Endogeneity in semiparametric threshold regression model with two threshold variables: Online supplement *

Chaoyi Chen[†‡] Thanasis Stengos[§] Yiguo Sun[¶]

July 12, 2022

Abstract

This paper considers a semiparametric threshold regression model with two threshold variables, extending Chen et al. (2012) and Kourtellos et al. (2022). The proposed model allows the endogeneity for both threshold variables and the slope regressors. Under the diminishing threshold effects assumption, we show the consistency and derive the asymptotic results of our proposed estimator for weakly dependent data. We study the finite sample performance of our proposed estimator via small Monte Carlo simulations and we apply our model to classify economic growth regimes based on both national public debt and national external debt.

JEL Classification: C14, C24, O40

Keywords: Control function; Public debt; Series estimation; Threshold regression.

*We would like to thank the editor and two anonymous referees for helpful comments and guidance on the revision of this paper. We also thank participants at the 29^{th} Annual Meeting of Midwest Econometrics Group for helpful discussion and comments. The views expressed here are the authors' and not necessarily those of the Central Bank of Hungary.

[†]Magyar Nemzeti Bank (Central Bank of Hungary), Budapest, 1054, Hungary; Email: chenc@mnb.hu.

[‡]MNB Institute, John von Neumann University, Kecskemét, 6000, Hungary.

[§]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.

This online supplement is composed of four parts. Section A uses a heuristic example to illustrate the inconsistency of the standard least squares estimator (LSE), which completely ignores the underlying endogeneity. Section B discusses the possibility to extend to the threshold model with three or more threshold variables and the curse of dimensionality. Section C contains the proofs of Lemmas C.1-C.4. Section D provides the proofs of Theorems 1-7.

Throughout the online supplement, let $\|.\|$ denote the Euclidean norm. All limits are taken as $n \to \infty$. $\overset{a.s.}{\to}$, $\overset{p}{\to}$, $\overset{d}{\to}$, and $\Longrightarrow$ denote almost sure convergence, convergence in probability, convergence in distribution and weak convergence respectively. Given two $n \times n$ Hermitian matrices A and B, let $B \preceq A$ denote $A - B$ is Hermitian positive semidefinite.

A The Inconsistency of the Standard LSE for the Endogenous Threshold Model with Two Threshold Variables: A Heuristic Example

In this section, we use a simplified version of model (2.1) to show how the threshold endogeneity affects the consistency of the standard least square estimator that ignores the underlying endogeneity. Our discussion is based on the following model

$$y = \sum_{i=2}^{4} \delta^{(i)} I^{(i)}(\gamma_1^0, \gamma_2^0) + \varepsilon, \quad E(\varepsilon|q) \neq 0. \tag{A.1}$$

For the illustration purpose, we assume that (q_1, q_2, ε) follows a joint normal distribution, $N\left(0, \begin{pmatrix} 1 & 0 & \rho_{10} \\ 0 & 1 & \rho_{20} \\ \rho_{10} & \rho_{20} & 1 \end{pmatrix}\right)$ with $\rho_{10} \neq 0$ and $\rho_{20} \neq 0$. Hence, $E(\varepsilon|q) = \rho_{10}q_1 + \rho_{20}q_2$. [1]

Now, we have $E(y|q) = \sum_{i=2}^{4} \delta^{(i)} E\left(I^{(i)}(\gamma_1^0, \gamma_2^0)\right) + \rho_{10}q_1 + \rho_{20}q_2$, and

$$y = \begin{cases} \rho_{10}q_1 + \rho_{20}q_2 + \sqrt{1-\rho_{10}^2-\rho_{20}^2}e, & \text{if } q_1 \leq \gamma_1^0,\ q_2 \leq \gamma_2^0, \\ \delta^{(2)} + \rho_{10}q_1 + \rho_{20}q_2 + \sqrt{1-\rho_{10}^2-\rho_{20}^2}e, & \text{if } q_1 \leq \gamma_1^0,\ q_2 > \gamma_2^0, \\ \delta^{(3)} + \rho_{10}q_1 + \rho_{20}q_2 + \sqrt{1-\rho_{10}^2-\rho_{20}^2}e, & \text{if } q_1 > \gamma_1^0,\ q_2 \leq \gamma_2^0, \\ \delta^{(4)} + \rho_{10}q_1 + \rho_{20}q_2 + \sqrt{1-\rho_{10}^2-\rho_{20}^2}e, & \text{if } q_1 > \gamma_1^0,\ q_2 > \gamma_2^0, \end{cases} \tag{A.2}$$

where $e \sim N(0,1)$ and is independent of q.

As we are in the interest of the inconsistency of the LSE of the threshold parameters, also considering the inconsistency of the slope estimator is standard, we hereinafter assume $\delta^{(i)}$ is

[1] This example essentially extends the discussion of Yu (2013) to the threshold regression model with two endogenous threshold variables.

known with $\delta^{(i)} \neq 0$ for $i = 2, 3, 4$ and $\delta^{(i)} \neq \delta^{(j)}$ for all $i \neq j$. Then, we can show

$$\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \sum_{i=2}^{4}\delta^{(i)} I_t^{(i)}(\gamma_1,\gamma_2)\right)^2 \xrightarrow{p} C + E\left[\left(\sum_{i=2}^{4}\delta^{(i)}\left(I_t^{(i)}(\gamma_1^0,\gamma_2^0) - I_t^{(i)}(\gamma_1,\gamma_2)\right) + \rho_{10}q_{1t} + \rho_{20}q_{2t}\right)^2\right], \tag{A.3}$$

where C is a bounded constant.

Let $\Phi(\cdot)$ and $\phi(\cdot)$ be the cumulative density function (CDF) and probability density function (PDF) for the standard normal distribution respectively. Then, by simple calculation, we have

$$E\left[\left(\sum_{i=2}^{4}\delta^{(i)}\left(I_t^{(i)}(\gamma_1^0,\gamma_2^0) - I_t^{(i)}(\gamma_1,\gamma_2)\right) + \rho_{10}q_{1t} + \rho_{20}q_{2t}\right)^2\right] = S_1(\gamma_1,\gamma_2) + S_2(\gamma_1,\gamma_2) + \rho_{10}^2 + \rho_{20}^2, \tag{A.4}$$

where

$$S_1(\gamma_1,\gamma_2) = \begin{cases} S_1^{(1)}(\gamma_1,\gamma_2), & \text{if } \gamma_1 \leq \gamma_1^0,\ \gamma_2 \leq \gamma_2^0, \\ S_1^{(2)}(\gamma_1,\gamma_2), & \text{if } \gamma_1 \leq \gamma_1^0,\ \gamma_2 > \gamma_2^0, \\ S_1^{(3)}(\gamma_1,\gamma_2), & \text{if } \gamma_1 > \gamma_1^0,\ \gamma_2 \leq \gamma_2^0, \\ S_1^{(4)}(\gamma_1,\gamma_2), & \text{if } \gamma_1 > \gamma_1^0,\ \gamma_2 > \gamma_2^0, \end{cases}$$

$$\begin{aligned} S_2(\gamma_1,\gamma_2) &= \left(\delta^{(2)} + \delta^{(3)} - \delta^{(4)}\right)\Big[\rho_{10}\left\{\left(\phi(\gamma_1^0) - \phi(\gamma_1)\right)\Phi(\gamma_2) + \phi(\gamma_1^0)\left[\Phi(\gamma_2^0) - \Phi(\gamma_2)\right]\right\} \\ &+ \rho_{20}\left\{\left[\Phi(\gamma_1^0) - \Phi(\gamma_1)\right]\phi(\gamma_2) + \Phi(\gamma_1^0)\left[\phi(\gamma_2^0) - \phi(\gamma_2)\right]\right\}\Big] + \left(\delta^{(2)} - \delta^{(4)}\right)\rho_{10}\left[\phi(\gamma_1) - \phi(\gamma_1^0)\right] \\ &+ \left(\delta^{(3)} - \delta^{(4)}\right)\rho_{20}\left[\phi(\gamma_2) - \phi(\gamma_2^0)\right], \end{aligned}$$

and

$$S_1^{(1)}(\gamma_1,\gamma_2) = \left(\delta^{(2)} - \delta^{(4)}\right)^2 \left[\Phi(\gamma_1^0) - \Phi(\gamma_1)\right](1-\Phi(\gamma_2)) + \delta^{(2)^2}\Phi(\gamma_1^0)\left[\Phi(\gamma_2^0) - \Phi(\gamma_2)\right]$$
$$+ \left(\delta^{(3)} - \delta^{(4)}\right)^2 (1-\Phi(\gamma_1^0))\left[\Phi(\gamma_2^0) - \Phi(\gamma_2)\right] + \delta^{(3)^2}\left[\Phi(\gamma_1^0) - \Phi(\gamma_1)\right]\Phi(\gamma_2),$$

$$S_1^{(2)}(\gamma_1,\gamma_2) = \left(\delta^{(2)} - \delta^{(4)}\right)^2 \left[\Phi(\gamma_1^0) - \Phi(\gamma_1)\right](1-\Phi(\gamma_2)) + \delta^{(2)^2}\Phi(\gamma_1^0)\left[\Phi(\gamma_2) - \Phi(\gamma_2^0))\right]$$
$$+ \left(\delta^{(3)} - \delta^{(4)}\right)^2 (1-\Phi(\gamma_1^0))\left[\Phi(\gamma_2) - \Phi(\gamma_2^0)\right] + \delta^{(3)^2}\left[\Phi(\gamma_1^0) - \Phi(\gamma_1)\right]\Phi(\gamma_2),$$

$$S_1^{(3)}(\gamma_1,\gamma_2) = \left(\delta^{(2)} - \delta^{(4)}\right)^2 \left[\Phi(\gamma_1) - \Phi(\gamma_1^0)\right](1-\Phi(\gamma_2)) + \delta^{(2)^2}\Phi(\gamma_1^0)\left[\Phi(\gamma_2^0) - \Phi(\gamma_2)\right]$$
$$+ \left(\delta^{(3)} - \delta^{(4)}\right)^2 (1-\Phi(\gamma_1^0))\left[\Phi(\gamma_2^0) - \Phi(\gamma_2)\right] + \delta^{(3)^2}\left[\Phi(\gamma_1) - \Phi(\gamma_1^0)\right]\Phi(\gamma_2),$$

$$S_1^{(4)}(\gamma_1,\gamma_2) = \left(\delta^{(2)} - \delta^{(4)}\right)^2 \left[\Phi(\gamma_1) - \Phi(\gamma_1^0)\right](1-\Phi(\gamma_2)) + \delta^{(2)^2}\Phi(\gamma_1^0)\left[\Phi(\gamma_2) - \Phi(\gamma_2^0)\right]$$
$$+ \left(\delta^{(3)} - \delta^{(4)}\right)^2 (1-\Phi(\gamma_1^0))\left[\Phi(\gamma_2) - \Phi(\gamma_2^0)\right] + \delta^{(3)^2}\left[\Phi(\gamma_1) - \Phi(\gamma_1^0)\right]\Phi(\gamma_2).$$

We make the following remarks for the limiting result of equation (A.4).

