Supplementary Material to "Endogeneity in Semiparametric Threshold Regression"*

Andros Kourtellos† **Thanasis Stengos**‡ **Yiguo Sun**§

April 24, 2021

Abstract

This paper estimates semiparametric threshold regression models with an endogenous threshold variable. Assuming diminishing threshold effects, we derive the consistency and limiting distribution of our proposed estimator constructed from the series approximation method for weakly dependent data. In addition, we propose a test for the endogeneity of the threshold variable, which is valid regardless of whether the threshold effects exist. We assess the performance of our methods using Monte Carlo simulations.

Keywords: control function, series estimation, threshold regression.

JEL Classification Codes: C14, C24, C51

---

*The first author acknowledges that this research has received funding from Marie Skłodowska-Curie Actions (Work Programme 2016-17) of the Excellence Science Pillar of the Horizon 2020 Research and Innovation Programme of the European Union under REA grant agreement No. 707990. We would also like to thank participants at the 69th European Meeting of the Econometric Society (ESEM 2017) and 26th Midwest Econometric Conference (MEG 2017).

†Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus, email: `andros@ucy.ac.cy`.

‡Department of Economics and Finance, University of Guelph, Guelph, Ontario N1G 2W1, Canada, email: `tstengos@uoguelph.ca`.

§Department of Economics and Finance, University of Guelph, Guelph, Ontario N1G 2W1, Canada, email: `yisun@uoguelph.ca`.

1 Supplementary Proofs

Lemma B.1 *Under Assumptions 1,2, and 3(i), we have*

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\left\|n^{-1}\sum_{t=1}^{n}\mathbf{x}_{t,\gamma}\mathbf{x}_t' I\left(q_t\leq\gamma_0\right)\boldsymbol{\delta}_0-\mathbf{g}_1\left(\gamma\right)\boldsymbol{\delta}_0\right\| = o_p\left(1\right), \tag{B.1}$$

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\left\|n^{-1}\sum_{t=1}^{n}\mathbf{x}_{t,\gamma}\eta_0\left(v_t\right)I\left(q_t\leq\gamma_0\right)-\mathbf{g}_2\left(\gamma\right)\right\| = o_p\left(1\right), \tag{B.2}$$

where $\mathbf{g}_1(\gamma) = E\left[\mathbf{x}_{t,\gamma}^{*}\mathbf{x}_t' I\left(q_t\leq\gamma_0\right)\right]$ and $\mathbf{g}_2(\gamma) = E\left[\mathbf{x}_{t,\gamma}^{*}\eta_0\left(v_t\right)I\left(q_t\leq\gamma_0\right)\right]$.

Proof: Applying Lemma 1 in Hansen (1996), we obtain

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\left\|n^{-1}\sum_{t=1}^{n}\boldsymbol{\omega}_t\mathbf{x}_t' I\left(q_t\leq\gamma_0\right)\boldsymbol{\delta}_0-E\left[\boldsymbol{\omega}_t\mathbf{x}_t' I\left(q_t\leq\gamma_0\right)\right]\boldsymbol{\delta}_0\right\| = o_p\left(1\right)$$

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\left\|n^{-1}\sum_{t=1}^{n}\boldsymbol{\omega}_t\eta_0\left(v_t\right)I\left(q_t\leq\gamma_0\right)-E\left[\boldsymbol{\omega}_t\eta_0\left(v_t\right)I\left(q_t\leq\gamma_0\right)\right]\right\| = o_p\left(1\right)$$

uniformly over $\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]$ for $\boldsymbol{\omega}_t = \mathbf{x}_t I\left(q_t\leq\gamma\right)$ and $\mathbf{x}_t I\left(q_t>\gamma\right)$. As $\mathbf{x}_{t,\gamma}$ is of dimension $2\left(d_x+L_n\right)$ with an increasing L_n as the sample size increases, we will extend Lemma 1 in Hansen (1996, p.428) for the finite dimension case to the infinite dimension case. Below, we only show (B.1) as (B.2) can be proved in the same way.

Denote an $L_n\times d_x$ matrix $\boldsymbol{\Delta}_n^{-}(\gamma) = n^{-1}\sum_{t=1}^{n}\boldsymbol{\Phi}_{L_n}\left(\hat{v}_t\right)\mathbf{x}_t' I\left(q_t\leq\gamma\right)I\left(q_t\leq\gamma_0\right)$ and its jth column vector equals

$$\boldsymbol{\Delta}_{n,j}^{-}(\gamma) = n^{-1}\sum_{t=1}^{n}\boldsymbol{\Phi}_{L_n}\left(\hat{v}_t\right)x_{j,t}I\left(q_t\leq\min\left(\gamma_0,\gamma\right)\right) = \mathbf{A}_{n1}(\gamma)+\mathbf{A}_{n2}(\gamma)$$

where

$$\mathbf{A}_{n1}(\gamma) = n^{-1}\sum_{t=1}^{n}\boldsymbol{\Phi}_{L_n}\left(v_t\right)x_{j,t}I\left(q_t\leq\min\left(\gamma_0,\gamma\right)\right)$$

and

$$\mathbf{A}_{n2}(\gamma) = n^{-1}\sum_{t=1}^{n}\left[\boldsymbol{\Phi}_{L_n}\left(\hat{v}_t\right)-\boldsymbol{\Phi}_{L_n}\left(v_t\right)\right]x_{j,t}I\left(q_t\leq\min\left(\gamma_0,\gamma\right)\right).$$

For a given $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$, we have

$$\begin{aligned}
\left\|\mathbf{A}_{n2}(\gamma)\right\| &\leq n^{-1}\sum_{t=1}^{n}\left\|\mathbf{\Phi}_{L_n}(\bar{v}_t)\,\mathbf{z}_t'\left(\hat{\boldsymbol{\pi}}_q-\boldsymbol{\pi}_q\right)x_{j,t}\right\| \\
&\leq n^{-1}\sum_{t=1}^{n}\left\|\mathbf{\Phi}_{L_n}\right\|_1\left\|\hat{\boldsymbol{\pi}}_q-\boldsymbol{\pi}_q\right\|\left\|\mathbf{z}_t x_{j,t}\right\| \\
&= O_p\left(\left\|\mathbf{\Phi}_{L_n}\right\|_1/\sqrt{n}\right) \qquad \text{(B.3)}
\end{aligned}$$

where $\bar{v}_t$ lies between $\hat{v}_t$ and v_t.

