilon] \\
&\leq 2\beta_{10}\delta_0 \max_{\gamma \in \Gamma} \left\| \frac{1}{N(T-1)}(\Delta v - \Delta \hat{v})' \Delta x(\gamma) \right\| \max_{\gamma \in \Gamma} \left\| \left[ \frac{1}{N(T-1)} \Delta x(\gamma)' \Delta x(\gamma) \right]^{-1} \right\| \\
&\quad \times \max_{\gamma \in \Gamma} \left\| \frac{1}{N(T-1)} \Delta x(\gamma)' [\chi(\gamma_0) - \chi(\gamma) + \Delta \varepsilon] \right\| \\
&= o_p(1)O_p(1)o_p(1)
\end{aligned} \tag{B.6}$$

by (C.4) in Lemma-1 and Assumption P2.

To sum up, we have  $\sup_{\phi \in B \times \Gamma} |S_{NT}(\phi) - S_{NT}^*(\phi)| \xrightarrow{p} 0$ .

Next, we show that the four conditions required by Section 4.3 Andrews (1994) also hold in our case. Denote  $\frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T M_{it}(\phi)$  as the first order partial derivative of  $S_{NT}^*(\phi)$  and the minimizer of  $S_{NT}^*(\phi)$  is given by solving  $\frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T M_{it}(\phi) = 0$ . By definition,  $M_{it}(\phi) = \Delta H_{it}(\phi) \Delta \varepsilon_{it}(\phi)$ , and  $\Delta \varepsilon_{it}(\phi) = \Delta y_{it} - \Delta x_{it}^*(\gamma) \theta$ . Under Assumption P7, the true threshold value  $\gamma_0$  is the unique minimizer of  $L^*(\theta^*(\gamma), \gamma)$ . And, by Assumption P3,  $\theta^*(\gamma)$  is uniquely defined for all  $\gamma \in \Gamma$ , where  $\Gamma$  is a compact set. Combining Assumptions T3 and T7, we have the true parameter  $\phi_0$  minimizes  $L^*(\phi) = L^*(\theta, \gamma)$  and is the unique solution of  $M(\phi_0) = E[M_{it}(\phi_0)] = 0$ , where  $L^*(\phi) = E[S_{NT}^*(\phi)]$ . Following we establish the four conditions one by one.

**Condition 1:**  $\hat{\phi} \xrightarrow{p} \phi_0$ .

For the kink regression model,  $\Delta x_{it}^*(\gamma)$  is continuous in  $\gamma$  and we have  $\Delta \varepsilon_{it}(\phi) = \Delta y_{it} - \Delta x_{it}^*(\gamma) \theta$ . Therefore,  $\Delta \varepsilon_{it}(\phi)$  and  $\Delta \varepsilon_{it}(\phi)^2$  are continuous in  $\phi$ . Using the Cauchy-Schwarz inequality, we have,

$$\Delta \varepsilon_{it}^2(\phi) \leq 2\Delta y_{it}^2 + 2|x_{it}^*(\gamma)\theta|^2 \leq 2\Delta y_{it}^2 + 2\bar{\theta}^2 \|\Delta x_{it}^*(\gamma)\|^2, \tag{B.7}$$

where  $\bar{\theta} = \sup\{\|\theta\| : \theta \in B\}$  and is bounded under Assumption P6. Recall the definition of  $\Delta x_{it}^*(\gamma)$ , under Assumptions P2 and P7, we have the finite bound  $\|\Delta x_{it}^*(\gamma)\|^2 \leq \Delta x_{it}^2 +$

$2(x_{it} - \gamma)^2 + \Delta z'_{it} \Delta z_{it} + \Delta v'_{it} \Delta v_{it}$ . Thus, for  $\phi \in B \times \Gamma$ ,  $E[\Delta \varepsilon_{it}^2(\phi)] = O(1)$ . Applying Lemma 2.4 of Newey and Mcfadden (1994), we can show  $\sup_{\phi \in B \times \Gamma} |S_{NT}^*(\phi) - L^*(\phi)| \xrightarrow{p} 0$  as  $NT \rightarrow \infty$ , where  $L^*(\phi) = E[S_{NT}^*(\phi)]$ .