Remark 1 *S_1 is uniquely minimized at $\gamma_1 = \gamma_1^0$ and $\gamma_2 = \gamma_2^0$.*

Remark 2 *The minimum point of the objective function S_2 is not fixed but jointly determined by the jump size, the endogenous degree, and the intensity of threshold variables at the true threshold level. As such, S_2 is an endogenous distortion term that contributes to the inconsistency of the LSE. Note that, if both threshold variables are exogenous, $\rho_{10} = \rho_{20} = 0$. As a result, $S_2 = 0$ and $\widehat{\gamma}^{LSE} = [\widehat{\gamma}_1^{LSE}, \widehat{\gamma}_2^{LSE}]$ converges to $\gamma^0 = [\gamma_1^0, \gamma_2^0]$.*

Remark 3 *The consistency of the LSE is determined by the tradeoff between the correct identification effect from S_1 and the endogeneous distortion effect from S_2. In particular, we observe S_2 dominates the overall effect when the jump sizes are small, the degree of endogeneity is large, and the true threshold level is not at the mode of the distribution of the threshold variable.*

To further elaborate our Remark 3, we experiment a numerical example by considering the case when only q_2 is endogenous and plot the figures of the corresponding minimizer, γ_2, of

$E\left[\left(\sum_{i=2}^{4}\delta^{(i)}I^{(i)}(\gamma_1,\gamma_2)+\rho_{10}q_{1t}+\rho_{20}q_{2t}\right)^2\right]$ as a function of γ_2^0 and ρ_{20} given a known γ_1^0. Figure 1 shows the results with large jump size, where $\delta^{(2)}=2$, $\delta^{(3)}=1$, $\delta^{(4)}=4$, and $\rho_{10}=0$. Top left and right figures show the results with $\rho_{20}=0.5$ and $\rho_{20}=1$, respectively. We observe γ_2 can be consistently estimated with lesser degree of endogeneity. However, as shown in bottom left ($\rho_{20}=2.5$) and right ($\rho_{20}=5$) figures, with high degree of endogeneity, $\widehat{\gamma}_2$ is generally inconsistent.

Figure 2 shows the results with relatively small jump size, $\delta^{(2)}=0.2$, $\delta^{(3)}=0.7$, $\delta^{(4)}=0.4$. Obviously, we observe more inconsistency than those in Figure 1. In particular, γ_2 is now inconsistent even under relatively small ρ_{20}. By comparison, for our special model (A.2), the smaller (but nonzero) the jump size is, the more significant the inconsistency of the LSE is.

Interestingly, Figures 1 -2 also reveal γ_2^0 can be consistently estimated when $\gamma_2^0=0$, irrespective of the jump size and the degree of endogeneity. Considering $q_2\sim N(0,1)$ under our setup, γ_2^0 is the mode of the distribution of q_2, which implies intensity of threshold variables at the true threshold level also affects the consistency of the LSE. Thus, we conclude the consistency of the LSE is jointly determined by the jump size, the endogenous degree, and the intensity of threshold variables at the true threshold level.

B Discussion on threshold model with three threshold variables

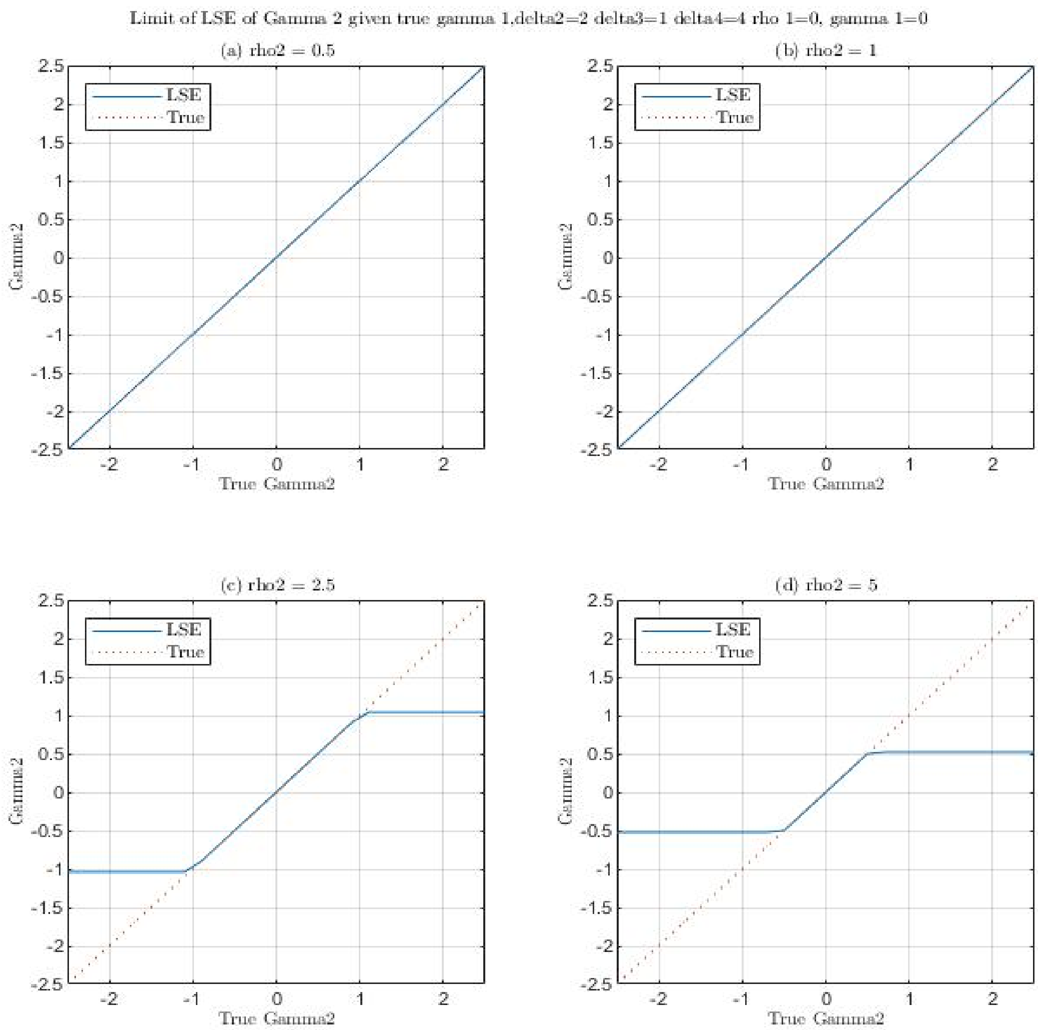
This section uses a parsimonious model to discuss the possible extension to three or more threshold variables. In particular, our discussion is mainly based on the following model

$$y_t=\sum_{i=1}^{8}\beta_i I_t^{(i)}(\gamma_1^0,\gamma_2^0,\gamma_3^0)+\varepsilon_t, \tag{B.1}$$

and

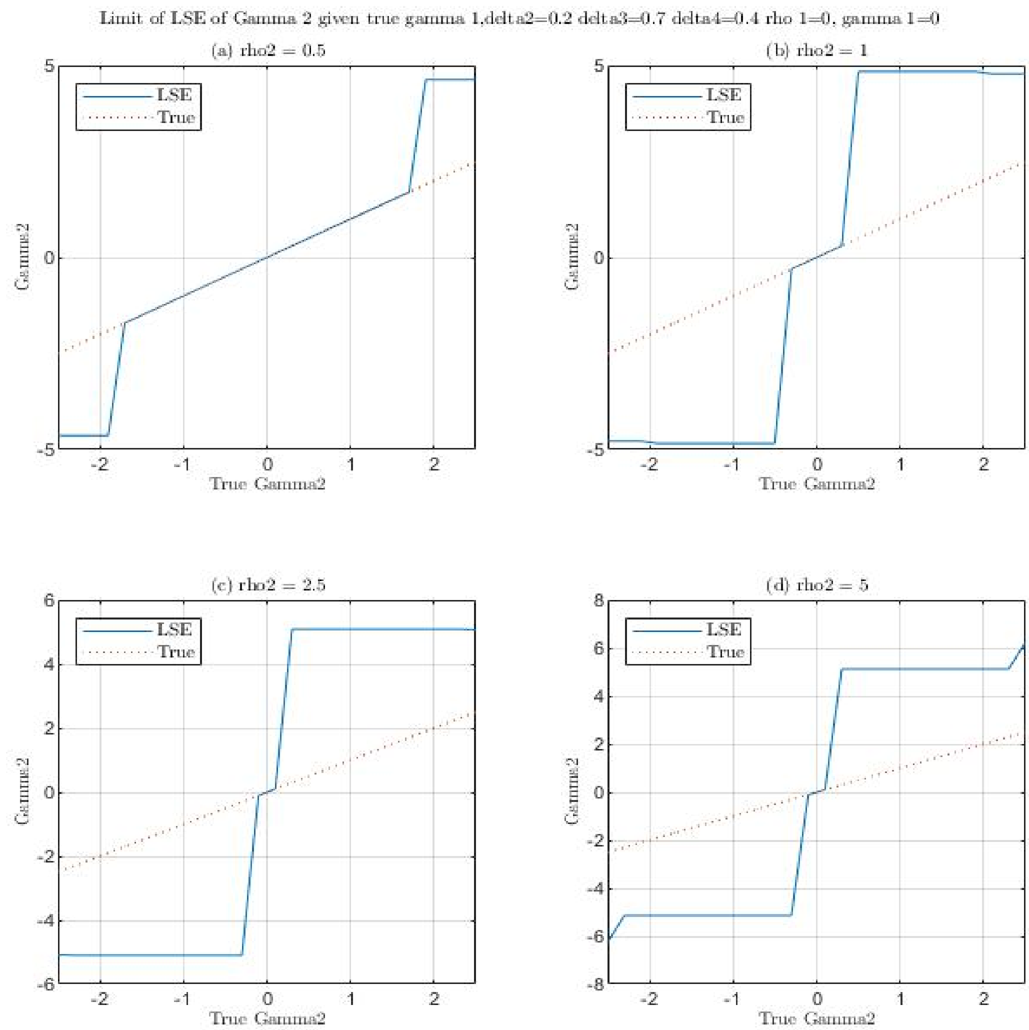
$$I_t^{(i)}(\gamma_1^0,\gamma_2^0,\gamma_3^0)=\begin{cases} I_t^{(1)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}\le\gamma_1^0,\ q_{2t}\le\gamma_1^0,\ q_{3t}\le\gamma_1^0,\\ I_t^{(2)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}>\gamma_1^0,\ q_{2t}\le\gamma_1^0,\ q_{3t}\le\gamma_1^0,\\ I_t^{(3)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}\le\gamma_1^0,\ q_{2t}>\gamma_1^0,\ q_{3t}\le\gamma_1^0,\\ I_t^{(4)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}>\gamma_1^0,\ q_{2t}>\gamma_1^0,\ q_{3t}\le\gamma_1^0,\\ I_t^{(5)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}\le\gamma_1^0,\ q_{2t}\le\gamma_1^0,\ q_{3t}>\gamma_1^0,\\ I_t^{(6)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}>\gamma_1^0,\ q_{2t}\le\gamma_1^0,\ q_{3t}>\gamma_1^0,\\ I_t^{(7)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}\le\gamma_1^0,\ q_{2t}>\gamma_1^0,\ q_{3t}>\gamma_1^0,\\ I_t^{(8)}(\gamma_1^0,\gamma_2^0,\gamma_3^0), & \text{if } q_{1t}>\gamma_1^0,\ q_{2t}>\gamma_1^0,\ q_{3t}>\gamma_1^0. \end{cases}$$

Figure 1: Limit of $\widehat{\gamma}_2^{LSE}$ given $\gamma = \gamma_1^0$, $\delta^{(2)} = 2$, $\delta^{(3)} = 1$, $\delta^{(4)} = 4$, $\rho_{10} = 0$, $\gamma_1^0 = 0$



Note: This figure provides the limit of $\widehat{\gamma}_2^{LSE}$ given γ_1^0 is known with $\delta^{(2)} = 2$, $\delta^{(3)} = 1$, $\delta^{(4)} = 4$, $\rho_{10} = 0$, $\gamma_1^0 = 0$. (a)-(d) show the limit of $\widehat{\gamma}_2^{LSE}$ when $\rho_{20} = 0.5$, $\rho_{20} = 1$, $\rho_{20} = 2.5$, and $\rho_{20} = 5$, respectively. The solid blue line stands for the $\widehat{\gamma}_2^{LSE}$. The dashed red line represents the true threshold points.