Applying Lemma 1 in Hansen (1996, p.428), we obtain that each element of $\mathbf{A}_{n1}(\gamma)$ uniformly converges to its population mean over $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$. Denoting $\ddot{\phi}_{l,t} = \phi_l(v_t)\,x_{j,t}I\left(q_t \leq \min(\gamma_0, \gamma)\right)$, for a given $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$, we have

$$\begin{aligned}
&E\left\|\mathbf{A}_{n1}(\gamma)-E\left[\mathbf{A}_{n1}(\gamma)\right]\right\|^2 = \sum_{l=1}^{L_n}E\left\{n^{-1}\sum_{t=1}^{n}\left[\ddot{\phi}_{l,t}-E\left(\ddot{\phi}_{l,t}\right)\right]\right\}^2 \\
&= n^{-1}\sum_{l=1}^{L_n}E\left[\ddot{\phi}_{l,t}-E\left(\ddot{\phi}_{l,t}\right)\right]^2 + n^{-2}\sum_{l=1}^{L_n}\sum_{t=1}^{n}\sum_{s\neq t}E\left\{\left[\ddot{\phi}_{l,t}-E\left(\ddot{\phi}_{l,t}\right)\right]\left[\ddot{\phi}_{l,s}-E\left(\ddot{\phi}_{l,s}\right)\right]\right\} \\
&\leq n^{-1}\sum_{l=1}^{L_n}E\left(\ddot{\phi}_{l,t}^2\right) + n^{-2}\sum_{l=1}^{L_n}\sum_{t=1}^{n}\sum_{s\neq t}\rho_{|t-s|}\left[E\left(\left|\ddot{\phi}_{l,t}-E\left(\ddot{\phi}_{l,t}\right)\right|^2\right)E\left|\ddot{\phi}_{l,s}-E\left(\ddot{\phi}_{l,s}\right)\right|^2\right]^{1/2} \\
&= O\left(L_n/n\right),
\end{aligned}$$

by Assumptions 1(i) and 2(i)(v). Then, by Markov's inequality, we obtain $\mathbf{A}_{n1}(\gamma) - E\left[\mathbf{A}_{n1}(\gamma)\right] = O_p\left(\sqrt{L_n/n}\right)$.

Now, we partition the finite interval $\left[\underline{\gamma}, \overline{\gamma}\right]$ into N_ζ non-overlapping intervals with equal length $\zeta = \left(\overline{\gamma}-\underline{\gamma}\right)/N_\zeta$; i.e., $\left[\underline{\gamma}, \overline{\gamma}\right] = \cup_{k=1}^{N_\zeta-1}[\gamma_k, \gamma_{k+1}) \cup \left[\gamma_{N_\zeta}, \gamma_{N_\zeta+1}\right]$. Then, we have

$$\begin{aligned}
&\max_{\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]}\left\|\mathbf{A}_{n1}(\gamma)-E\left[\mathbf{A}_{n1}(\gamma)\right]\right\| \\
&\leq \max_{|\gamma'-\gamma|<\zeta}\left\|\left(\mathbf{A}_{n1}(\gamma)-E\left[\mathbf{A}_{n1}(\gamma)\right]\right)-\left(\mathbf{A}_{n1}(\gamma')-E\left[\mathbf{A}_{n1}(\gamma')\right]\right)\right\| \\
&+\sum_{k=1}^{N_\zeta+1}\left\|\mathbf{A}_{n1}(\gamma_k)-E\left[\mathbf{A}_{n1}(\gamma_k)\right]\right\| \\
&= O_p\left(\sqrt{L_n/n}\zeta\right)+O_p\left(N_\zeta\sqrt{L_n/n}\right) = o_p(1)
\end{aligned}$$

if we set ζ to be a small finite constant and $L_n/n = o(1)$.

Therefore, we obtain

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} |\mathbf{A}_{n1}(\gamma) - E[\mathbf{A}_{n1}(\gamma)]| = o_p(1)$$

provided that $\|\mathbf{\Phi}_{L_n}\|_1/\sqrt{n}$ and $L_n/n \to 0$ as $n \to \infty$.

Similarly, we can show that $\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} |\mathbf{\Delta}_n^+(\gamma) - E[\mathbf{\Delta}_n^+(\gamma)]| = o_p(1)$, where $\mathbf{\Delta}_n^+(\gamma) = n^{-1}\sum_{t=1}^n \mathbf{\Phi}_{L_n}(\hat{v}_t)\mathbf{x}_t' I(q_t > \gamma) I(q_t \leq \gamma_0)$. This completes the proof of this lemma.

Lemma B.2 *Under Assumptions 1,2, and 3(i), we have* $\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} \left\|\mathbf{\Sigma}_{n,\mathbf{xx}',\gamma}^{-1} - \mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp} = o_p(1)$, *where* $\mathbf{\Sigma}_{n,\mathbf{xx}',\gamma} = n^{-1}\mathbf{X}_\gamma'\mathbf{X}_\gamma$, $\mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma} = E\left(\mathbf{x}_{t,\gamma}^*\mathbf{x}_{t,\gamma}^{*\prime}\right)$, *and* $\|A\|_{sp} = \boldsymbol{\lambda}_{\max}^{1/2}(\mathbf{A}'\mathbf{A})$ *is the spectral norm of a square matrix* $\mathbf{A}$.

Proof: Applying the triangular inequality gives $\left\|\mathbf{\Sigma}_{n,\mathbf{xx}',\gamma}^{-1} - \mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp} \leq \mathbf{A}_{n1}(\gamma) + \mathbf{A}_{n2}(\gamma)$, where $\mathbf{A}_{n1}(\gamma) = \left\|\mathbf{\Sigma}_{n,\mathbf{xx}',\gamma}^{-1} - \mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp}$ and $\mathbf{A}_{n2}(\gamma) = \left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1} - \mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp}$. As $\|\cdot\|_{sp}$ is submultiplicative, we have

$$\begin{aligned}\mathbf{A}_{n2}(\gamma) &= \left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\left(\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma} - \Sigma_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right)\mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp} \\ &\leq \left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp}\left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1} - \mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right\|\left\|\mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp}\end{aligned}$$

where applying Weyl's theorem in Seber (2008) gives

$$\left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp} = \boldsymbol{\lambda}_{\min}^{-1}\left(\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right) = \boldsymbol{\lambda}_{\min}^{-1}\left(\mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right) + O\left(\left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma} - \mathbf{\Sigma}_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right\|\right). \quad \text{(B.4)}$$

Closely following the proof of Lemma B.1, we can show that

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} \left\|\mathbf{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma} - \Sigma_{\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right\| = o_p(1)$$

under Assumptions 1, 2, and 3(i). It then follows $\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} \mathbf{A}_{n2}(\gamma) = o_p(1)$ under Assumption 2(ii).