Finally, by Assumptions P3, P6, and P7,  $B \times \Gamma$  is compact and  $\phi_0$  is the unique minimizer of  $L^*(\phi)$ . Thus, we conclude the proof of Condition 1 by applying Theorem 2.1 of Newey and Mcfadden (1994).

**Condition 2:**  $\frac{1}{\sqrt{N(T-1)}} \sum_{i=1}^N \sum_{t=2}^T \Delta H_{it} \Delta \varepsilon_{it} \xrightarrow{d} N(0, \mathcal{S})$ .

Following Herrndorf (1984), we complete the proof for Condition 2 by applying the CLT for the strong mixing process under Assumptions P1 and P2.

**Condition 3:**  $\mathcal{Q}(\phi)$  is continuous in  $\phi$  for  $\phi \in B \times \Gamma$  and  $\mathcal{Q}(\phi_0) = \mathcal{Q}$ , where

$$\begin{aligned} \mathcal{Q}(\phi) = & -\frac{\partial}{\partial \phi'} E[\Delta H_{it}(\phi) \Delta \varepsilon_{it}] = E[\Delta H_{it}(\phi) \Delta H'_{it}(\phi)] \\ & + E \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta \varepsilon_{it}(\phi)[I(x_t \geq \gamma) - I(x_{i,t-1} \geq \gamma)] \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta \varepsilon_{it}(\phi)[I(x_t \geq \gamma) - I(x_{i,t-1} \geq \gamma)] & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Obviously,  $\mathcal{Q}(\phi)$  is continuous w.r.t.  $\theta$ . For  $\gamma$ , note that the parameter  $\gamma$  enters  $\mathcal{Q}(\phi)$  through one of the following forms:  $W_{i,t-D_1} I(x_{i,t-D_2} \geq \gamma)$ ,  $x_{i,t-D_1} I(x_{it} \geq \gamma) I(x_{i,t-1} \geq \gamma)$ , or  $I(x_{it} \geq \gamma) I(x_{i,t-1} \geq \gamma)$ , where  $D_j \in \{0, 1\}$ ,  $j = 1, 2$  and  $W_{i,t-D_j}$  can be any vector whose elements are pairwise products of  $(1, y_{it}, y_{i,t-1}, x_{it}, x_{i,t-1}, z_{it}, z_{i,t-1}, v_{it}, v_{i,t-1})$ . By Assumptions P2, P4 and applying the law of iterated expectation, we can show that  $\mathcal{Q}(\phi)$  is also continuous w.r.t  $\gamma$ . By definition, the second term of  $\mathcal{Q}(\phi)$  equals zero when we evaluate it at  $\phi = \phi_0$ , which implies  $\mathcal{Q}(\phi_0) = \mathcal{Q}$ . Hence, we conclude our proof for Condition 3.

**Condition 4:**  $g_{NT}(\phi) = \frac{1}{\sqrt{N(T-1)}} \sum_{i=1}^N \sum_{t=2}^T (M_{it}(\phi) - M(\phi))$  is stochastic equicontinuous, where  $M_{it}(\phi) = \Delta H_{it}(\phi) \Delta \varepsilon_{it}(\phi)$  and  $M(\phi) = E[M_{it}(\phi)]$ .

Note that  $\theta$  enters  $M_{it}(\phi)$  with linear or quadratic forms. Therefore, we only need to establish the stochastic equicontinuity w.r.t.  $\gamma$ . Hence, without loss of generality, we temporarily ignore  $\theta$  and focus on  $\gamma$  in  $M_{it}(\phi)$  and write  $M_{it}(\phi)$  as  $M_{it}(\gamma)$ . Then, we have

$$\begin{aligned} M_{it}(\gamma) &= \Delta H_{it}(\gamma) \Delta \varepsilon_{it}(\gamma) \\ &= [\Delta x_{it}^*(\gamma)(\Delta y_{it} - \Delta x_{it}^{*'}(\gamma)\theta), \delta[I(x_{it} \geq \gamma) - I(x_{i,t-1} \geq \gamma)(\Delta y_{it} - \Delta x_{it}^{*'}(\gamma)\theta)]. \end{aligned} \quad (\text{B.8})$$

Note that the parameter  $\gamma$  enters  $M_{it}(\gamma)$  through one of the following forms:  $W_{i,t-D_1}I(x_{i,t-D_2} \geq \gamma)$ ,  $x_{i,t-D_1}I(x_{it} \geq \gamma)I(x_{i,t-1} \geq \gamma)$ , or  $I(x_{it} \geq \gamma)I(x_{i,t-1} \geq \gamma)$ , where  $D_j \in \{0, 1\}$ ,  $j = 1, 2$ , which defined in Condition 3. Then we construct a bound which helps establishing Condition 4.