Figure 2: Limit of $\widehat{\gamma}_2^{LSE}$ given $\gamma = \gamma_1^0$, $\delta^{(2)} = 0.2$, $\delta^{(3)} = 0.7$, $\delta^{(4)} = 0.4$, $\rho_{10} = 0$, $\gamma_1^0 = 0$



Note: This figure provides the limit of $\widehat{\gamma}_2^{LSE}$ given γ_1^0 is known with $\delta^{(2)} = 0.2$, $\delta^{(3)} = 0.7$, $\delta^{(4)} = 0.4$, $\rho_{10} = 0$, $\gamma_1^0 = 0$. (a)-(d) show the limit of $\widehat{\gamma}_2^{LSE}$ when $\rho_{20} = 0.5$, $\rho_{20} = 1$, $\rho_{20} = 2.5$, and $\rho_{20} = 5$, respectively. The solid blue line stands for the $\widehat{\gamma}_2^{LSE}$. The dashed red line represents the true threshold points.

B.1 Consistency of LSE under exogenous threshold variables

In this subsection, we show the consistency of the LSE with exogenous threshold variables. As we are in the interest of the the least squares threshold estimators, we hereinafter assume the knowledge of β_1 to β_8. We make the following regularity assumptions to support model (B.1).

Assumption B.1 ***(a)*** $\{(q_{1t}, q_{2t}, q_{3t}, \varepsilon_t)\}$ *is a strictly stationary, ergodic, and* ρ*-mixing sequence with the mixing coefficient* $\{\rho_r\}$ *satisfying* $\sum_{r=1}^{\infty} \rho_r^{1/2} < \infty$ *and* $\gamma_g^0 \in \Gamma \equiv [\underline{\gamma}, \overline{\gamma}]$ *for* $g = 1, 2, 3$*, where* $\underline{\gamma}$ *and* $\overline{\gamma}$ *are two bounded constants;*

(b) *Let* $\mathcal{F}_{n,t}^*$ *be the smallest sigma-field generated by* $\{(q_{1s+1}, q_{2s+1}, , q_{3s+1}, \varepsilon_s) : 1 \leq s < t \leq n\}$. *Then,* $\left\{\left(\varepsilon_t, \mathcal{F}_{n,t}^*\right)\right\}_{t=1}^n$ *is martingale difference sequence;*

(c) *The threshold variables* q_{1t}*,* q_{2t}*, and* q_{3t} *have a common continuous joint distribution* $F(\gamma_1, \gamma_2, \gamma_3)$ *and a corresponding joint density* $f(\gamma_1, \gamma_2, \gamma_3)$*. For* $j = 1, 2, 3$*,* $f_j(\gamma_1, \gamma_2, \gamma_3) = \frac{\partial F(\gamma_1, \gamma_2, \gamma_3)}{\partial \gamma_j}$ *and* $0 < f(\gamma_1, \gamma_2, \gamma_3) \leq \bar{f} < \infty$ *and* $0 < f_j(\gamma_1, \gamma_2, \gamma_3) \leq \bar{f}_j < \infty$

(d) *For* $i = 1, \ldots, 8$ *and* $j = 1, \ldots, 8$*,* $\beta_i \neq \beta_j$ *for all* $i \neq j$*.*

Theorem B.1 *Under Assumption B.1, as* $n \longrightarrow \infty$*, we have* $\widehat{\gamma} - \gamma_0 = o_p(1)$*.*

Proof. First, by Assumption B.1 (a)-(b), uniformly for all $\gamma \in \Theta_\gamma$, we can show

$$\begin{aligned}\frac{1}{n}\left(SSR(\gamma) - SSR(\gamma_0)\right) &= \frac{1}{n}\sum_{t=1}^{n}\left(y_t - \sum_{i=1}^{8}\beta_i I_t^{(i)}(\gamma_1, \gamma_2, \gamma_3)\right)^2 - \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}\left[\sum_{i=1}^{8}\beta_i A_t^i(\gamma_1, \gamma_2, \gamma_3)\right]^2 + o_p(1) \xrightarrow{p} E\left(\sum_{i=1}^{8}\beta_i A_t^i(\gamma_1, \gamma_2, \gamma_3)\right)^2,\end{aligned}$$

where $A_t^i(\gamma_1, \gamma_2, \gamma_3) = I_t^{(i)}(\gamma_1^0, \gamma_2^0, \gamma_3^0) - I_t^{(i)}(\gamma_1, \gamma_2, \gamma_3)$.

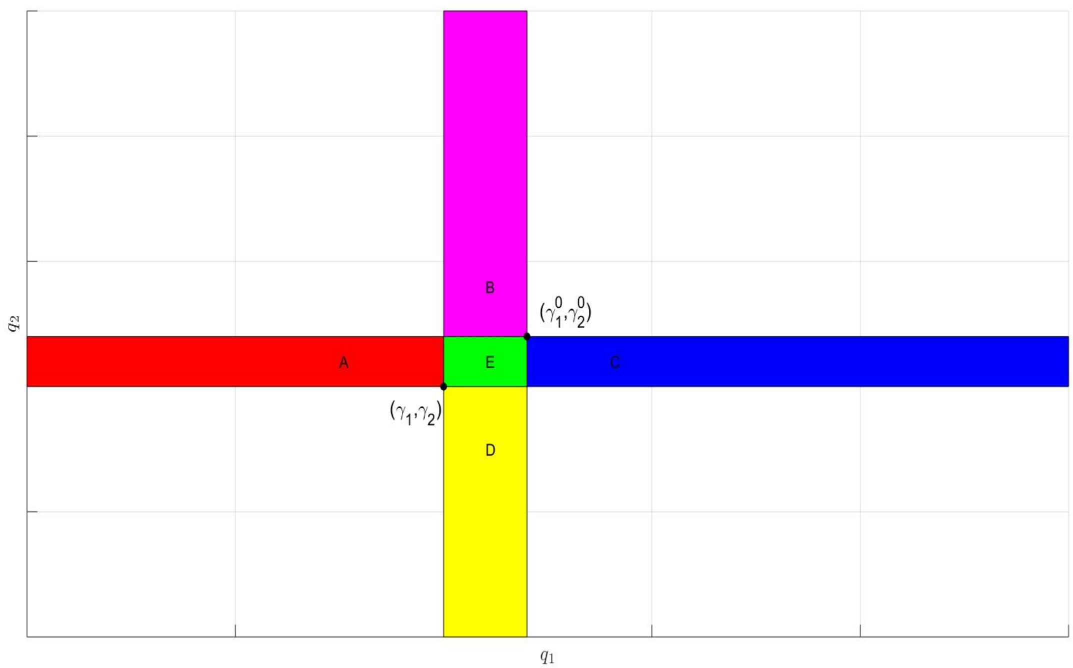
Next, we partition our discussion into eight different cases.

Case 1: $\gamma_1 < \gamma_1^0$, $\gamma_2 < \gamma_2^0$, $\gamma_3 < \gamma_3^0$.

Following the decomposing method, we used to derive the consistency of LSE for the threshold model with one or two thresholds[2], we decompose each $A_i(\gamma_1, \gamma_2, \gamma_3)$ into some of the nineteen

[2]Figure 3 shows the intuition on decomposing $I^j(\gamma_1^0, \gamma_2^0) - I^j(\gamma_1, \gamma_2)$ into five mutually independent terms.

Figure 3: Intuition on decomposing $I^j(\gamma_1^0, \gamma_2^0) - I^j(\gamma_1, \gamma_2)$



Note: This figure presents an example on how we decompose $I^j(\gamma_1^0, \gamma_2^0) - I^j(\gamma_1, \gamma_2)$ into five mutually independent terms for the case of $\gamma_1 < \gamma_1^0$ and $\gamma_2 < \gamma_0^2$. Specifically, the area of $A + E + D$ stands for $I^1(\gamma_1^0, \gamma_2^0) - I^1(\gamma_1, \gamma_2)$. The area of $C - D$ shows $I^2(\gamma_1^0, \gamma_2^0) - I^2(\gamma_1, \gamma_2)$. The area of $B - A$ captures $I^3(\gamma_1^0, \gamma_2^0) - I^3(\gamma_1, \gamma_2)$. The area of $-B - E - C$ represents $I^4(\gamma_1^0, \gamma_2^0) - I^4(\gamma_1, \gamma_2)$.

mutually independent terms. We have

$$
\begin{aligned}
&\sum_{i=1}^{8} \beta_i A_t^i(\gamma_1,\gamma_2,\gamma_3) \\
&= (\beta_1-\beta_2)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} \le \gamma_2, q_{3t} \le \gamma_3\right) + (\beta_1-\beta_3)\, I\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} \le \gamma_3\right) \\
&+ (\beta_1-\beta_4)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} \le \gamma_3\right) + (\beta_1-\beta_5)\, I\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) \\
&+ (\beta_1-\beta_6)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} \le \gamma_2, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_1-\beta_7)\, I\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_1-\beta_8)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_2-\beta_4)\, I\left(q_{1t} > \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} < \gamma_3\right) \\
&+ (\beta_2-\beta_6)\, I\left(q_{1t} > \gamma_1^0, q_{2t} \le \gamma_2, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_2-\beta_8)\, I\left(q_{1t} > \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_3-\beta_4)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} > \gamma_2^0, q_{3t} < \gamma_3\right) + (\beta_3-\beta_7)\, I\left(q_{1t} \le \gamma_1, q_{2t} > \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_3-\beta_8)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} > \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_4-\beta_8)\, I\left(q_{1t} > \gamma_1^0, q_{2t} > \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_5-\beta_6)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} \le \gamma_2, q_{3t} > \gamma_3^0\right) + (\beta_5-\beta_7)\, I\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) \\
&+ (\beta_5-\beta_8)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) + (\beta_6-\beta_7)\, I\left(q_{1t} > \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) \\
&+ (\beta_7-\beta_8)\, I\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} > \gamma_2^0, q_{3t} > \gamma_3^0\right).
\end{aligned}
$$

Thus, by Assumption B.1(a)-(c), we have

$$
\begin{aligned}
&E\left(\sum_{i=1}^{8} \beta_i A_t^i(\gamma_1,\gamma_2,\gamma_3)\right)^2 \\
&= (\beta_1-\beta_2)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} \le \gamma_2, q_{3t} \le \gamma_3\right) + (\beta_1-\beta_3)^2 F\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} \le \gamma_3\right) \\
&+ (\beta_1-\beta_4)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} \le \gamma_3\right) + (\beta_1-\beta_5)^2 F\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) \\
&+ (\beta_1-\beta_6)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} \le \gamma_2, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_1-\beta_7)^2 F\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_1-\beta_8)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_2-\beta_4)^2 F\left(q_{1t} > \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} < \gamma_3\right) \\
&+ (\beta_2-\beta_6)^2 F\left(q_{1t} > \gamma_1^0, q_{2t} \le \gamma_2, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_2-\beta_8)^2 F\left(q_{1t} > \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_3-\beta_4)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} > \gamma_2^0, q_{3t} < \gamma_3\right) + (\beta_3-\beta_7)^2 F\left(q_{1t} \le \gamma_1, q_{2t} > \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_3-\beta_8)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} > \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) + (\beta_4-\beta_8)^2 F\left(q_{1t} > \gamma_1^0, q_{2t} > \gamma_2^0, \gamma_3 \le q_{3t} \le \gamma_3^0\right) \\
&+ (\beta_5-\beta_6)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} \le \gamma_2, q_{3t} > \gamma_3^0\right) + (\beta_5-\beta_7)^2 F\left(q_{1t} \le \gamma_1, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) \\
&+ (\beta_5-\beta_8)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) + (\beta_6-\beta_7)^2 F\left(q_{1t} > \gamma_1^0, \gamma_2 \le q_{2t} \le \gamma_2^0, q_{3t} > \gamma_3^0\right) \\
&+ (\beta_7-\beta_8)^2 F\left(\gamma_1 \le q_{1t} \le \gamma_1^0, q_{2t} > \gamma_2^0, q_{3t} > \gamma_3^0\right).
\end{aligned}
$$

Therefore, by Assumption B.1(d), for all $\gamma \in \Theta_\gamma \backslash (\gamma_1^0, \gamma_2^0, \gamma_3^0)$, we have

$$
E\left(\sum_{i=1}^{8} \beta_i A_t^i(\gamma_1,\gamma_2,\gamma_3)\right)^2 > E\left(\sum_{i=1}^{8} \beta_i A_t^i(\gamma_1^0,\gamma_2^0,\gamma_3^0)\right)^2 = 0.
$$

For all eight cases, we can show above inequality by analogue, which concludes the proof. ■

B.2 Inconsistency of the LSE under endogenous threshold variables and the nonparametric control function approach

Now we assume some of the threshold variables are endogenous and show the inconsistency of the LSE under endogenous threshold variables. We also outline how to deliver the consistency of LSE with the nonparametric control function approach. We make the following assumptions for model (B.1) to allow for the endogeneity.