Similarly, we have

$$
\begin{aligned}
&\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} \mathbf{A}_{n1}(\gamma) \\
&\leq \max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} \left\|\boldsymbol{\Sigma}_{n,\mathbf{xx}',\gamma}^{-1}\right\|_{sp} \left\|\boldsymbol{\Sigma}_{n,\mathbf{xx}',\gamma} - \boldsymbol{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right\| \left\|\boldsymbol{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}^{-1}\right\|_{sp} \\
&= O_p\left(\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} \left\|\boldsymbol{\Sigma}_{n,\mathbf{xx}',\gamma} - \boldsymbol{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right\|\right) = O_p\left(\left\|\boldsymbol{\Phi}_{L_n}\right\|_1^2 L_n/n\right),
\end{aligned}
$$

as

$$
\left\|\boldsymbol{\Sigma}_{n,\mathbf{xx}',\gamma} - \boldsymbol{\Sigma}_{n,\mathbf{x}^*\mathbf{x}^{*\prime},\gamma}\right\|^2 = \left\|n^{-1}\sum_{t=1}^{n}\left(\mathbf{x}_{t,\gamma}^{-}\mathbf{x}_{t,\gamma}^{-\prime} - \mathbf{x}_{t,\gamma}^{-*}\mathbf{x}_{t,\gamma}^{-*\prime}\right)\right\|^2 + \left\|n^{-1}\sum_{t=1}^{n}\left(\mathbf{x}_{t,\gamma}^{+}\mathbf{x}_{t,\gamma}^{+\prime} - \mathbf{x}_{t,\gamma}^{+*}\mathbf{x}_{t,\gamma}^{+*\prime}\right)\right\|^2,
$$

where $\mathbf{x}_{t,\gamma} = \left[\mathbf{x}_{t,\gamma}^{-}, \mathbf{x}_{t,\gamma}^{+}\right]$, $\mathbf{x}_{t,\gamma}^{*} = \left[\mathbf{x}_{t,\gamma}^{*-}, \mathbf{x}_{t,\gamma}^{*+}\right]$, and

$$
\begin{aligned}
\boldsymbol{\Delta}_n(\gamma) &= \left\|n^{-1}\sum_{t=1}^{n}\left(\mathbf{x}_{t,\gamma}^{-}\mathbf{x}_{t,\gamma}^{-\prime} - \mathbf{x}_{t,\gamma}^{-*}\mathbf{x}_{t,\gamma}^{-*\prime}\right)\right\|^2 \\
&= \frac{2}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\mathbf{x}_t'\mathbf{x}_s I(q_t\leq\gamma) I(q_s\leq\gamma)\left[\boldsymbol{\Phi}_{L_n}(\hat{v}_t) - \boldsymbol{\Phi}_{L_n}(v_t)\right]'\left[\boldsymbol{\Phi}_{L_n}(\hat{v}_s) - \boldsymbol{\Phi}_{L_n}(v_s)\right] \\
&+ \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n} I(q_t\leq\gamma) I(q_s\leq\gamma)\,\mathrm{tr}\,\left\{\left[\boldsymbol{\Phi}_{L_n}(\hat{v}_t)\boldsymbol{\Phi}_{L_n}(\hat{v}_t)' - \boldsymbol{\Phi}_{L_n}(v_t)\boldsymbol{\Phi}_{L_n}(v_t)'\right]\right. \\
&\times \left.\left[\boldsymbol{\Phi}_{L_n}(\hat{v}_t)\boldsymbol{\Phi}_{L_n}(\hat{v}_s)' - \boldsymbol{\Phi}_{L_n}(v_s)\boldsymbol{\Phi}_{L_n}(v_s)'\right]\right\} \\
&= O_p\left(\left\|\boldsymbol{\Phi}_{L_n}\right\|_1^2 L_n/n\right)
\end{aligned}
$$

by Assumption 2 and $\boldsymbol{\Phi}_{L_n}(\hat{v}_t) - \boldsymbol{\Phi}_{L_n}(v_t) = \boldsymbol{\Phi}_{L_n}^{(1)}(\bar{v}_t)(\hat{v}_t - v_t) = \boldsymbol{\Phi}_{L_n}^{(1)}(\bar{v}_t)\mathbf{z}_t'(\boldsymbol{\pi}_q - \hat{\boldsymbol{\pi}}_q)$. Using the same notation as in the proof of Lemma B.1, we obtain

$$
\begin{aligned}
\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\boldsymbol{\Delta}_n(\gamma) &\leq \max_{|\gamma'-\gamma|<\zeta}\left|\boldsymbol{\Delta}_n(\gamma) - \boldsymbol{\Delta}_n(\gamma')\right| + \sum_{k=1}^{N_\zeta+1}\boldsymbol{\Delta}_n(\gamma_k) \\
&= O_p\left(\left\|\boldsymbol{\Phi}_{L_n}\right\|_1^2 L_n/n\zeta\right) + O_p\left(N_\zeta\left\|\boldsymbol{\Phi}_{L_n}\right\|_1^2 L_n/n\right) = o_p(1)
\end{aligned}
$$

if we set ζ to be a small finite number. This completes the proof of this lemma.

Lemma B.3 *Under Assumptions 1, 2, and 3(i), we have* $\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]} n^{-1}\boldsymbol{\varepsilon}'\mathbf{P}_\gamma\boldsymbol{\varepsilon} = O_p(L_n/n)$.

Proof: We have

$$\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\frac{\boldsymbol{\varepsilon}'\mathbf{P}_\gamma\boldsymbol{\varepsilon}}{n}\leq\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\boldsymbol{\lambda}_{\max}\left(\left(\frac{\mathbf{X}_\gamma'\mathbf{X}_\gamma}{n}\right)^{-1}\right)\frac{\boldsymbol{\varepsilon}'\mathbf{X}_\gamma\mathbf{X}_\gamma'\boldsymbol{\varepsilon}}{n^2}\leq ML_n/n$$

as $\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\boldsymbol{\lambda}_{\max}\left(\left(n^{-1}\mathbf{X}_\gamma'\mathbf{X}_\gamma\right)^{-1}\right)=O_p(1)$ from the proof of Lemma B.2, and under Assumption 2(i) and we can show that $\max_{\gamma\in[\underline{\gamma},\overline{\gamma}]}\boldsymbol{\varepsilon}'\mathbf{X}_\gamma\mathbf{X}_\gamma'\boldsymbol{\varepsilon}/n^2=O_p(L_n/n)$ by closely following the proof of Lemma B.2. This completes the proof of this lemma.

Lemma B.4 *Under Assumptions 1, 2 and 3(i), we have* $\left\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right\|=O_p(\vartheta_n)+o_p\left(n^{-\min(\varsigma,\varrho)}\right)$.

Proof: Denote $d_t(\gamma,\gamma_0)=I(q_t\leq\gamma)-I(q_t\leq\gamma_0)$. By definition, we have $\hat{\boldsymbol{\theta}}(\gamma)-\boldsymbol{\theta}=\left(\mathbf{X}_\gamma'\mathbf{X}_\gamma\right)^{-1}\mathbf{X}_\gamma'(\mathbf{y}-\mathbf{X}_\gamma\boldsymbol{\theta})$, where

$$\begin{aligned}y_t-\mathbf{x}_{t,\gamma}'\boldsymbol{\theta}=&h_2(v_t)-h_2^*(\hat{v}_t)+n^{-\varrho}\left[\eta_0(v_t)-\eta_0^*(\hat{v}_t)\right]I(q_t\leq\gamma)+\varepsilon_t\\&-\left[n^{-\varsigma}\boldsymbol{\delta}_0'\mathbf{x}_t+n^{-\varrho}\eta_0(v_t)\right]d_t(\gamma,\gamma_0)\end{aligned}\tag{B.5}$$