We denote  $\mathbf{F}(\cdot)$  as the cumulative distribution of  $x_{it}$ . Then, for any  $\gamma_2 \geq \gamma_1$  and  $\gamma_1, \gamma_2 \in \Gamma$ , by Assumption P4 and employing Taylor expansion, we have  $\mathbf{F}(\gamma_2) - \mathbf{F}(\gamma_1) \leq \bar{f}|\gamma_2 - \gamma_1|$ . Thus, the bound equation for  $W_{i,t-D_1}I(x_{i,t-D_2} \geq \gamma)$ ,  $x_{i,t-D_1}I(x_{it} \geq \gamma)I(x_{i,t-1} \geq \gamma)$  and  $I(x_{it} \geq \gamma)I(x_{i,t-1} \geq \gamma)$  are given as follows by applying the Hölder's inequality:

$$\begin{aligned} &E|W_{i,t-D_1}I(\gamma_1 \leq x_{i,t} \leq \gamma_2)I(\gamma_1 \leq x_{i,t-1} \leq \gamma_2)|^2 \\ &\leq E(|W_{i,t-D_1}|^{2r})^{1/r}(E|I(\gamma_1 \leq x_{it} \leq \gamma_2)|)^{1/q_1}(E|I(\gamma_1 \leq x_{i,t-1} \leq \gamma_2)|)^{1/q_2} \\ &\leq \mathcal{C}(\mathbf{F}(\gamma_2) - \mathbf{F}(\gamma_1))^{(1/q_1+1/q_2)} \\ &\leq \mathcal{C}\bar{f}^{(1/q_1+1/q_2)}|\gamma_2 - \gamma_1|^{(1/q_1+1/q_2)}, \end{aligned} \quad (\text{B.9})$$

where  $\frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{r} = 1$  and  $E(\|W_{i,t-D_1}\|^{2r})^{1/r} \leq \mathcal{C}$ .

Following we establishing Condition 4 by applying Doukhan et al. (1995) Theorem 1. In our case, with martingale differences sequences, their condition (2.15) holds. Let  $\gamma_k$  be an equally spaced grid search point on  $\Gamma$ , where  $k = 1, \dots, \mathcal{N}$ . Then, for any  $\gamma \in \Gamma$ , there exist a bracket such that  $\min[M_{it}(\gamma_{k-1}), M_{it}(\gamma_k)] \leq M_{it}(\gamma) \leq \max[M_{it}(\gamma_{k-1}), M_{it}(\gamma_k)]$ .

Denote  $\mu$  as any positive number and  $\mathcal{N}(\mu) = \mu^{-2/q}$ . By the bound equation (B.9), we have  $E \|M_{it}(\gamma_k) - M_{it}(\gamma_{k-1})\|^2 \leq \mathcal{C} \bar{f}^{1/q} |\gamma_k - \gamma_{k-1}|^{1/q} \leq O(\mathcal{N}(\mu)^{-q}) = O(\mu^2)$ , where  $O(\mathcal{N}(\mu)^{-1})$  is the distance between grid points and  $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$ . Thus, we have  $\mathcal{N}(\mu)$  are the  $L^2$  bracketing numbers and  $H_2(\mu) = \ln \mathcal{N}(\mu) = |\log \mu|$  is the metric entropy for the class  $\{m_{it}(\gamma) | \gamma \in \Gamma\}$ . Combining this with Assumption P1, we can apply Theorem 1 of Doukhan et al. (1995)<sup>1</sup> to establish the stochastic equicontinuity of  $g_{NT}$  w.r.t  $\gamma$ , which finish the proof of condition 4.

As Conditions 1-4 are proved above, it is sufficient for us to establish Theorem 1-panel.