Assumption B.2 ***(a)*** *Assumption B.1 (a), (c), and (d) holds;*
(b) *Let* $\mathcal{F}^{*'}_{n,t}$ *be the smallest sigma-field generated by* $\left\{\left(q_{1s}, q_{2s}, q_{3s}, z^T_{q_{s+1}}, \varepsilon_s\right) : 1 \le s < t \le n\right\}$, *where* $z^T_{q_t}$ *is the instrumental variables for* q_t *and we have reduced-form* $q_t = \pi_q z_{q_t} + v_{q_t}$ *with* $q_t = [q_{1t}, q_{2t}, q_{3t}]^T$ *and* $v_{q_t} = [v_{q_{1t}}, v_{q_{2t}}, v_{q_{3t}}]^T$*. Then,* $\left\{\left(\varepsilon_t, \mathcal{F}^{*'}_{n,t}\right)\right\}_{t=1}^n$ *is martingale difference sequence.*
(c) $E(y_t|\mathcal{F}^{*'}_{n,t}, q_{1t}, q_{2t}, q_{3t}) = \sum_{i=1}^8 h_i(v_{q_{1t}}, v_{q_{2t}}, v_{q_{3t}}) I^{(i)}_t(\gamma^0_1, \gamma^0_2, \gamma^0_3)$.

Thus, under Assumption B.2, we have

$$\frac{1}{n}\left(SSR(\gamma) - SSR(\gamma_0)\right) = \frac{1}{n}\sum_{t=1}^n \left(y_t - \sum_{i=1}^8 \beta_i I^{(i)}_t(\gamma_1, \gamma_2, \gamma_3)\right)^2 - \frac{1}{n}\sum_{t=1}^n \varepsilon_t^2$$
$$= \frac{1}{n}\sum_{t=1}^n \left[\sum_{i=1}^8 \beta_i A^i_t(\gamma_1, \gamma_2, \gamma_3)\right]^2 + \frac{2}{n}\sum_{t=1}^n \left[\sum_{i=1}^8 \beta_i A^i_t(\gamma_1, \gamma_2, \gamma_3)\right]\varepsilon_t = A(\gamma_1, \gamma_2, \gamma_3) + B(\gamma_1, \gamma_2, \gamma_3).$$

While $A(\gamma_1, \gamma_2, \gamma_3)$ is always minimized at γ_0, and $B(\gamma_1, \gamma_2, \gamma_3)$ is the endogenous distortion term and cannot achieve minimum at $\gamma = \gamma_0$ in general due to $E(\varepsilon|q) \neq 0$. Therefore, the LSE of the threshold regression model with three endogenous threshold variables is inconsistent.

Next, we show how the nonparametric control function approach helps us correct the endogenous bias term in a nutshell. Note that, assuming the knowledge of v_{q_t} and by Assumption B.2(c), we can rewrite model (B.1) as $y_t = \sum_{i=1}^8 h_i(v_{q_{1t}}, v_{q_{2t}}, v_{q_{3t}}) I^{(i)}_t(\gamma^0_1, \gamma^0_2, \gamma^0_3) + \varepsilon^*_t$, where $\left\{\left(\varepsilon^*_t, \mathcal{F}^{*''}_{n,t}\right)\right\}$ with $\mathcal{F}^{*''}_{n,t} = \sigma\left(z^T_{q_t}, q^T_{t-1}, \varepsilon_{t-1}, \ldots\right)$ is a martingale difference sequence with $E\left(\varepsilon^*_t|\mathcal{F}^{*''}_{n,t}\right) = 0$ almost surely. For $i = 1, \ldots, 8$, suppose we can approximate $h_i(v_{q_{1t}}, v_{q_{2t}}, v_{q_{3t}})$ by $h^*_i(v_{q_{1t}}, v_{q_{2t}}, v_{q_{3t}}) = \alpha^T_{Ln,i}\Phi_{Ln}(v_{q_{1t}}, v_{q_{2t}}, v_{q_{3t}})$ such that $\sup_{w\in W} |h_i(w) - h^*_i(w)\,| < ML_n^{-p}$, where $W = W_1 \times W_2 \times W_3$

and W_1, W_2, W_3 are all compact subset of $\mathcal{R}$. Then, we can show

$$\frac{1}{n}\left(SSR(\gamma)-SSR(\gamma_0)\right)=\frac{1}{n}\sum_{t=1}^{n}\left(y_t-\sum_{i=1}^{8}h_i^*(v_{q_{1t}},v_{q_{2t}},v_{q_{3t}})I_t^{(i)}(\gamma_1,\gamma_2,\gamma_3)\right)^2-\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^{*2}$$
$$=\frac{1}{n}\sum_{t=1}^{n}\left(\sum_{i=1}^{8}A_t^{*i}(\gamma_1,\gamma_2,\gamma_3)+\varepsilon_t^*\right)^2-\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^{*2}=\frac{1}{n}\sum_{t=1}^{n}\left[\sum_{i=1}^{8}A_t^{*i}(\gamma_1,\gamma_2,\gamma_3)\right]^2+o_p(1),$$

where $A_t^{*i}(\gamma_1,\gamma_2,\gamma_3)=h_i(v_{q_{1t}},v_{q_{2t}},v_{q_{3t}})I_t^{(i)}(\gamma_1^0,\gamma_2^0,\gamma_3^0)-h_i^*(v_{q_{1t}},v_{q_{2t}},v_{q_{3t}})I_t^{(i)}(\gamma_1,\gamma_2,\gamma_3)$. Thus, given sieve approximation error is small enough and negligible such that $A_t^{*i}(\gamma_1,\gamma_2,\gamma_3)\approx h_i(v_{q_{1t}},v_{q_{2t}},v_{q_{3t}})$ $\times A_t^{i}(\gamma_1,\gamma_2,\gamma_3)$, we can apply the same lines of steps in the proof of the consistency of the LSE without endogeneity above by decomposing the $A_t^{*i}(\gamma_1,\gamma_2,\gamma_3)$ into some of nineteen mutually independent terms and show the consistency of the LSE of the semiparametric threshold regression model. However, the sieve approximation error is more prominent here than the case with two endogenous threshold variables, pointing to the issue of the curse-of-dimensionality. The subsection below discusses this problem more.

B.3 Threshold model with more than three threshold variables and the curse of dimensionality

Naturally, we can extend our discussion to the threshold regression model with more than three threshold variables. As shown above, the most crucial step is to decompose the leading term of the least-squares objective function into several mutually independent terms such that we can split the square of sum into some mutually independent sum of squares terms.

However, the more threshold variables used in a threshold variable, the less accurate the estimate will be. First, a sufficiently large sample size needed to ensure that each regime has enough observations for estimation accuracy. Second, if all the threshold variables are endogenous, our proposed semiparametric approach requires the inclusion of endogenous threshold correction terms, which are unknown functions of d arguments with d being the number of the endogenous threshold variables. It is well-known that the sieve approximation error is of order $O(L_n^{-p})$, where $p=p_0/d$ and p_0 measures the smoothness of the endogenous threshold correction terms; see, e.g., Newey (1997) and Chen (2007). Hence, the higher the dimension of endogenous threshold variable, q_t, the larger the sieve approximation error, which results the curse-of-dimensionality of sieve estimation. To derive the consistency and asymptotic distribution, we need Assumption 2 (f) also holds for the semiparametric threshold regression model with three or more endogenous threshold variables. Let $L_n=c_0n^l$ for some constants $c_0>0$ and $l\in(0,1)$. Assumption 2 (f) essentially requires

$l \in \left(\frac{d}{2p_0}, min\left(1/4, 1-2min(\alpha,\rho)\right)\right)$. Thus, the upper bound of d is $2p_0 min\left(1/4, 1-2min(\alpha,\rho)\right)$. The two reasons limit the use of the threshold model with more than two endogenous threshold variables, but a threshold regression model whose regime is determined by a linear combination of multiple threshold variables can be used to reduce the dimensionality concern; see, e.g., Seo & Linton (2007).

C Lemmas

Let $\chi_t = \left[x_t^T, \Phi_{Ln}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})\right]^T$ and $\chi_t^* = \left[x_t^T, \Phi_{Ln}^T(v_{q_1t}, v_{q_2t})\right]^T$. Also, let M and $M_j(\gamma_1,\gamma_2)$ denote the moment functionals defined as $M = E\left(\chi_t^*\chi_t^{*T}\right)$ and $M_j(\gamma_1,\gamma_2) = E\left(\chi_t^*\chi_t^{*T} I_t^{(j)}(\gamma_1,\gamma_2)\right)$.

Lemma C.1 *Under Assumptions 1-2, we have*

$$\max_{\gamma\in\Theta_\gamma} \|n^{-1}\sum_{t=1}^n \chi_t\chi_t^T I_t^{(j)}(\gamma_1,\gamma_2) - M_j(\gamma_1,\gamma_2)\| = o_p(1),$$

for $j = 1,2,3,4$.

Proof. Applying Lemma 2 in Chen et al. (2012), we obtain

$$\max_{\gamma\in\Theta_\gamma} \|n^{-1}\sum_{t=1}^n x_t x_t^T I_t^{(j)}(\gamma_1,\gamma_2) - E(x_t x_t^T I_t^{(j)}(\gamma_1,\gamma_2))\| = o_p(1). \tag{C.1}$$

As χ_t is of dimension $k + L_n$, which grows as the sample size increases, we will extend Lemma 2 of Chen et al. (2012) to infinite dimension case.

First, under Assumptions 1 (a)-(b), applying the first order Taylor expansion, we have

$$\Phi_{L_n}(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) = \Phi_{L_n}(v_{q_1t}, v_{q_2t}) + \mathbb{D}\Phi_{L_n}(\tilde{v}_{q_1t}, \tilde{v}_{q_2t})\begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1} - \pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2} - \pi_{q_2})\end{bmatrix},$$

where $\tilde{v}_{q_jt}$ lies between v_{q_jt} and $\widehat{v}_{q_jt}$, $v_{q_jt} - \widehat{v}_{q_jt} = z_{jt}^T(\widehat{\pi}_{q_j} - \pi_{q_j})$ for $j = 1,2$ and $\mathbb{D}\Phi_{L_n}(v_{q_1t}, v_{q_2t})$ is the $L_n \times 2$ first order derivative matrix w.r.t v_{q_1t} and v_{q_2t}.