By equation (A.9), Lemmas B.2 and B.3, we show that

$$\left\|\left(\mathbf{X}_\gamma'\mathbf{X}_\gamma\right)^{-1}\mathbf{X}_\gamma'\boldsymbol{\Delta}_n\right\|\leq\left\|\left(\mathbf{X}_\gamma'\mathbf{X}_\gamma/n\right)^{-1/2}\right\|\left\|\left(\mathbf{X}_\gamma'\mathbf{X}_\gamma/n\right)^{-1/2}\mathbf{X}_\gamma'\boldsymbol{\Delta}_n/n\right\|=O_p(\vartheta_n)$$

holds uniformly over $\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]$, where we denote an $n\times1$ vector $\boldsymbol{\Delta}_n$ whose t^{th} element equal to $h_2(v_t)-h_2^*(\hat{v}_t)+n^{-\varrho}\left[\eta_0(v_t)-\eta_0^*(\hat{v}_t)\right]I(q_t\leq\gamma)+\varepsilon_t$. In addition, following the proof of Lemma B.1 and by Lemma A.1 in Hansen (2000), we have $\left\|\hat{\boldsymbol{\theta}}(\gamma)-\boldsymbol{\theta}\right\|=O_p(\vartheta_n)+O_p\left(n^{-min(\varsigma,\varrho)}|\gamma-\gamma_0|\right)$ uniformly over $\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]$. As $\hat{\gamma}-\gamma_0=o_p(1)$ by Lemma 1, we then complete the proof of this lemma.

Lemma B.5 *(i)* $R_n(\upsilon)\Rightarrow\sqrt{\tau_1^2}W_1(-\upsilon)$ *if* $\upsilon\in[\underline{\upsilon},0]$ *and* $R_n(\upsilon)\Rightarrow\sqrt{\tau_2^2}W_2(\upsilon)$ *if* $\upsilon\in[0,\overline{\upsilon}]$, *where* $W_1(\upsilon)$ *and* $W_2(\upsilon)$ *are independent standard Brownian motion process over* $[0,\infty)$, *and* $\tau_j^2=E\left\{\boldsymbol{\varepsilon}_{jt}^2\left[I(\varsigma\leq\varrho)\boldsymbol{\delta}_0'\mathbf{x}_t+I(\varsigma\geq\varrho)\eta_0(v_t)\right]^2|q_t=\gamma_0\right\}f_q(\gamma_0)$ *for j=1,2; (ii)* $G_n(\upsilon)\xrightarrow{p}\mu|\upsilon|$, *where* $\mu=E\left\{\left[I(\varsigma\leq\varrho)\boldsymbol{\delta}_0'\mathbf{x}_t+I(\varsigma\geq\varrho)\eta_0(v_t)\right]^2|q_t=\gamma_0\right\}f_q(\gamma_0)$.

Proof: We first verify (i). Denoting $\epsilon_{nt}(\upsilon)=\sqrt{n}\varepsilon_t\boldsymbol{\kappa}_n'\boldsymbol{\chi}_t^*d_t(\gamma_0+\upsilon/a_n,\gamma_0)$, we have $R_n(\upsilon)=n^{-1/2}\sum_{t=1}^n\epsilon_{nt}(\upsilon)$, where $E(\epsilon_{nt}(\upsilon))=0$ and

$$E\left(\epsilon_{nt}^2(\upsilon)\right)=\begin{cases}n^{2\min(\varsigma,\varrho)}|\upsilon|E\left[\boldsymbol{\varepsilon}_{1t}^2\left(\boldsymbol{\kappa}_n'\boldsymbol{\chi}_t^*\right)^2|q_t=\gamma_0\right]f_q(\gamma_0)[1+o(1)]=|\upsilon|\tau_1^2+o(1), & \upsilon\in[\underline{\upsilon},0]\\n^{2\min(\varsigma,\varrho)}|\upsilon|E\left[\boldsymbol{\varepsilon}_{2t}^2\left(\boldsymbol{\kappa}_n'\boldsymbol{\chi}_t^*\right)^2|q_t=\gamma_0\right]f_q(\gamma_0)[1+o(1)]=|\upsilon|\tau_2^2+o(1), & \upsilon\in[0,\overline{\upsilon}]\end{cases}$$

as $\left\{\left(\varepsilon_t, \mathcal{F}_{n,t}^*\right)\right\}$ is a martingale difference sequence under Assumption **1**(iii) as explained in Section 2. Under Assumption 1(iii), $\left\{\left(\epsilon_{nt}\left(\upsilon\right), \mathcal{F}_{n,t}^*\right)\right\}$ is a stationary ergodic martingale difference array. Applying Theorem 5.16 in White (2001), we have $R_n\left(\upsilon\right) \xrightarrow{d} N\left(0, \left|\upsilon\right| \tau_1^2\right)$ if $\upsilon \in [\underline{\upsilon}, 0]$. Therefore, $\left[R_n\left(\upsilon_1\right), \ldots, R_n\left(\upsilon_k\right)\right] \xrightarrow{d} \left[B_1\left(\upsilon_1\right), \ldots, B_1\left(\upsilon_k\right)\right]$ for any finite positive integer k, where $B_1\left(\upsilon\right)$ is normally distributed with zero mean and variance $\left|\upsilon\right| \tau_1^2$. Similarly, if $\upsilon \in [0, \overline{\upsilon}]$, we have $R_n\left(\upsilon\right) \xrightarrow{d} N\left(0, \left|\upsilon\right| \tau_2^2\right)$ and $\left[R_n\left(\upsilon_1\right), \ldots, R_n\left(\upsilon_k\right)\right] \xrightarrow{d} \left[B_2\left(\upsilon_1\right), \ldots, B_2\left(\upsilon_k\right)\right]$ for any finite positive integer k, where $B_2\left(\upsilon\right)$ is normally distributed with zero mean and variance $\left|\upsilon\right| \tau_2^2$ and is independent of $B_1\left(\upsilon\right)$. Closely following the proof of Lemma A.3 in Hansen (2000), we can show that $R_n\left(\upsilon\right)$ is tight over $\upsilon \in [\underline{\upsilon}, \overline{\upsilon}]$. This completes the proof of (i).