## C Lemma

**Lemma 1** *Under Assumptions P2, P3, P5(b) and P7, we have  $\max_{\gamma \in \Gamma} \|P^*(\gamma) - P(\gamma)\|_{sp} = o_p(1)$ .*

**Proof.** By triangular inequality and simple calculations, we have

$$\begin{aligned}
& \max_{\gamma \in \Gamma} \|P(\gamma) - P^*(\gamma)\|_{sp} \\
&= \max_{\gamma \in \Gamma} \left\| \Delta x(\gamma) [\Delta x(\gamma)' \Delta x(\gamma)]^{-1} \Delta x(\gamma)' - \Delta x^*(\gamma) [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \Delta x^*(\gamma)' \right\|_{sp} \\
&\leq \max_{\gamma \in \Gamma} \left\| \Delta x(\gamma) \{ [\Delta x(\gamma)' \Delta x(\gamma)]^{-1} - [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \} \Delta x(\gamma)' \right\|_{sp} \\
&\quad + \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma) - \Delta x^*(\gamma)] [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} [\Delta x(\gamma) - \Delta x^*(\gamma)]' \right\|_{sp} \\
&\quad + 2 \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma) - \Delta x^*(\gamma)] [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \Delta x^*(\gamma)' \right\|_{sp} \\
&= \mathcal{S}_{p1} + \mathcal{S}_{p2} + 2\mathcal{S}_{p3} = o_p(1).
\end{aligned} \tag{C.1}$$

First, we verify that  $\mathcal{S}_{p2}$  and  $\mathcal{S}_{p3}$  are  $o_p(1)$ . Under Assumptions P2, P3 and P7, by equation

---

<sup>1</sup>Note that, by collecting the bounded condition we show above and the Assumption P1, condition (2.15) in Doukhan et al. (1995) is satisfied. Thus, Theorem 1 of Doukhan et al. (1995) holds here.

(B.2), we have

$$\begin{aligned}
\mathcal{S}_{p2} &= \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma) - \Delta x^*(\gamma)] [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} [\Delta x(\gamma) - \Delta x^*(\gamma)]' \right\|_{sp} \\
&\leq \frac{1}{N(T-1)} \max_{\gamma \in \Gamma} \|\Delta x(\gamma) - \Delta x^*(\gamma)\| \max_{\gamma \in \Gamma} \left\| \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) \right]^{-1} \right\|_{sp} \\
&\quad \times \max_{\gamma \in \Gamma} \|\Delta x(\gamma) - \Delta x^*(\gamma)\| \\
&= \frac{1}{N(T-1)} \|\Delta v - \Delta \hat{v}\|^2 \max_{\gamma \in \Gamma} \left\| \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) \right]^{-1} \right\|_{sp} \\
&= O_p([N(T-1)]^{-1}) O_p(1) = o_p(1),
\end{aligned} \tag{C.2}$$

and

$$\begin{aligned}
\mathcal{S}_{p3} &= \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma) - \Delta x^*(\gamma)] [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \Delta x^*(\gamma)' \right\|_{sp} \\
&\leq \|\Delta \hat{v} - \Delta v\| \max_{\gamma \in \Gamma} \left\| \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) \right]^{-1} \right\|_{sp} \max_{\gamma \in \Gamma} \frac{1}{N(T-1)} \|\Delta x^*(\gamma)\| \\
&= O_p(1) O_p(1) o_p(1) = o_p(1),
\end{aligned} \tag{C.3}$$

where we have

$$\begin{aligned}
&\left\| \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) \right]^{-1} \right\|_{sp} \\
&= \left\| \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) \right]^{-1} \right\|_{sp} = \lambda_{min}^{-1} \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) \right] \\
&= \lambda_{min}^{-1} (E[\Delta x_{it}^*(\gamma) \Delta x_{it}^{*'}(\gamma)]) + O_p \left( \left\| \frac{1}{N(T-1)} \Delta x^*(\gamma)' \Delta x^*(\gamma) - E[\Delta x_{it}^*(\gamma) \Delta x_{it}^{*'}(\gamma)] \right\| \right) \\
&= O_p(1).
\end{aligned} \tag{C.4}$$

by Assumption P3 and applying Weyl's theorem in Seber (2008).  $\lambda_{min}(A)$  denotes the smallest eigenvalue of a symmetric matrix  $A$ . Next, we show  $\mathcal{S}_{p1}$  is  $o_p(1)$ . Under Assumption P2, we have