Hence, we have

$$\begin{aligned} n^{-1}\sum_{t=1}^n x_t\Phi_{L_n}(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})^T I_t^{(j)}(\gamma_1,\gamma_2) &= n^{-1}\sum_{t=1}^n x_t\Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I_t^{(j)}(\gamma_1,\gamma_2) \\ + n^{-1}\sum_{t=1}^n x_t \begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1} - \pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2} - \pi_{q_2})\end{bmatrix}^T &\mathbb{D}^T\Phi_{L_n}(\tilde{v}_{q_1t}, \tilde{v}_{q_2t}) I_t^{(j)}(\gamma_1,\gamma_2) = A_{n,1}^{(j)}(\gamma_1,\gamma_2) + A_{n,2}^{(j)}(\gamma_1,\gamma_2), \end{aligned} \tag{C.2}$$

where, under Assumptions 2 (b) and (e), as $\widehat{\pi}_{q_1} - \pi_{q_1} = O_p(n^{-1/2})$ and $\widehat{\pi}_{q_2} - \pi_{q_2} = O_p(n^{-1/2})$, for all $\gamma \in \Theta_\gamma$,

$$\|A_{n,2}^{(j)}(\gamma_1, \gamma_2)\| \leq \|n^{-1} \sum_{t=1}^{n} x_t \begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1} - \pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2} - \pi_{q_2}) \end{bmatrix}^T \mathbb{D}^T \Phi_{L_n}(\tilde{v}_{q_1t}, \tilde{v}_{q_2t})\| \leq n^{-1} \sum_{t=1}^{n} \|x_t\| \|z_{1t}^T$$
$$\times (\widehat{\pi}_1 - \pi_1)\| \|\Phi_{L_n}\|_1 + n^{-1} \sum_{t=1}^{n} \|x_t\| \|z_{2t}^T(\widehat{\pi}_2 - \pi_2)\| \|\Phi_{L_n}\|_1 = O_p(\|\Phi_{L_n}\|_1/\sqrt{n}), \tag{C.3}$$

for $j = 2, 3, 4$.

Next, replacing $I_t(\gamma)$ by $I_t^{(1)}(\gamma_1, \gamma_2)$ in proof of Lemma 1 of Kourtellos et al. (2022) and applying similar argument, we can show, for all $\gamma \in \Theta_\gamma$,

$$\left\| A_{n,1}^{(1)}(\gamma_1, \gamma_2) - E\left[x_t \Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I_t^{(1)}(\gamma_1, \gamma_2) \right] \right\| = O_p(\sqrt{L_n/n}).$$

Then, closely following the interval partition method in proof of Lemma 1 of Kourtellos et al. (2022), we have

$$\underset{\gamma \in \Theta_\gamma}{max} \left\| A_{n,1}^{(1)}(\gamma_1, \gamma_2) - E\left[x_t \Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I_t^{(1)}(\gamma_1, \gamma_2) \right] \right\| = o_p(1).$$

For $j = 2$, we have

$$A_{n,1}^{(2)}(\gamma_1, \gamma_2) = n^{-1} \sum_{t=1}^{n} x_t \Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I(q_{1t} \leq \gamma_1) - A_{n,1}^{(1)}(\gamma_1, \gamma_2),$$

where, closely following the proof of Lemma 1 in Kourtellos et al. (2022), we can show the first term

$$\underset{\gamma_1 \in \Theta_{\gamma_1}}{max} \left\| n^{-1} \sum_{t=1}^{n} x_t \Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I(q_{1t} \leq \gamma_1) - E\left[x_t \Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I(q_{1t} \leq \gamma_1) \right] \right\| = o_p(1).$$

A similar proof can be applied to the cases where j = 3 and 4. Hence, combining with (C.3), we have

$$\max_{\gamma \in \Theta_\gamma} \left\| n^{-1} \sum_{t=1}^{n} x_t \Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) I_t^{(j)}(\gamma_1, \gamma_2) - E\left(x_t \Phi_{L_n}^T(v_{q_1t}, v_{q_2t}) I_t^{(j)}(\gamma_1, \gamma_2) \right) \right\| = o_p(1) \tag{C.4}$$

for all $j = 1, 2, 3, 4$.

Next, note that

$$
\begin{aligned}
& n^{-1}\sum_{t=1}^{n}\Phi_{L_n}(\widehat{v}_{q_1t},\widehat{v}_{q_2t})\Phi_{L_n}^T(\widehat{v}_{q_1t},\widehat{v}_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2) = n^{-1}\sum_{t=1}^{n}\Phi_{L_n}(v_{q_1t},v_{q_2t})\Phi_{L_n}^T(v_{q_1t},v_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2) \\
& + n^{-1}\sum_{t=1}^{n}\mathbb{D}\Phi_{L_n}(\tilde{v}_{q_1t},\tilde{v}_{q_2t})\begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1}-\pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2}-\pi_{q_2})\end{bmatrix}\Phi_{L_n}^T(v_{q_1t},v_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2) \\
& + n^{-1}\sum_{t=1}^{n}\Phi_{L_n}(v_{q_1t},v_{q_2t})\begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1}-\pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2}-\pi_{q_2})\end{bmatrix}^T\mathbb{D}^T\Phi_{L_n}(\tilde{v}_{q_1t},\tilde{v}_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2) \\
& + n^{-1}\sum_{t=1}^{n}\mathbb{D}\Phi_{L_n}(\tilde{v}_{q_1t},\tilde{v}_{q_2t})\begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1}-\pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2}-\pi_{q_2})\end{bmatrix}\begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1}-\pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2}-\pi_{q_2})\end{bmatrix}^T\mathbb{D}^T\Phi_{L_n}(\tilde{v}_{q_1t},\tilde{v}_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2) \\
& = B_{n,1}^j(\gamma_1,\gamma_2)+B_{n,2}^j(\gamma_1,\gamma_2)+B_{n,2}^{jT}(\gamma_1,\gamma_2)+B_{n,4}^j(\gamma_1,\gamma_2)
\end{aligned}
$$

where, with very similar proof to $A_{n,1}^{(j)}$ and $A_{n,2}^{(j)}$, we can show, for all $\gamma\in\Theta_\gamma$,

$$
\begin{aligned}
& \left\|B_{n,1}^j(\gamma_1,\gamma_2)-E\left(\Phi_{L_n}(v_{q_1t},v_{q_2t})\Phi_{L_n}^T(v_{q1t},v_{q2t})I_t^{(j)}(\gamma_1,\gamma_2)\right)\right\| = O_p(L_n/n), \\
& \left\|B_{n,2}^j(\gamma_1,\gamma_2)\right\| = O_p(\|\Phi_{L_n}\|_1\sqrt{L_n/n}), \\
& \left\|B_{n,3}^j(\gamma_1,\gamma_2)\right\| = O_p(\|\Phi_{L_n}\|_1^2/n)
\end{aligned}
$$

for $j=1,2,3,4$.

Under Assumption 2 (f), closely following the interval partition trick used in the proof of Lemma 1 in Kourtellos et al. (2022) yields

$$
\max_{\gamma\in\Theta_\gamma}\left\|B_{n,1}^j(\gamma_1,\gamma_2)-E\left(\Phi_{L_n}(v_{q_1t},v_{q_2t})\Phi_{L_n}^T(v_{q_1t},v_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2)\right)\right\| = o_p(1).
$$

Hence, under Assumption 2 (f), we have

$$
\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\Phi_{L_n}(\widehat{v}_{q_1t},\widehat{v}_{q_2t})\Phi_{L_n}^T(\widehat{v}_{q_1t},\widehat{v}_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2)-E\left(\Phi_{L_n}(v_{q_1t},v_{q_2t})\Phi_{L_n}^T(v_{q_1t},v_{q_2t})I_t^{(j)}(\gamma_1,\gamma_2)\right)\right\| = o_p(1). \tag{C.5}
$$

Combining equation (C.5) with (C.1) and (C.4) , this completes the proof of this Lemma. ■

Lemma C.2 *Under Assumptions 1-2, we have*

$$
\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_{t,\gamma}x_t^T-E\left(\chi_{t,\gamma}^*x_t^T\right)\right\| = o_p(1), \tag{C.6}
$$

$$
\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_{t,\gamma}\eta_0^{(i)^T}(z_{1t}^T\pi_1,z_{2t}^T\pi_2)-E\left(\chi_{t,\gamma}^*\eta_0^{(i)^T}(z_{1t}^T\pi_1,z_{2t}^T\pi_2)\right)\right\| = o_p(1), \tag{C.7}
$$

for $i = 2, 3, 4$, where $\chi_{t,\gamma}$ is of dimension $4(k + L_n)$ and defined in (2.6) in section 2 and $\chi^*_{t,\gamma}$ is defined in Assumption 2 (b).

Proof. We only show (C.6) as (C.7) can be proved in the same way.

Following the proof of Lemma C.1, we can show

$$\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_t x_t^T I_t^{(j)}(\gamma_1,\gamma_2) - E\left(\chi_t^* x_t^T I_t^{(j)}(\gamma_1,\gamma_2)\right)\right\| = o_p(1),$$

for $j = 1, 2, 3, 4$. Hence, we have

$$\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_{t,\gamma}x_t^T - E\left(\chi^*_{t,\gamma}x_t^T\right)\right\| \le \max_{\gamma\in\Theta_{\gamma_1}\times\Theta_{\gamma_2}}\sum_{j=1}^{4}\left\|n^{-1}\sum_{t=1}^{n}\chi_t x_t^T I_t^{(j)}(\gamma_1,\gamma_2) - E\left(\chi_t^* x_t^T I_t^{(j)}(\gamma_1,\gamma_2)\right)\right\| = o_p(1).$$

This completes the proof of this Lemma. ■

Now, re-denote $\chi_t = \left[z_{xt}^T\widehat{\pi}_x^T, \Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})^T\right]^T$ and $\chi_t^* = \left[z_{xt}^T\pi_x^T, \Phi_{L_n}^T(v_{q_1t}, v_{q2t})\right]^T$. Let M and $M_j(\gamma_1,\gamma_2)$ re-denote the moment functionals defined as $M = E\left(\chi_t^*\chi_t^{*^T}\right)$ and $M_j(\gamma_1,\gamma_2) = E\left(\chi_t^*\chi_t^{*^T} I_t^{(j)}(\gamma_1,\gamma_2)\right)$.

Lemma C.3 *Under Assumptions 3-4, we have*

$$\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_t\chi_t^T I_t^{(j)}(\gamma_1,\gamma_2) - M_j(\gamma_1,\gamma_2)\right\| = o_p(1),$$

for $j = 1, 2, 3, 4$.

Proof. Note that $\widehat{\pi}_x - \pi_x = O_p(n^{-1/2})$. Hence, for all $\gamma_1 \in \Theta_{\gamma_1}$ and $\gamma_2 \in \Theta_{\gamma_2}$ and $j = 1, 2, 3, 4$, we have

$$n^{-1}\sum_{t=1}^{n}\widehat{x}_t\widehat{x}_t^T I_t^{(j)}(\gamma_1,\gamma_2) = n^{-1}\sum_{t=1}^{n}\pi_x z_{xt} z_{xt}^T\pi_x^T I_t^{(j)}(\gamma_1,\gamma_2) + O_p(n^{-1}),$$
$$n^{-1}\sum_{t=1}^{n}\widehat{x}_t\Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) I_t^{(j)}(\gamma_1,\gamma_2) = n^{-1}\sum_{t=1}^{n}\pi_x z_{xt}\Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) I_t^{(j)}(\gamma_1,\gamma_2) + O_p(\|\Phi_{L_n}\|_1/\sqrt{n}).$$

Following the proof of Lemma C.1 with x_t being replaced with $\pi_x z_{xt}$ completes the proof of this Lemma. ■

Lemma C.4 *Under Assumptions 3-4, we have*

$$\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_{t,\gamma}\widehat{x}_t^T - E\left(\chi_{t,\gamma}^{*} z_{xt}^T\pi_x^T\right)\right\| = o_p(1),$$
$$\max_{\gamma\in\Theta_\gamma}\left\|n^{-1}\sum_{t=1}^{n}\chi_{t,\gamma}\eta_0^{(i)^T}(v_{q_{1t}}, v_{q_{2t}}) - E\left(\chi_{t,\gamma}^{*}\eta_0^{(i)^T}(v_{q_{1t}}, v_{q_{2t}})\right)\right\| = o_p(1),$$

for $i = 2, 3, 4$, where $\chi_{t,\gamma}$ is of dimension $4(k + L_n)$ and defined in (3.3) in section 3 and $\chi_{t,\gamma}^{*}$ is defined in Assumption 4 (b).