Now, we verify (ii). For any given $\upsilon \in [\underline{\upsilon}, \overline{\upsilon}]$, we have

$$
\begin{aligned}
E\left[G_n\left(\upsilon\right)\right] &= \sum_{t=1}^{n} E\left[\left(\boldsymbol{\kappa}_n' \boldsymbol{\chi}_t^*\right)^2 d_t^2\left(\gamma_0 + \frac{\upsilon}{a_n}, \gamma_0\right)\right] \\
&= \left|\upsilon\right| n^{2\min(\varsigma, \varrho)} E\left[\left(\boldsymbol{\kappa}_n' \boldsymbol{\chi}_t^*\right)^2 | q_t = \gamma_0\right] f_q\left(\gamma_0\right)\left[1 + o\left(1\right)\right]
\end{aligned}
$$

and denoting $\eta_{nt} = \left(\boldsymbol{\kappa}_n' \boldsymbol{\chi}_t^*\right)^2 d_t^2\left(\gamma_0 + \upsilon/a_n, \gamma_0\right)$, we have

$$
\begin{aligned}
& E\left\{G_n\left(\upsilon\right) - E\left[G_n\left(\upsilon\right)\right]\right\}^2 \\
&= nE\left(\eta_{nt} - E\eta_{nt}\right)^2 + \sum_{t=1}^{n} \sum_{s \neq t} E\left[\left(\eta_{nt} - E\eta_{nt}\right)\left(\eta_{ns} - E\eta_{ns}\right)\right] \\
&= A_{n1} + A_{n2},
\end{aligned}
$$

where

$$
\begin{aligned}
A_{n1} &\leq nE\left(\eta_{nt}^2\right) = \frac{n}{a_n} \left|\upsilon\right| E\left[\left(\boldsymbol{\kappa}_n' \boldsymbol{\chi}_t^*\right)^4 | q_t = \gamma_0\right] f_q\left(\gamma_0\right)\left[1 + o\left(1\right)\right] \\
&= O\left(n^{-2\min(\varsigma, \varrho)}\right) = o\left(1\right)
\end{aligned}
$$

and

$$
\begin{aligned}
A_{n2} &= \sum_{t=1}^{n} \sum_{s \neq t} E\left[\left(\eta_{nt} - E\eta_{nt}\right)\left(\eta_{ns} - E\eta_{ns}\right)\right] \\
&\leq \sum_{t=1}^{n} \sum_{s \neq t} \rho_{|t-s|} E\left(\eta_{nt} - E\eta_{nt}\right)^2 = O\left(a_n^{-1}\right) = o\left(1\right)
\end{aligned}
$$

by Assumption 1(i) . Hence, we obtain $G_n(\upsilon) = E[G_n(\upsilon)] + o_p(1) = |\upsilon|\mu + o_p(1)$. As $G_n(\upsilon)$ is monotonically increasing in $\upsilon \in [\underline{\upsilon}, \overline{\upsilon}]$ and the limiting function is continuous in υ, closely following the interval split method used in the proof of Lemma B.1, we can show that $G_n(\upsilon) = |\upsilon|\mu + o_p(1)$ holds uniformly over $\upsilon \in [\underline{\upsilon}, \overline{\upsilon}]$. This completes the proof of this lemma.

Lemma B.6 *Under Assumptions 1-3(i) and $\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$ and under H_0, we have*

$$\text{(i)}\ \left\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1} - \boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\right\| = O_p(c_n)\ ;$$

$$\text{(ii)} \left\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},2} - n^{-1}\sum_{t=1}^{n}\boldsymbol{\omega}_t^*\mathbf{x}_t' I(q_t \leq \gamma_0)\right\| = O_p(c_n)\ ;$$

$$\text{(iii)} \left\|\boldsymbol{\lambda}_{n,\boldsymbol{\omega}} - n^{-1}\sum_{t=1}^{n}\varepsilon_t\boldsymbol{\omega}_t^* I(q_t \leq \gamma_0)\right\| = O_p(b_n)$$

for $\boldsymbol{\omega}_t = \boldsymbol{\Phi}_{L_n}(\hat{v}_t)$ and $\mathbf{x}_t$, and $\boldsymbol{\omega}_t^* = \boldsymbol{\Phi}_{L_n}(v_t)$ and $\mathbf{x}_t$, where we denote $c_n = E[\|\boldsymbol{\omega}_1^*\mathbf{x}_1'\| \,| q_1 = \gamma_0]\, n^{-1+2\varsigma}$ and $b_n = \sqrt{E(\boldsymbol{\omega}_t^{*\prime}\boldsymbol{\omega}_t^*|q_t = \gamma_0)/n}\, n^{-1+2\varsigma}$.

Proof: We verify (i) first. It is readily seen that $\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1} - \boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\| \leq n^{-1}\sum_{t=1}^{n}\|\boldsymbol{\omega}_t\mathbf{x}_t'\|\, I(\gamma_0 - |\hat{\gamma} - \gamma_0| < q_t < \gamma_0 + |\hat{\gamma} - \gamma_0|)$. As $\hat{\gamma} = \gamma_0 + O_p(n^{-1+2\varsigma})$, for any small $\epsilon > 0$, there exist some constant M_ϵ and an integer N_ϵ such that for any $n > N_\epsilon$ we have $\Pr\{|\hat{\gamma} - \gamma_0| > M_\epsilon n^{-1+2\varsigma}\} \leq \epsilon$. We then partition the finite interval $[\gamma_0 - M_\epsilon n^{-1+2\varsigma}, \gamma_0 + M_\epsilon n^{-1+2\varsigma}]$ into N_ϵ non-overlapping intervals with equal length $\epsilon = 2M_\epsilon n^{-1+2\varsigma}/N_\epsilon$; i.e., $[\gamma_0 - M_\epsilon n^{-1+2\varsigma},\ \gamma_0 + M_\epsilon n^{-1+2\varsigma}] = \cup_{k=1}^{N_\epsilon - 1}[\gamma_k, \gamma_{k+1}) \cup [\gamma_{N\epsilon}, \gamma_{N_\epsilon+1}]$. Then, we have

$$\begin{aligned}
&\max_{\hat{\gamma}\in[\gamma_0 - M_\epsilon n^{-1+2\varsigma}, \gamma_0 + M_\epsilon n^{-1+2\varsigma}]} \left\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1} - \boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\right\| \\
&\leq \max_{|\gamma' - \gamma| < \epsilon}\left\|n^{-1}\sum_{t=1}^{n}\boldsymbol{\omega}_t\mathbf{x}_t' d_t(\gamma', \gamma_0)\right\| + \sum_{k=1}^{N_\epsilon+1}\left\|n^{-1}\sum_{t=1}^{n}\boldsymbol{\omega}_t\mathbf{x}_t' d_t(\gamma_k, \gamma_0)\right\| \\
&= O_p\left(n^{-1+2\varsigma}\sqrt{E(\|\boldsymbol{\omega}_t^*\mathbf{x}_t'\| \,| q_t = \gamma_0)}\right),
\end{aligned}$$

where we use B.3 and Markov's inequality to obtain the second inequality. It follows that for any $n > N_\epsilon$ and any finite $M_\epsilon > 0$ such that

$$
\begin{aligned}
&\Pr\left\{\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1}-\boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\| > M_\epsilon c_n\right\} \\
&= \Pr\left\{\left\{\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1}-\boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\| > M_\epsilon c_n\right\} \cap \left\{|\hat{\gamma}-\gamma_0| \le M_\epsilon n^{-1+2\varsigma}\right\}\right\} \\
&+ \Pr\left\{\left\{\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1}-\boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\| > M_\epsilon c_n\right\} \cap \left\{|\hat{\gamma}-\gamma_0| > M_\epsilon n^{-1+2\varsigma}\right\}\right\} \\
&\le \Pr\left\{\max_{|\hat{\gamma}-\gamma_0|\le M_\epsilon n^{-1+2\varsigma}} \left\| n^{-1}\sum_{t=1}^{n} \boldsymbol{\omega}_t \mathbf{x}_t' d_t(\hat{\gamma},\gamma_0)\right\| > M_\epsilon c_n\right\} + \Pr\left\{|\hat{\gamma}-\gamma_0| > M_\epsilon n^{-1+2\varsigma}\right\} \\
&\le 2\epsilon.
\end{aligned}
$$

This gives $\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},1}-\boldsymbol{\chi}_{n,\boldsymbol{\omega},2}\| = O_p(c_n)$. Similarly, we obtain $\|\boldsymbol{\chi}_{n,\boldsymbol{\omega},2} - n^{-1}\sum_{t=1}^{n}\boldsymbol{\omega}_t^* \mathbf{x}_t' I(q_t \le \gamma_0)\| = O_p(c_n)$.