$$\begin{aligned}
\mathcal{S}_{p1} &= \max_{\gamma \in \Gamma} \left\| \Delta x(\gamma) \left\{ [\Delta x(\gamma)' \Delta x(\gamma)]^{-1} - [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \right\} \Delta x(\gamma)' \right\|_{sp} \\
&\leq \max_{\gamma \in \Gamma} \frac{1}{N(T-1)} \|\Delta x(\gamma)\|^2 N(T-1) \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma)' \Delta x(\gamma)]^{-1} - [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \right\|_{sp} \\
&= O_p(1) o_p(1)
\end{aligned} \tag{C.5}$$

whereby (C.4) and closely following the proof of Theorem 1-time series (A.11), (A.12) and (A.13), we obtain

$$\begin{aligned}
& N(T-1) \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma)' \Delta x(\gamma)]^{-1} - [\Delta x^*(\gamma)' \Delta x^*(\gamma)]^{-1} \right\|_{sp} \\
&= N(T-1) \max_{\gamma \in \Gamma} \left\| [\Delta x(\gamma)' \Delta x(\gamma)]^{-1} \{ [\Delta x(\gamma)' \Delta x(\gamma)] - [\Delta x^*(\gamma)' \Delta x^*(\gamma)] \} [\Delta x^*(\gamma) \Delta x^*(\gamma)]^{-1} \right\|_{sp} \\
&\leq \max_{\gamma \in \Gamma} \left\| \left[ \frac{1}{N(T-1)} \Delta x(\gamma)' \Delta x(\gamma) \right]^{-1} \right\|_{sp} \max_{\gamma \in \Gamma} \frac{1}{N(T-1)} \left\| [\Delta x(\gamma)' \Delta x(\gamma)] - [\Delta x^*(\gamma)' \Delta x^*(\gamma)] \right\| \\
&\quad \times \max_{\gamma \in \Gamma} \left\| \left[ \frac{1}{N(T-1)} \Delta x^*(\gamma) \Delta x^*(\gamma) \right]^{-1} \right\|_{sp} \\
&= O_p(1) o_p(1) O_p(1) = o_p(1)
\end{aligned} \tag{C.6}$$

Given  $\mathcal{S}_{p1}$ ,  $\mathcal{S}_{p2}$  and  $\mathcal{S}_{p3}$  are both  $o_p(1)$ , we have  $\max_{\gamma \in \Gamma} \|P^*(\gamma) - P(\gamma)\|_{sp} = o_p(1)$ .



## D Monte Carlo simulation results for DGP2 and DGP4

Table 1: Estimation Results for DGP2

TABLE 1: Estimation Results for DGP 2								
	MEAN- $\beta_1$	RMSE- $\beta_1$	MEAN- $\delta$	RMSE- $\delta$	MEAN- $\beta_3$	RMSE- $\beta_3$	MEAN- $\gamma$	RMSE- $\gamma$
	$(\beta_{10} = 1)$		$(\delta_0 = 1)$		$(\beta_{30} = 1)$		$(\gamma_0 = 2)$	
$\kappa = 0.05$	NO CF							
T=100	1.025	0.0263	0.9999	0.0134	1.025	0.0253	2	0.0072
T=200	1.0252	0.0257	0.9997	0.0092	1.025	0.0251	2	0
T=300	1.025	0.0254	1.0001	0.0076	1.025	0.0251	2	0
T=400	1.025	0.0253	1.0001	0.0064	1.025	0.0251	2	0
$\kappa = 0.05$	CF							
T=100	0.9973	0.0158	1	0.0036	0.9973	0.0144	2	0
T=200	0.9986	0.0092	1	0.0026	0.9987	0.0089	2	0
T=300	0.9991	0.007	1	0.0021	0.9989	0.0071	2	0
T=400	0.9993	0.0059	1	0.0018	0.9992	0.006	2	0
$\kappa = 0.5$	NO CF							
T=100	1.2398	0.2639	1.0225	0.1387	1.2498	0.2529	2.0055	0.2354
T=200	1.2458	0.256	1.0061	0.0934	1.2498	0.2514	1.9966	0.1509
T=300	1.2475	0.2536	1.0061	0.0764	1.2502	0.2513	2.0004	0.1139
T=400	1.2475	0.2521	1.0046	0.0648	1.2502	0.251	1.9991	0.1003
$\kappa = 0.5$	CF							
T=100	0.9718	0.1592	1.0011	0.037	0.9734	0.1439	1.9995	0.0581
T=200	0.9862	0.0929	1.0001	0.0261	0.9871	0.0888	2.0002	0.0392
T=300	0.9902	0.0709	1.0004	0.0209	0.9889	0.0709	1.9994	0.0283
T=400	0.9929	0.0593	1.0001	0.0181	0.9923	0.0599	2.0001	0.0205