Proof. Closely following the proof of Lemma C.2 and Lemma C.3 completes the proof of this Lemma. ■

D Proof of Theorems

Proof of Theorem 1: For simplicity, denote $\theta_j = \left[\beta_j^T, \alpha_{L_n,j}^T\right]^T$ for $j = 1, 2, 3, 4$. Denote $\widehat{\theta}_j(\gamma)$ as the least square estimator of θ_j given γ. Hence, we can rewrite the model (2.4) as

$$y_t = \sum_{j=1}^{4}\chi_t^T\theta_j I_t^{(j)}(\gamma_1^0, \gamma_2^0) + \tilde{\varepsilon}_t,$$

where

$$\tilde{\varepsilon}_t = \sum_{j=1}^{4}\left(h_j(v_{q_{1t}}, v_{q_{2t}}) - h_j^{*}(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}})\right) I_t^{(j)}(\gamma_1^0, \gamma_2^0) + \varepsilon_t.$$

Given γ, the sum of squared residuals can be written as

$$\begin{aligned} S_n(\gamma_1, \gamma_2) &= Y^TY - \widehat{Y}^T\widehat{Y} \\ &= \sum_{j=1}^{4}\left[\theta_j^T\chi^T\mathbf{I}^{(j)}(\gamma_1^0, \gamma_2^0)\chi\theta_j - \widehat{\theta}_j^T\chi^T\mathbf{I}^{(j)}(\gamma_1, \gamma_2)\chi\widehat{\theta}_j\right] + 2\sum_{j=1}^{4}\tilde{\varepsilon}^T\mathbf{I}^{(j)}(\gamma_1^0, \gamma_2^0)\chi\theta_j + \tilde{\varepsilon}^T\tilde{\varepsilon} \\ &= S_{n1}(\gamma_1, \gamma_2) + S_{n2}(\gamma_1, \gamma_2) + S_{n3}(\gamma_1, \gamma_2), \end{aligned} \tag{D.1}$$

where $\mathbf{I}^{(j)}(\gamma)$ is a $n \times n$ diagonal matrix with t^{th} diagonal element being $I^{(j)}(\gamma)$, χ stacks up χ_t^T, and $\tilde{\varepsilon}$ is an $n \times (k + L_n)$ matrix stacking up $\tilde{\varepsilon}_t$.

First, we consider $S_{n3}(\gamma_1, \gamma_2)$. Note that, for $j = 1, 2, 3, 4$,

$$h_j(v_{q_{1t}}, v_{q_{2t}}) - h_j^{*}(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}}) = \left(h_j(v_{q_{1t}}, v_{q_{2t}}) - h_j^{*}(v_{q_{1t}}, v_{q_{2t}})\right) + \left(h_j^{*}(v_{q_{1t}}, v_{q_{2t}}) - h_j^{*}(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}})\right). \tag{D.2}$$

For the first term, by Assumption 2 (d), we have

$$h_j(v_{q_1t}, v_{q_2t}) - h_j^*(v_{q_1t}, v_{q_2t}) = O(L_n^{-p}), \tag{D.3}$$

for all $j = 1, 2, 3, 4$ and $t = 1, ..., n$.

Applying first order Taylor expansion to the second term, we have

$$h_j^*(v_{q_1t}, v_{q_2t}) - h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) = \alpha_{L_n,j}^T \mathbb{D}\Phi_{L_n}(\tilde{v}_{q_1t}, \tilde{v}_{q_2t}) \begin{bmatrix} z_{1t}^T(\widehat{\pi}_{q_1} - \pi_{q_1}) \\ z_{2t}^T(\widehat{\pi}_{q_2} - \pi_{q_2}) \end{bmatrix}, \tag{D.4}$$

which implies, for $j = 1, 2, 3, 4$,

$$n^{-1} \sum_{t=1}^{n} \left[h_j^*(v_{q_1t}, v_{q_2t}) - h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) \right] = O_p(\|\Phi_{L_n}\|_1/\sqrt{n}).$$

Hence, for $j = 1, 2, 3, 4$, we have

$$n^{-1} \sum_{t=1}^{n} \left[h_j(v_{q_1t}, v_{q_2t}) - h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) \right] = O(L_n^{-p}) + O_p(\|\Phi_{L_n}\|_1/\sqrt{n}). \tag{D.5}$$

Under Assumptions 1 (a)-(c) and 2 (d)-(f) and by (D.5) we have

$$n^{-1} S_{n3}(\gamma_1, \gamma_2) = n^{-1}\varepsilon^T\varepsilon + O(L_n^{-2p}) + O_p(\|\Phi_{L_n}\|_1^2/n) = n^{-1}\varepsilon^T\varepsilon + o_p(1). \tag{D.6}$$

Next, we consider $S_{n2}(\gamma_1, \gamma_2)$. Note that, by triangular inequality, for $j = 1, 2, 3, 4$, we have

$$\sum_{j=1}^{4} \left| n^{-1}\tilde{\varepsilon}^T \mathbf{I}^{(j)}(\gamma_1^0, \gamma_2^0)\chi\theta_j \right| \leq \sum_{j=1}^{4} \left| n^{-1} \sum_{t=1}^{n} \chi_{j,t,\gamma}^T \theta_j \left[h_j(v_{q_1t}, v_{q_2t}) - h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) \right] \right| + \sum_{j=1}^{4} \left| n^{-1}\varepsilon^T\chi\theta_j \right|, \tag{D.7}$$

where $\chi_{j,t,\gamma}^T$ is defined in section 2.1

For the first term of (D.7), by (D.5), we have

$$\begin{aligned} \sum_{j=1}^{4} \left| n^{-1} \sum_{t=1}^{n} \chi_{j,t,\gamma}^T \theta_j \left[h_j(v_{q_1t}, v_{q_2t}) - h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) \right] \right| &\leq O_p(L_n^{-p} + \|\Phi_{L_n}\|_1/\sqrt{n}) \\ \times \sum_{j=1}^{4} \left(n^{-1} \sum_{t=1}^{n} \left| x_t^T\beta_j \right| + n^{-1} \sum_{t=1}^{n} \left| h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) \right| \right). \end{aligned}$$

where $\sum_{j=1}^{4} \left| n^{-1} \sum_{t=1}^{n} x_t^T\beta_j \right| = O_p(1)$ under Assumption 2 (a), and

$$n^{-1} \sum_{t=1}^{n} h_j^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t}) = n^{-1} \sum_{t=1}^{n} h_j(v_{q_1t}, v_{q_2t}) + O_p(L_n^{-p} + \|\Phi_{L_n}\|_1/\sqrt{n}) = O_p(1) \tag{D.8}$$

under Assumptions 2 (d)-(f) for $j = 1, 2, 3, 4$. Hence, we have

$$\sum_{j=1}^{4} \left| n^{-1} \sum_{t=1}^{n} \chi_{j,t,\gamma}^{T} \theta_j \left[h_j(v_{q_{1t}}, v_{q_{2t}}) - h_j^*(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}}) \right] \right| = O_p\left(L_n^{-p} + \|\Phi_{L_n}\|_1/\sqrt{n} \right),$$

Similarly, under Assumptions 1 (c), 2 (a), and 2 (d), we can show, for $j = 1, 2, 3, 4$,

$$\left| n^{-1} \varepsilon^T \chi \theta_j \right| \leq n^{-1} \sum_{t=1}^{n} \left| x_t^T \beta_j \varepsilon_t \right| + n^{-1} \sum_{t=1}^{n} \left| h_j^*(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}}) \varepsilon_t \right| = O_p(n^{-1/2}).$$

To sum up, we have

$$n^{-1} S_{n2}(\gamma_1, \gamma_2) = O_p\left(L_n^{-p} + n^{-1/2} \|\Phi_{L_n}\|_1 + n^{-1/2} \right) = o_p(1). \tag{D.9}$$

where the second equality holds under Assumption 2 (f).

Next, to find the probability limit of $S_{n1}(\gamma_1, \gamma_2)$, we partition the threshold space into four regimes, which is closely related to the proof of Theorem 1 in Chen et al. (2012). Below, we only show the first regime as the remaining three cases can be obtained by analogy.

Case 1: $\gamma_1 \leq \gamma_1^0$ and $\gamma_2 \leq \gamma_2^0$.

Note that, by (D.5) and under Assumptions 1 (c) and 2 (a), we have

$$\begin{aligned} &\left\| n^{-1} \tilde{\varepsilon}^T \chi \chi^T \tilde{\varepsilon} \right\| \leq \sum_{j=1}^{4} \left\| h_j(v_{q_{1t}}, v_{q_{1t}}) - h_j^*(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}}) \right\|^2 + \left\| n^{-1} \chi^T \varepsilon \right\|^2 \\ &= O_p\left(L_n^{-2p} + \|\Phi_{L_n}\|_1^2/n + L_n/n \right), \end{aligned}$$

which implies, under Assumption 2 (f) for $j = 1, 2, 3, 4$,

$$\left\| n^{-1} \chi^T \mathbf{I}^{(j)}(\gamma_1, \gamma_2) \tilde{\varepsilon} \right\| \leq \left\| n^{-1} \chi^T \tilde{\varepsilon} \right\| = O_p\left(L_n^{-p} + \|\Phi_{L_n}\|_1/\sqrt{n} + \sqrt{L_n/n} \right) = o_p(1).$$

Hence, under Assumptions 1-2 and by Lemma C.1, we can show

$$\begin{aligned} &\widehat{\theta}_1 = \left[\chi^T \mathbf{I}^{(1)}(\gamma_1, \gamma_2) \chi \right]^{-1} \chi^T \mathbf{I}^{(1)}(\gamma_1, \gamma_2) \mathbf{y} = \theta_1 + \left[n^{-1} \chi^T \mathbf{I}^{(1)}(\gamma_1, \gamma_2) \chi \right]^{-1} \left[n^{-1} \chi^T \mathbf{I}^{(1)}(\gamma_1, \gamma_2) \tilde{\varepsilon} \right] \\ &= \theta_1 + o_p(1), \end{aligned} \tag{D.10}$$

$$\begin{aligned} &\widehat{\theta}_2 = \left[\chi^T \mathbf{I}^{(2)}(\gamma_1, \gamma_2) \chi \right]^{-1} \chi^T \mathbf{I}^{(2)}(\gamma_1, \gamma_2) \mathbf{y} = \theta_2 + \left[\chi^T \mathbf{I}^{(2)}(\gamma_1, \gamma_2) \chi \right]^{-1} \chi^T \left[\mathbf{I}^{(2)}(\gamma_1, \gamma_2) - \mathbf{I}^{(2)}(\gamma_1, \gamma_2^0) \right] \\ &\times \chi\left(\theta_1 - \theta_2 \right) + \left[\chi^T \mathbf{I}^{(2)}(\gamma_1, \gamma_2) \chi \right]^{-1} \chi^T \mathbf{I}^{(2)}(\gamma_1, \gamma_2) \tilde{\varepsilon} = \theta_2 + M_2(\gamma_1, \gamma_2)^{-1} \left(M_2(\gamma_1, \gamma_2) - M_2(\gamma_1, \gamma_2^0) \right) \\ &\times \left(\theta_1 - \theta_2 \right) + o_p(1), \end{aligned} \tag{D.11}$$

$$\widehat{\theta}_3 = \left[\chi^T \mathbf{I}^{(3)}(\gamma_1,\gamma_2)\chi\right]^{-1} \chi^T \mathbf{I}^{(3)}(\gamma_1,\gamma_2)\mathbf{y} = \theta_2 + M_3(\gamma_1,\gamma_2)^{-1} M_3(\gamma_1^0,\gamma_2)\,(\theta_3-\theta_2) + M_3(\gamma_1,\gamma_2)^{-1}$$
$$\times \left(M_3(\gamma_1,\gamma_2) - M_3(\gamma_1^0,\gamma_2)\right)(\theta_1-\theta_2) + o_p(1), \tag{D.12}$$

$$\widehat{\theta}_4 = \left[\chi^T \mathbf{I}^{(4)}(\gamma_1,\gamma_2)\chi\right]^{-1} \chi^T \mathbf{I}^{(4)}(\gamma_1,\gamma_2)\mathbf{y} = \theta_2 + M_4(\gamma_1,\gamma_2)^{-1}\big(M_4(\gamma_1,\gamma_2) - M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1,\gamma_2^0)$$
$$+ M_4(\gamma_1^0,\gamma_2^0)\big)(\theta_1-\theta_2) + M_4(\gamma_1,\gamma_2)^{-1}\left(M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1^0,\gamma_2^0)\right)(\theta_3-\theta_2) + M_4(\gamma_1,\gamma_2)^{-1}$$
$$\times M_4(\gamma_1^0,\gamma_2^0) + o_p(1). \tag{D.13}$$