Now, we verify (iii). Using the proof method above, we have

$$
\begin{aligned}
&\max_{\hat{\gamma}\in[\gamma_0 - M_\epsilon n^{-1+2\varsigma},\, \gamma_0 + M_\epsilon n^{-1+2\varsigma}]} \left\|\boldsymbol{\lambda}_{n,\boldsymbol{\omega}} - n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t I(q_t \le \gamma_0)\right\| \\
&\le \max_{|\gamma'-\gamma|<\epsilon} \left\| n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t d_t(\gamma',\gamma_0)\right\| + \sum_{k=1}^{N_\epsilon+1} \left\| n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t d_t(\gamma_k,\gamma_0)\right\| \\
&= O_p\left(n^{-1+2\varsigma}\sqrt{E\left(\varepsilon_t^2 \boldsymbol{\omega}_t^{*\prime}\boldsymbol{\omega}_t^* | q_t = \gamma_0\right)/n}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\Pr\left\{\left\|\boldsymbol{\lambda}_{n,\boldsymbol{\omega}} - n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t^* I(q_t \le \gamma_0)\right\| > M b_n\right\} \\
&= \Pr\left\{\left\{\left\|\boldsymbol{\lambda}_{n,\boldsymbol{\omega}} - n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t^* I(q_t \le \gamma_0)\right\| > M_\epsilon b_n\right\} \cap \left\{|\hat{\gamma}-\gamma_0| \le M_\epsilon n^{-1+2\varsigma}\right\}\right\} \\
&+ \Pr\left\{\left\{\left\|\boldsymbol{\lambda}_{n,\boldsymbol{\omega}} - n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t^* I(q_t \le \gamma_0)\right\| > M_\epsilon b_n\right\} \cap \left\{|\hat{\gamma}-\gamma_0| > M_\epsilon n^{-1+2\varsigma}\right\}\right\} \\
&\le \Pr\left\{\max_{|\hat{\gamma}-\gamma_0|\le M_\epsilon n^{-1+2\varsigma}} \left\| n^{-1}\sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t^* d_t(\hat{\gamma},\gamma_0)\right\| > M_\epsilon b_n\right\} \\
&+ \Pr\left\{|\hat{\gamma}-\gamma_0| > M_\epsilon n^{-1+2\varsigma}\right\} = \epsilon + \epsilon = 2\epsilon.
\end{aligned}
$$

This gives

$$\left\| \boldsymbol{\lambda}_{n,\boldsymbol{\omega}} - n^{-1} \sum_{t=1}^{n} \varepsilon_t \boldsymbol{\omega}_t^* I\left(q_t \leq \gamma_0\right) \right\| = O_p\left(b_n\right).$$

This complete the proof of this lemma.

2 Supplementary Monte Carlo Simulations

Our alternative DGP adds an endogenous regressor to model (4.1)

$$y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + (\delta_1 + \delta_2 x_{1i} + \delta_3 x_{2i}) I\{q_i \le \gamma\} + u_i, \tag{B.6}$$

where

$$x_{1i} = z_{xi} + v_{xi},$$

with

$$z_{xi} = (w x_{2i} + (1-w)\varsigma_{zi}) \,/ \sqrt{w^2 + (1-w)^2}, \tag{B.7}$$

and

$$u_i = (c_{xu} v_{xi} + c_{qu} v_{qi} + (1 - c_{xu} - c_{qu})\varsigma_{ui}) \,/ \sqrt{c_{xu}^2 + c_{qu}^2 + (1 - c_{xu} - c_{qu})^2}, \tag{B.8}$$

where x_{2i}, v_{xi}, ς_{zi} and ς_{ui} are independent $i.i.d.$ $N(0,1)$ random variables. The degree of endogeneity of the threshold variable is controlled by the correlation coefficient between u_i and v_{qi} given by $c_{qu}/\sqrt{c_{xu}^2 + c_{qu}^2 + (1 - c_{xu} - c_{qu})^2}$. Similarly, the degree of endogeneity of x_{1i} is determined by the correlation between u_i and v_{xi} given by $c_{xu}/\sqrt{c_{xu}^2 + c_{qu}^2 + (1 - c_{xu} - c_{qu})^2}$. We vary δ_3 and fix $c_{xu} = 0.45$, $w = 0.5$, $\beta_1 = \beta_2 = \beta_3 = 1$, and $\delta_1 = \delta_2 = 0$. We set c_{qu} at 0.45, which corresponds to correlation of 0.7 between q_i and u_i.

Table B.1: Threshold Parameter and Threshold Effect: $L_n = 2$

This table presents Monte Carlo quantiles of the estimates of the true threshold parameter $\gamma = 2$ and variant true threshold effects, using a 2nd order Hermite basis function and sample sizes n. Columns (2)-(7) present the results for the DGP with only an exogenous regressor in (4.1)-(4.4) and the remaining columns present the results for the DGP that also includes an endogenous regressor in (B.6)-(B.8).

	Exogenous Regressor						Endogenous Regressor					
	Threshold Parameter			Threshold Effect			Threshold Parameter			Threshold Effect		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
n			$\delta_2 = 1$						$\delta_3 = 1$			
100	1.825	1.978	2.086	0.881	0.995	1.084	0.679	1.954	3.109	0.358	0.949	1.363
250	1.920	1.992	2.057	0.924	0.999	1.057	1.120	1.982	2.648	0.711	0.976	1.180
500	1.954	1.996	2.033	0.946	0.997	1.039	1.644	1.994	2.284	0.819	0.984	1.128
1000	1.968	1.998	2.017	0.964	0.998	1.033	1.764	1.995	2.106	0.885	0.989	1.086
			$\delta_2 = 2$						$\delta_3 = 2$			
100	1.884	1.977	2.024	1.900	2.000	2.094	1.373	1.977	2.480	1.473	1.959	2.333
250	1.950	1.991	2.015	1.935	2.003	2.063	1.824	1.992	2.151	1.772	1.998	2.183
500	1.975	1.995	2.003	1.954	2.000	2.046	1.902	1.996	2.062	1.865	1.991	2.126
1000	1.989	1.998	2.003	1.968	2.000	2.035	1.957	1.998	2.030	1.903	1.995	2.087
			$\delta_2 = 3$						$\delta_3 = 3$			
100	1.886	1.977	2.001	2.902	3.001	3.092	1.709	1.980	2.212	2.626	2.977	3.311
250	1.954	1.991	2.003	2.937	3.004	3.064	1.888	1.992	2.082	2.799	3.004	3.175
500	1.977	1.996	2.001	2.954	3.001	3.046	1.946	1.996	2.032	2.867	2.993	3.123
1000	1.990	1.998	2.001	2.968	3.000	3.035	1.978	1.998	2.018	2.905	2.996	3.088
			$\delta_2 = 4$						$\delta_3 = 4$			
100	1.885	1.977	2.000	3.903	4.001	4.092	1.793	1.980	2.136	3.662	3.985	4.308
250	1.954	1.991	2.001	3.937	4.004	4.064	1.923	1.992	2.047	3.818	4.004	4.180
500	1.978	1.996	2.001	3.954	4.001	4.046	1.966	1.996	2.025	3.874	3.995	4.129
1000	1.990	1.998	2.001	3.968	4.000	4.035	1.983	1.998	2.012	3.909	3.996	4.088