NOTE: CF=Control function approach



Table 2: Estimation Results for DGP2(continue)

	MEAN- $\beta_1$ ( $\beta_{10} = 1$ )	RMSE- $\beta_1$	MEAN- $\delta$ ( $\delta_0 = 1$ )	RMSE- $\delta$	MEAN- $\beta_3$ ( $\beta_{30} = 1$ )	RMSE- $\beta_3$	MEAN- $\gamma$ ( $\gamma_0 = 2$ )	RMSE- $\gamma$
$\kappa = 0.95$ NO CF								
T=100	1.4286	0.4884	1.0985	0.291	1.4745	0.4805	2.0119	0.496
T=200	1.4544	0.4791	1.0371	0.185	1.4747	0.4777	1.995	0.3374
T=300	1.4639	0.4776	1.026	0.1481	1.4754	0.4774	2.0036	0.2525
T=400	1.4646	0.4744	1.0184	0.1241	1.4754	0.4769	1.9957	0.2076
$\kappa = 0.95$ CF								
T=100	0.9456	0.3031	1.0045	0.0702	0.9494	0.2735	2.0004	0.1086
T=200	0.9733	0.177	1.0015	0.0497	0.9756	0.1687	2.0009	0.0756
T=300	0.9809	0.1354	1.0017	0.0398	0.979	0.1347	1.9988	0.063
T=400	0.9861	0.1131	1.0008	0.0344	0.9853	0.1138	2.0004	0.0544
$\kappa = 2$ NO CF								
T=100	1.8606	1.0037	1.2899	0.7563	1.9988	1.0113	2.0196	0.7939
T=200	1.905	0.9702	1.1844	0.4575	1.9992	1.0055	1.9949	0.6813
T=300	1.9323	0.9709	1.1381	0.3534	2.0009	1.005	1.995	0.5806
T=400	1.9453	0.9746	1.1051	0.292	2.0008	1.0039	1.9941	0.5139
$\kappa = 2$ CF								
T=100	0.8766	0.6408	1.0267	0.1509	0.8935	0.576	2	0.2634
T=200	0.9411	0.3734	1.01	0.1048	0.9485	0.3551	2.0037	0.1688
T=300	0.958	0.2849	1.008	0.0839	0.9557	0.2836	1.9983	0.1342
T=400	0.969	0.2384	1.0049	0.0726	0.9691	0.2396	2.0002	0.1122

NOTE: CF=Control function approach

Table 3: Estimation Results for DGP4

		MEAN- $\beta_1$	RMSE- $\beta_1$	MEAN- $\delta$	RMSE- $\delta$	MEAN- $\beta_3$	RMSE- $\beta_3$	MEAN- $\gamma$	RMSE- $\gamma$
		$(\beta_{10} = 1)$		$(\delta_0 = 1)$		$(\beta_{30} = 1)$		$(\gamma_0 = 2)$	
NO	CF	FD							
T=10	N=10	1.7416	0.9354	1.3258	0.8619	1.9018	0.9184	1.9902	0.8476
	N=20	1.7905	0.874	1.2138	0.4971	1.9009	0.9087	1.9916	0.7119
	N=30	1.8294	0.8789	1.149	0.3836	1.9001	0.9049	2.0128	0.6166
	N=40	1.8449	0.8812	1.1172	0.3118	1.8986	0.9024	2.0096	0.5386
T=20	N=10	1.7918	0.8753	1.208	0.4986	1.9042	0.9118	1.9933	0.7021
	N=20	1.8465	0.8823	1.119	0.3122	1.9018	0.9057	2.0081	0.5382
	N=30	1.8651	0.8883	1.0694	0.2445	1.8996	0.9021	1.9976	0.4546
	N=40	1.8752	0.89	1.0503	0.1984	1.9001	0.9019	2.0002	0.3741
T=30	N=10	1.823	0.8713	1.1592	0.3758	1.9019	0.9069	2.0049	0.6126
	N=20	1.8632	0.8854	1.0746	0.2411	1.9005	0.9028	2.0009	0.4428
	N=30	1.8758	0.8896	1.0469	0.1885	1.9007	0.9024	1.998	0.3572
	N=40	1.8851	0.8947	1.0329	0.1637	1.8994	0.9006	2.0015	0.2947
T=40	N=10	1.8471	0.8811	1.1107	0.3088	1.9008	0.9045	2.0049	0.5353
	N=20	1.8685	0.8845	1.0586	0.2051	1.8992	0.9011	1.993	0.3747
	N=30	1.8835	0.893	1.0318	0.1602	1.9003	0.9015	1.9987	0.2959
	N=40	1.8898	0.8962	1.0213	0.1372	1.9003	0.9012	1.9985	0.2416