Note that equations (D.10)-(D.13) imply

$$\sum_{j=1}^{4} \widehat{\theta}_j^T M_j(\gamma_1,\gamma_2)\theta_2 = \theta_1^T M_1(\gamma_1,\gamma_2)\theta_2 + (\theta_1-\theta_2)^T\left[M_2(\gamma_1,\gamma_2) - M_2(\gamma_1,\gamma_2^0)\right]\theta_2 + \theta_2^T M_2(\gamma_1,\gamma_2)\theta_2$$
$$+ (\theta_1-\theta_2)^T\left[M_3(\gamma_1,\gamma_2) - M_3(\gamma_1^0,\gamma_2)\right]\theta_2 + (\theta_3-\theta_2)^T M_3(\gamma_1^0,\gamma_2)\theta_2 + \theta_2^T M_2(\gamma_1,\gamma_2)\theta_2$$
$$+ (\theta_1-\theta_2)^T\left[M_4(\gamma_1^0,\gamma_2^0) - M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1,\gamma_2^0) + M_4(\gamma_1,\gamma_2)\right]\theta_2 + (\theta_4-\theta_3)^T M_4(\gamma_1^0,\gamma_2^0)\theta_2$$
$$+ (\theta_3-\theta_2)^T\left[M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1^0,\gamma_2^0)\right]\theta_2 + \theta_2^T M_4(\gamma_1,\gamma_2)\theta_2 + o_p(1)$$
$$= \theta_1^T\left[M_1(\gamma_1,\gamma_2) + M_2(\gamma_1,\gamma_2) - M_2(\gamma_1,\gamma_2^0) + M_3(\gamma_1,\gamma_2) - M_3(\gamma_1^0,\gamma_2)\right]\theta_2$$
$$+ \theta_2^T M_2(\gamma_1,\gamma_2)\theta_2$$
$$+ \theta_3\left[M_3(\gamma_1^0,\gamma_2) + M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1^0,\gamma_2^0)\right]\theta_2$$
$$+ \theta_4^T M_4(\gamma_1^0,\gamma_2^0)\theta_2 + o_p(1)$$
$$= \sum_{j=1}^{4} \theta_j M_j(\gamma_1^0,\gamma_2^0)\theta_2 + o_p(1).$$

Hence, applying simple calculation, we have

$$
\begin{aligned}
&n^{-1}S_{n1}(\gamma_1,\gamma_2) = \sum_{j=1}^{4}\theta_j^T M_j(\gamma_1^0,\gamma_2^0)\left(\theta_j-\theta_2\right) + \sum_{j=1}^{4}\widehat{\theta}_j^T M_j(\gamma_1,\gamma_2)\theta_2 - \sum_{j=1}^{4}\widehat{\theta}_j^T M_j(\gamma_1,\gamma_2)\widehat{\theta}_j + o_p(1)\\
&= (\theta_1-\theta_2)^T\Big\{M_1(\gamma_1^0,\gamma_2^0) - M_1(\gamma_1,\gamma_2) - M_2(\gamma_1,\gamma_2)^{-1}\left[M_2(\gamma_1,\gamma_2)-M_2(\gamma_1^0,\gamma_2^0)\right]\left[M_2(\gamma_1,\gamma_2)-M_2(\gamma_1^0,\gamma_2^0)\right]^T\\
&- M_3(\gamma_1,\gamma_2)^{-1}\left[M_3(\gamma_1,\gamma_2)-M_3(\gamma_1^0,\gamma_2)\right]\left[M_3(\gamma_1,\gamma_2)-M_3(\gamma_1^0,\gamma_2)\right]^T - M_4(\gamma_1,\gamma_2)^{-1}\big[M_4(\gamma_1^0,\gamma_2^0)-M_4(\gamma_1^0,\gamma_2)\\
&- M_4(\gamma_1,\gamma_2^0)+M_4(\gamma_1,\gamma_2)\big]\left[M_4(\gamma_1^0,\gamma_2^0)-M_4(\gamma_1^0,\gamma_2)-M_4(\gamma_1,\gamma_2^0)+M_4(\gamma_1,\gamma_2)\right]^T\Big\}(\theta_1-\theta_2)\\
&+ (\theta_3-\theta_2)^T\Big\{M_3(\gamma_1^0,\gamma_2^0) - M_4(\gamma_1,\gamma_2)^{-1}\left[M_4(\gamma_1^0,\gamma_2)-M_4(\gamma_1^0,\gamma_2^0)\right]\left[M_4(\gamma_1^0,\gamma_2)-M_4(\gamma_1^0,\gamma_2^0)\right]^T\\
&- M_3(\gamma_1,\gamma_2)^{-1}M_3(\gamma_1^0,\gamma_2)M_3(\gamma_1^0,\gamma_2)^T\Big\}(\theta_3-\theta_2)\\
&+ (\theta_4-\theta_2)^T\left\{M_4(\gamma_1^0,\gamma_2^0) - M_4(\gamma_1,\gamma_2)^{-1}M_4(\gamma_1^0,\gamma_2^0)M_4(\gamma_1^0,\gamma_2^0)^T\right\}(\theta_4-\theta_2) + o_p(1)\\
&= (\theta_1-\theta_2)^T Q_1\,(\theta_1-\theta_2) + (\theta_3-\theta_2)\,Q_2\,(\theta_3-\theta_2) + (\theta_4-\theta_2)^T Q_3\,(\theta_4-\theta_2) + o_p(1)\\
&= S_1(\gamma_1,\gamma_2) + o_p(1).
\end{aligned}
$$

Note that,

$$
\begin{aligned}
&M_1(\gamma_1^0,\gamma_2^0) - M_1(\gamma_1,\gamma_2) = M_2(\gamma_1,\gamma_2) - M_2(\gamma_1^0,\gamma_2^0) + M_3(\gamma_1,\gamma_2) - M_3(\gamma_1^0,\gamma_2)\\
&+ M_4(\gamma_1^0,\gamma_2^0) - M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1,\gamma_2^0) + M_4(\gamma_1,\gamma_2),
\end{aligned}
$$

and

$$
M_3(\gamma_1^0,\gamma_2^0) = M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1^0,\gamma_2^0) + M_3(\gamma_1^0,\gamma_2).
$$

Also, for $\gamma_1 \le \gamma_1^0$ and $\gamma_2 \le \gamma_2^0$, we have

$$
\begin{aligned}
&M_2(\gamma_1,\gamma_2) - M_2(\gamma_1^0,\gamma_2^0) \preceq M_2(\gamma_1,\gamma_2),\\
&M_3(\gamma_1,\gamma_2) - M_3(\gamma_1^0,\gamma_2) \preceq M_3(\gamma_1,\gamma_2),\\
&M_4(\gamma_1^0,\gamma_2^0) - M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1,\gamma_2^0) + M_4(\gamma_1,\gamma_2) \preceq M_4(\gamma_1,\gamma_2),\\
&M_4(\gamma_1^0,\gamma_2) - M_4(\gamma_1^0,\gamma_2^0) \preceq M_4(\gamma_1,\gamma_2),\\
&M_3(\gamma_1^0,\gamma_2 \preceq M_3(\gamma_1,\gamma_2),\\
&M_4(\gamma_1^0,\gamma_2^0) \preceq M_4(\gamma_1,\gamma_2).
\end{aligned}
$$

Therefore, we can show that Q_1, Q_2, and Q_3 are all positive semi-definite matrices. Thus, we have

$$
S_1(\gamma_1,\gamma_2) \ge S_1(\gamma_1^0,\gamma_2^0) = 0. \tag{D.14}
$$

Similarly, for all four regimes ($j = 1, 2, 3, 4$) and all $\gamma \in \Theta_\gamma$,

$$n^{-1}S_n(\gamma_1, \gamma_2) = S_1^j(\gamma_1, \gamma_2) + o_p(1),$$

where $S_1^j(\gamma_1, \gamma_2)$ is a continuous function of γ and is uniquely minimized at γ_0. Then, following Theorem 2.1 in Newey & Mcfadden (1994), we have $\widehat{\gamma} \xrightarrow{p} \gamma_0$.

Now, we verify that $\|\widehat{\theta} - \theta\| = O_p(L_n^{-p} + n^{-1/2}\|\Phi_{L_n}\|_1 + \sqrt{L_n/n} + n^{-min(\alpha,\varrho)})$. By definition, we have

$$\widehat{\theta}(\gamma_1, \gamma_2) - \theta = (n^{-1}\chi_\gamma^T\chi_\gamma)^{-1}n^{-1}\chi_\gamma^T(\mathbf{y} - \chi_\gamma\theta),$$

where

$$\mathbf{y} - \chi_\gamma\theta = h_1(v_{q_1}, v_{q_2}) - h_1^*(\widehat{v}_{q_1}, \widehat{v}_{q_2}) + n^{-\rho}\sum_{j=2}^{4}\mathbf{I}^{(j)}(\gamma_1, \gamma_2)\left[\eta_0^{(j)}(v_{q_1}, v_{q_2}) - \eta_0^{*(j)}(\widehat{v}_{q_1}, \widehat{v}_{q_2})\right]$$
$$+ \varepsilon + \sum_{j=2}^{4}\left[\mathbf{I}^{(j)}(\gamma_1^0, \gamma_2^0) - \mathbf{I}^{(j)}(\gamma_1, \gamma_2)\right]\left[n^{-\alpha}x\delta_0^{(j)} + n^{-\rho}\eta_0^{(j)}(v_{q_1}, v_{q_2})\right],$$

where, x is an $n \times k$ matrix stacking up x_t^T, $h_1(v_{q_1}, v_{q_2})$, $h_1^*(\widehat{v}_{q_1}, \widehat{v}_{q_2})$, $\eta_n^{(j)}(v_{q_1}, v_{q_2})$, and $\eta_n^{*(j)}(\widehat{v}_{q_1}, \widehat{v}_{q_2})$ stack up $h_1(v_{q_1t}, v_{q_2t})$, $h_1^*(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})$, $\eta_n^{(j)}(v_{q_1t}, v_{q_2t})$, and $\eta_n^{*(j)}(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})$ respectively for $j = 2, 3, 4$.