Table B.2: Threshold Parameter and Threshold Effect: $L_n = 3$

This table presents Monte Carlo quantiles of the estimates of the true threshold parameter $\gamma = 2$ and variant true threshold effects, using a 3rd order Hermite basis function and sample sizes n. Columns (2)-(7) present the results for the DGP with only an exogenous regressor in (4.1)-(4.4) and the remaining columns present the results for the DGP that also includes an endogenous regressor in (B.6)-(B.8).

	Exogenous Regressor						Endogenous Regressor					
	Threshold Parameter			Threshold Effect			Threshold Parameter			Threshold Effect		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
n			$\delta_2 = 1$						$\delta_3 = 1$			
100	1.877	1.977	2.042	0.935	0.999	1.057	0.645	1.968	3.190	0.341	0.926	1.371
250	1.950	1.991	2.022	0.957	0.999	1.040	0.984	1.994	2.927	0.665	0.969	1.183
500	1.976	1.996	2.008	0.969	1.000	1.028	1.654	1.995	2.333	0.805	0.982	1.120
1000	1.989	1.998	2.005	0.978	1.000	1.022	1.829	1.996	2.140	0.887	0.991	1.087
			$\delta_2 = 2$						$\delta_3 = 2$			
100	1.885	1.976	2.001	1.935	2.000	2.057	1.350	1.984	2.556	1.439	1.965	2.326
250	1.954	1.991	2.002	1.960	2.000	2.040	1.817	1.991	2.161	1.773	1.995	2.180
500	1.978	1.996	2.001	1.969	2.000	2.028	1.911	1.996	2.068	1.870	1.992	2.118
1000	1.989	1.998	2.001	1.978	2.000	2.022	1.958	1.998	2.031	1.902	1.996	2.087
			$\delta_2 = 3$						$\delta_3 = 3$			
100	1.887	1.977	2.000	2.936	3.000	3.057	1.709	1.981	2.225	2.602	2.979	3.313
250	1.955	1.991	2.000	2.959	3.000	3.040	1.890	1.992	2.090	2.793	3.000	3.177
500	1.978	1.996	2.000	2.970	3.000	3.028	1.946	1.996	2.031	2.871	2.995	3.118
1000	1.990	1.998	2.001	2.979	3.000	3.022	1.978	1.998	2.018	2.906	2.996	3.086
			$\delta_2 = 4$						$\delta_3 = 4$			
100	1.888	1.976	1.999	3.936	4.000	4.057	1.793	1.979	2.142	3.654	3.987	4.312
250	1.955	1.991	2.000	3.959	4.000	4.040	1.920	1.992	2.049	3.816	4.000	4.179
500	1.978	1.996	2.000	3.970	4.000	4.028	1.963	1.996	2.025	3.873	3.996	4.122
1000	1.990	1.998	2.001	3.979	4.000	4.022	1.983	1.998	2.012	3.907	3.997	4.086

Table B.3: Threshold Parameter and Threshold Effect: $L_n = 4$

This table presents Monte Carlo quantiles of the estimates of the true threshold parameter $\gamma = 2$ and variant true threshold effects, using a 4th order Hermite basis function and sample sizes n. Columns (2)-(7) present the results for the DGP with only an exogenous regressor in (4.1)-(4.4) and the remaining columns present the results for the DGP that also includes an endogenous regressor in (B.6)-(B.8).

	Exogenous Regressor						Endogenous Regressor					
	Threshold Parameter			Threshold Effect			Threshold Parameter			Threshold Effect		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
n			$\delta_2 = 1$						$\delta_3 = 1$			
100	1.881	1.978	2.039	0.943	1.001	1.049	0.678	1.984	3.198	0.336	0.926	1.364
250	1.952	1.991	2.013	0.965	1.000	1.032	0.974	1.994	2.907	0.655	0.964	1.178
500	1.976	1.996	2.007	0.979	1.000	1.022	1.636	1.996	2.370	0.812	0.980	1.114
1000	1.989	1.998	2.003	0.985	1.000	1.017	1.837	1.997	2.165	0.888	0.991	1.083
			$\delta_2 = 2$						$\delta_3 = 2$			
100	1.888	1.978	2.000	1.944	2.001	2.051	1.312	1.986	2.566	1.414	1.954	2.327
250	1.954	1.991	2.000	1.968	2.001	2.033	1.824	1.992	2.159	1.784	1.995	2.180
500	1.978	1.996	2.001	1.979	2.000	2.022	1.908	1.996	2.068	1.867	1.992	2.113
1000	1.990	1.998	2.001	1.985	2.000	2.017	1.959	1.998	2.033	1.901	1.996	2.084
			$\delta_2 = 3$						$\delta_3 = 3$			
100	1.888	1.978	2.000	2.945	3.001	3.050	1.720	1.981	2.234	2.619	2.973	3.315
250	1.954	1.991	2.000	2.968	3.001	3.033	1.884	1.992	2.089	2.799	2.999	3.179
500	1.978	1.996	2.001	2.979	3.000	3.022	1.948	1.996	2.032	2.872	2.994	3.117
1000	1.990	1.998	2.001	2.985	3.000	3.017	1.978	1.998	2.018	2.902	2.997	3.085
			$\delta_2 = 4$						$\delta_3 = 4$			
100	1.892	1.977	2.000	3.945	4.001	4.050	1.800	1.981	2.141	3.666	3.984	4.314
250	1.954	1.991	2.000	3.968	4.001	4.033	1.918	1.992	2.050	3.809	4.001	4.181
500	1.978	1.996	2.000	3.979	4.000	4.022	1.962	1.996	2.025	3.874	3.994	4.121
1000	1.990	1.998	2.001	3.985	4.000	4.017	1.983	1.998	2.012	3.906	3.998	4.086

Table B.4: Threshold Parameter and Threshold Effect: $L_n = 5$

This table presents Monte Carlo quantiles of the estimates of the true threshold parameter $\gamma = 2$ and variant true threshold effects, using a 5th order Hermite basis function and sample sizes n. Columns (2)-(7) present the results for the DGP with only an exogenous regressor in (4.1)-(4.4) and the remaining columns present the results for the DGP that also includes an endogenous regressor in (B.6)-(B.8).