NOTE: FD=first difference; CF=control function approach

		MEAN- $\beta_1$	RMSE- $\beta_1$	MEAN- $\delta$	RMSE- $\delta$	MEAN- $\beta_3$	RMSE- $\beta_3$	MEAN- $\gamma$	RMSE- $\gamma$
		$(\beta_{10} = 1)$		$(\delta_0 = 1)$		$(\beta_{30} = 1)$		$(\gamma_0 = 2)$	
CF	FD								
T=10	N=10	0.9636	0.2427	1.0296	0.1492	0.9752	0.2144	2.0028	0.2642
	N=20	0.9831	0.1608	1.0093	0.098	0.9922	0.1437	1.9959	0.1591
	N=30	0.9875	0.1327	1.0035	0.0802	0.9919	0.1168	1.9979	0.1253
	N=40	0.9942	0.1113	1.0045	0.0671	0.9938	0.0993	2.0006	0.106
T=20	N=10	0.9845	0.1584	1.0072	0.0875	0.9895	0.1459	1.9958	0.1443
	N=20	0.9948	0.1103	1.0026	0.0615	0.9961	0.1015	2.0018	0.0959
	N=30	0.9964	0.0878	1.0017	0.0499	0.9976	0.0815	1.9994	0.0773
	N=40	0.9986	0.0764	1.0016	0.0426	0.9987	0.07	2.0007	0.0689
T=30	N=10	0.9918	0.1271	1.0031	0.0687	0.9929	0.1185	2.0003	0.1061
	N=20	0.997	0.0879	1.0019	0.0479	0.9972	0.0832	2.0004	0.075
	N=30	0.9967	0.0706	1.0009	0.039	0.9982	0.0667	1.9999	0.0629
	N=40	0.9975	0.0612	1.0011	0.0337	0.9967	0.0585	1.9985	0.0541
T=40	N=10	0.9926	0.1086	1.0026	0.0579	0.9962	0.1014	2.0002	0.0894
	N=20	0.9978	0.0752	1.0008	0.0398	0.9987	0.0711	2.0013	0.0647
	N=30	0.9996	0.0612	1.0011	0.0333	0.998	0.0573	2.001	0.0516
	N=40	0.9982	0.0536	1.0011	0.0287	0.9995	0.0493	2.0011	0.0455

NOTE: FD=first difference; CF=control function approach

## References

- Andrews, D. W. (1994). Chapter 37 empirical process methods in econometrics. *Handbook of Econometrics*, 2247–2294.
- Doukhan, P., P. Massart, and E. Rio (1995). Invariance principles for absolutely regular empirical processes. *Annales de l'I.H.P. Probabilités et statistiques* 31(2), 393–427.
- Hansen, B. E. (2017). Regression kink with an unknown threshold. *Journal of Business & Economic Statistics* 35(2), 228–240.
- Herrndorf, N. (1984). A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables. *The Annals of Probability* 12(1), 141 – 153.
- Newey, W. K. and D. Mcfadden (1994). Chapter 36 large sample estimation and hypothesis testing. *Handbook of Econometrics*, 2111–2245.
- Seber, G. A. F. (2008). *A matrix handbook for statisticians*. J. Wiley.