	Exogenous Regressor						Endogenous Regressor					
	Threshold Parameter			Threshold Effect			Threshold Parameter			Threshold Effect		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
n			$\delta_2 = 1$						$\delta_3 = 1$			
100	1.880	1.979	2.037	0.944	1.000	1.046	0.659	2.012	3.283	0.325	0.922	1.385
250	1.952	1.991	2.010	0.972	1.000	1.029	0.943	1.997	3.028	0.617	0.965	1.183
500	1.976	1.996	2.004	0.981	1.000	1.019	1.625	1.996	2.370	0.816	0.982	1.118
1000	1.989	1.998	2.002	0.987	1.000	1.014	1.843	1.998	2.145	0.890	0.993	1.082
			$\delta_2 = 2$						$\delta_3 = 2$			
100	1.888	1.978	2.008	1.945	2.001	2.046	1.274	1.985	2.667	1.373	1.950	2.327
250	1.954	1.991	2.000	1.972	2.000	2.029	1.814	1.992	2.191	1.763	1.995	2.178
500	1.978	1.996	2.001	1.981	1.999	2.019	1.905	1.996	2.071	1.867	1.991	2.116
1000	1.990	1.998	2.001	1.987	2.000	2.014	1.958	1.998	2.033	1.903	1.995	2.083
			$\delta_2 = 3$						$\delta_3 = 3$			
100	1.890	1.978	2.000	2.945	3.001	3.046	1.681	1.980	2.278	2.568	2.970	3.317
250	1.954	1.991	2.000	2.972	3.000	3.029	1.889	1.992	2.088	2.798	2.997	3.179
500	1.978	1.996	2.001	2.981	2.999	3.019	1.948	1.997	2.034	2.870	2.993	3.119
1000	1.990	1.998	2.001	2.987	3.000	3.014	1.978	1.998	2.018	2.904	2.996	3.083
			$\delta_2 = 4$						$\delta_3 = 4$			
100	1.894	1.977	2.000	3.945	4.001	4.046	1.794	1.979	2.157	3.631	3.980	4.316
250	1.954	1.991	2.000	3.972	4.000	4.029	1.915	1.991	2.054	3.813	3.999	4.175
500	1.978	1.995	2.000	3.981	3.999	4.019	1.963	1.996	2.026	3.873	3.995	4.123
1000	1.990	1.998	2.000	3.987	4.000	4.014	1.983	1.998	2.012	3.907	3.997	4.085

Table B.5: Threshold Parameter and Threshold Effect: $L_n = 6$

This table presents Monte Carlo quantiles of the estimates of the true threshold parameter $\gamma = 2$ and variant true threshold effects, using a 6th order Hermite basis function and sample sizes n. Columns (2)-(7) present the results for the DGP with only an exogenous regressor in (4.1)-(4.4) and the remaining columns present the results for the DGP that also includes an endogenous regressor in (B.6)-(B.8).

	Exogenous Regressor						Endogenous Regressor					
	Threshold Parameter			Threshold Effect			Threshold Parameter			Threshold Effect		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
n			$\delta_2 = 1$						$\delta_3 = 1$			
100	1.880	1.979	2.039	0.943	1.000	1.045	0.628	1.984	3.238	0.311	0.914	1.379
250	1.953	1.991	2.011	0.970	1.001	1.028	0.960	1.997	2.977	0.616	0.959	1.185
500	1.977	1.996	2.004	0.983	1.000	1.019	1.590	1.998	2.398	0.804	0.980	1.117
1000	1.990	1.998	2.002	0.988	1.000	1.013	1.844	1.997	2.145	0.889	0.991	1.083
			$\delta_2 = 2$						$\delta_3 = 2$			
100	1.887	1.978	2.007	1.943	2.000	2.045	1.260	1.986	2.632	1.338	1.945	2.323
250	1.954	1.991	2.000	1.970	2.001	2.028	1.799	1.992	2.176	1.769	1.993	2.181
500	1.978	1.996	2.001	1.983	2.000	2.019	1.912	1.996	2.074	1.870	1.991	2.118
1000	1.990	1.998	2.001	1.988	2.000	2.012	1.958	1.998	2.035	1.906	1.995	2.082
			$\delta_2 = 3$						$\delta_3 = 3$			
100	1.888	1.978	2.000	2.942	3.000	3.045	1.676	1.984	2.313	2.554	2.967	3.330
250	1.954	1.991	2.000	2.970	3.001	3.028	1.887	1.992	2.092	2.792	2.995	3.179
500	1.978	1.996	2.000	2.983	3.000	3.019	1.948	1.996	2.035	2.868	2.993	3.118
1000	1.990	1.998	2.001	2.988	3.000	3.012	1.978	1.998	2.019	2.905	2.996	3.084
			$\delta_2 = 4$						$\delta_3 = 4$			
100	1.893	1.978	2.000	3.942	4.000	4.046	1.785	1.981	2.158	3.616	3.973	4.325
250	1.954	1.991	2.000	3.970	4.001	4.028	1.915	1.992	2.053	3.810	3.999	4.179
500	1.978	1.995	2.000	3.983	4.000	4.019	1.962	1.996	2.026	3.874	3.995	4.124
1000	1.990	1.998	2.000	3.988	4.000	4.012	1.983	1.998	2.012	3.906	3.997	4.087

Table B.6: Confidence Interval Coverage of the Threshold Parameter

This table presents Monte Carlo results about the nominal 95% confidence interval coverage of the threshold parameter for true threshold effect $\delta_3 = 1, 2, 3, 4$ and order of Hermite basis function $L_n = 2, 3, 4, 5, 6$. The results are based on the DGP that also includes an endogenous regressor (B.6)-(B.8).

δ_3	1	2	3	4
		$L_n = 2$		
100	0.66	0.76	0.85	0.89
250	0.76	0.89	0.93	0.94
500	0.90	0.94	0.95	0.95
1000	0.94	0.97	0.97	0.96
		$L_n = 3$		
100	0.58	0.73	0.83	0.88
250	0.70	0.86	0.92	0.93
500	0.85	0.91	0.93	0.94
1000	0.91	0.94	0.94	0.94
		$L_n = 4$		
100	0.57	0.72	0.83	0.88
250	0.70	0.85	0.91	0.92
500	0.86	0.91	0.93	0.93
1000	0.92	0.95	0.94	0.95
		$L_n = 5$		
100	0.58	0.70	0.81	0.87
250	0.68	0.84	0.90	0.91
500	0.80	0.91	0.92	0.92
1000	0.90	0.92	0.93	0.94
		$L_n = 6$		
100	0.55	0.68	0.78	0.85
250	0.66	0.82	0.88	0.90
500	0.79	0.89	0.92	0.92
1000	0.88	0.93	0.92	0.93