Let $\|A\|_{sp}$ denote the spectral norm for matrix A. As $\|.\|_{sp}$ is submultiplicative, we have

$$\left\|\widehat{\theta}(\gamma_1, \gamma_2) - \theta\right\|_{sp} = \left\|\left(n^{-1}\chi_\gamma^T\chi_\gamma\right)^{-1}n^{-1}\chi_\gamma^T\left(\mathbf{y} - \chi_\gamma\theta\right)\right\|_{sp}$$
$$\leq \left\|\left(n^{-1}\chi_\gamma^T\chi_\gamma\right)^{-1/2}\right\|_{sp}\left\|\left(n^{-1}\chi_\gamma^T\chi_\gamma\right)^{-1/2}n^{-1}\chi_\gamma^T\left(\mathbf{y} - \chi_\gamma\theta\right)\right\|_{sp}. \tag{D.15}$$

First, under Assumption 2 (b) and by Lemma C.1, for all $\gamma \in \Theta_\gamma$, we have

$$\left\|\left(n^{-1}\chi_\gamma^T\chi_\gamma\right)^{-1/2}\right\|_{sp} = \left\|E^{-1/2}\left(\chi_{t,\gamma}^*\chi_{t,\gamma}^{*T}\right)\right\|_{sp} + o_p(1) = \lambda_{max}^{1/2}\left(E^{-1}\left(\chi_{t,\gamma}^*\chi_{t,\gamma}^{*T}\right)\right) + o_p(1) = O_p(1). \tag{D.16}$$

Next, similar to the proof of Lemma 3 in Kourtellos et al. (2022), under Assumptions 1-2, we have, for all $\gamma \in \Theta_\gamma$,

$$n^{-1}\varepsilon^T\chi_\gamma\left(\chi_\gamma^T\chi_\gamma\right)^{-1}\chi_\gamma^T\varepsilon = O_p(L_n/n). \tag{D.17}$$

Similar to the proof of Lemma C.1, under Assumption 2 (a), we have, for all $\gamma \in \Theta_\gamma$

$$n^{-1}\sum_{j=2}^{4}\left[\mathbf{I}^{(j)}(\gamma_1^0, \gamma_2^0) - \mathbf{I}^{(j)}(\gamma_1, \gamma_2)\right]\left[n^{-\alpha}x\delta_0^{(j)} + n^{-\rho}\eta_0^{(j)}(v_{q_1}, v_{q_2})\right]$$
$$\leq n^{-1}\sum_{j=2}^{4}\left[n^{-\alpha}x\delta_0^{(j)} + n^{-\rho}\eta_0^{(j)}(v_{q_1}, v_{q_2})\right] = O_p\left(n^{-\alpha} + n^{-\rho}\right). \tag{D.18}$$

Hence, by equations (D.2), (D.16), (D.17), and (D.18) and closely following the proof of Lemma 4 in Kourtellos et al. (2022), we have

$$\left\|\widehat{\theta}(\gamma_1,\gamma_2)-\theta\right\|_{sp} \leq O_p(1)\lambda_{max}^{1/2}\left\{n^{-1}\left(\mathbf{y}-\chi_\gamma\theta\right)^T\chi_\gamma\left(\chi_\gamma^T\chi_\gamma\right)^{-1}\chi_\gamma^T\left(\mathbf{y}-\chi_\gamma\theta\right)\right\}$$
$$= O_p\left(L_n^{-p}+n^{-1/2}\|\Phi_{L_n}\|_1+\sqrt{L_n/n}+n^{-min(\alpha,\varrho)}\right).$$

This completes the proof of the Theorem.

Proof of Theorem 2: In the matrix form, we have

$$\mathbf{y}-\chi_\gamma\widehat{\theta}(\gamma) = x\beta_1+\sum_{j=2}^{4}x_{\gamma_0}^{(j)}\delta_n^{(j)}+h_1(v_{q_1},v_{q_2})+\sum_{j=2}^{4}\eta_{n,\gamma_0}^{(j)}(v_{q_1},v_{q_2})+\varepsilon$$
$$-x\widehat{\beta}_1-\sum_{j=2}^{4}x_\gamma^{(j)}\widehat{\delta}_n^{(j)}-h_1^*(\widehat{v}_{q_1},\widehat{v}_{q_2})-\sum_{j=2}^{4}\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})$$
$$=\varepsilon+\Delta_n-\sum_{j=2}^{4}\Delta x_\gamma^{(j)}\widehat{\delta}_n^{(j)}-\sum_{j=2}^{4}\Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2}),$$

where $\Delta_n = x\left(\beta_1-\widehat{\beta}_1\right)+\sum_{j=2}^4 x_{\gamma_0}^{(j)}\left(\delta_n^{(j)}-\widehat{\delta}_n^{(j)}\right)+h_1(v_{q_1},v_{q_2})-h_1^*\left(\widehat{v}_{q_1},\widehat{v}_{q_2}\right)+\sum_{j=2}^4\left(\eta_{n,\gamma_0}^{(j)}(v_{q_1},v_{q_2})-\eta_{n,\gamma_0}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2}\right)$, $\Delta x_\gamma^{(j)}=x_\gamma^{(j)}-x_{\gamma_0}^{(j)}$, $\Delta\widehat{\eta}_{n,\gamma}^{*(j)}=\widehat{\eta}_{n,\gamma}^{*(j)}-\widehat{\eta}_{n,\gamma_0}^{*(j)}$, $x_\gamma^{(j)}$ stacks up $x_t^T I_t^{(j)}(\gamma)$, and $\eta_{n,\gamma}^{(j)}(v_{q_1},v_{q_2})$ stacks up $\eta_{n,\gamma}^{(j)}(v_{q_1t},v_{q_2t})I_t^{(j)}(\gamma)$.

Note that for all $j=2,3,4,$, we have $\Delta x_{\gamma_0}^{(j)}=0$ and $\Delta\eta_{n,\gamma_0}^{*(j)}=0$. For $i\neq j$, we have $\left(x_\gamma^{(j)}\widehat{\delta}_n^{(j)}\right)^T\left(x_\gamma^{(i)}\widehat{\delta}_n^{(i)}\right)=0$, $\left[\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right]^T\left[\eta_{n,\gamma}^{*(i)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right]=0$. Let $\widehat{\kappa}_{n,j}=\left[\widehat{\delta}_n^T,n^{-\varrho}\widehat{\alpha}_{L_n,0,j}^T\right]^T$ and $\kappa_{n,j}=\left[n^{-\alpha}\delta_0^T,n^{-\varrho}\alpha_{L_n,0,j}^T\right]^T$ for $j=2,3,4$.

Now, let $\gamma_1=\gamma_1^0+\frac{v}{a_n}$ and $\gamma_2=\gamma_2^0+\frac{w}{a_n}$. Similar to the proof of Theorem 1 of Chen et al. (2012), we partition the model into four cases and show the proof for the case where $v>0$ and $w>0$. The remaining three cases can be proved accordingly.

By simple calculation, we have

$$
\begin{aligned}
&\Delta x_{\gamma}^{(2)}\widehat{\delta}_n^{(2)} + \Delta\eta_{n,\gamma}^{*(2)}(\widehat{v}_{q_1},\widehat{v}_{q_2}) \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,2}^{T}\chi_t\left[I_t^{(2)}(\gamma) - I_t^{(2)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,2}^{T}\chi_t\left[I_t^{(2)}(\gamma) - I_t^{(2)}(\gamma_1^0,\gamma_2) + I_t^{(2)}(\gamma_1^0,\gamma_2) - I_t^{(2)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,2}^{T}\chi_t\left[I(\gamma_1^0 \le q_{1t} < \gamma_1, \gamma_2 \le q_{2t}) - I(q_{1t} < \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\right],
\end{aligned}
$$

$$
\begin{aligned}
&\Delta x_{\gamma}^{(3)}\widehat{\delta}_n^{(3)} + \Delta\eta_{n,\gamma}^{*(3)}(\widehat{v}_{q_1},\widehat{v}_{q_2}) \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,3}^{T}\chi_t\left[I_t^{(2)}(\gamma) - I_t^{(2)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\kappa_{n,3}^{T}\chi_t\left[I_t^{(3)}(\gamma) - I_t^{(3)}(\gamma_1^0,\gamma_2) + I_t^{(3)}(\gamma_1^0,\gamma_2) - I_t^{(3)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,3}^{T}\chi_t\left[-I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} < \gamma_2) + I(\gamma_1^0 \le q_{1t}, \gamma_2^0 \le q_{2t} < \gamma_2)\right],
\end{aligned}
$$

$$
\begin{aligned}
&\Delta x_{\gamma}^{(4)}\widehat{\delta}_n^{(4)} + \Delta\eta_{n,\gamma}^{*(4)}(\widehat{v}_{q_1},\widehat{v}_{q_2}) \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,4}^{T}\chi_t\left[I_t^{(4)}(\gamma) - I_t^{(4)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,4}^{T}\chi_t\left[I_t^{(4)}(\gamma) - I_t^{(4)}(\gamma_1^0,\gamma_2) + I_t^{(4)}(\gamma_1^0,\gamma_2) - I_t^{(4)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\widehat{\kappa}_{n,4}^{T}\chi_t\left[-I(\gamma_1^0 \le q_{1t} < \gamma_1, \gamma_2 \le q_{2t}) - I(\gamma_1^0 \le q_{1t}, \gamma_2^0 \le q_{2t} < \gamma_2)\right].
\end{aligned}
$$

Summing up, we have

$$
\begin{aligned}
&\sum_{j=2}^{4}\left(\Delta x_{\gamma}^{(j)}\widehat{\delta}_n^{(j)} + \Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right) \\
&= \sum_{t=1}^{n}\Big\{\left(\widehat{\kappa}_{n,2} - \widehat{\kappa}_{n,4}\right)^{T}\chi_t I(\gamma_1^0 \le q_{1t} < \gamma_1, \gamma_2 \le q_{2t}) - \widehat{\kappa}_{n,2}^{T}\chi_t I(q_{1t} < \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2) \\
&- \widehat{\kappa}_{n,3}^{T}\chi_t I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} < \gamma_2) + \left(\widehat{\kappa}_{n,3} - \widehat{\kappa}_{n,4}\right)^{T}\chi_t I(\gamma_1^0 \le q_{1t}, \gamma_2^0 \le q_{2t} < \gamma_2)\Big\}.
\end{aligned}
$$

Since the four terms are orthogonal to each other, we can show the centered process as

$$
\begin{aligned}
& S_n(\gamma) - S_n(\gamma_0) \\
&= \left[\sum_{j=2}^{4}\left(\Delta x_\gamma^{(j)}\widehat{\delta}_n^{(j)} + \Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1}, \widehat{v}_{q_2})\right)\right]^T \left[\sum_{j=2}^{4}\left(\Delta x_\gamma^{(j)}\widehat{\delta}_n^{(j)} + \Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1}, \widehat{v}_{q_2})\right)\right] \\
&- 2(\varepsilon + \Delta_n)^T \left[\sum_{j=2}^{4}\Delta x_\gamma^{(j)}\widehat{\delta}_n^{(j)} + \sum_{j=2}^{4}\Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1}, \widehat{v}_{q_2})\right] \\
&= \sum_{t=1}^{n}\Big\{ [\kappa_{n,2} - \kappa_{n,4}]^T \chi_t\chi_t^T [\kappa_{n,2} - \kappa_{n,4}] I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2) + \kappa_{n,2}^T\chi_t\chi_t^T\kappa_{n,2} I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2) \\
&+ \kappa_{n,3}^T\chi_t\chi_t^T\kappa_{n,3} I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} \le \gamma_2) + [\kappa_{n,3} - \kappa_{n,4}]^T \chi_t\chi_t^T [\kappa_{n,3} - \kappa_{n,4}] I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Big\} \\
&- 2\sum_{j=2}^{4}\widehat{\kappa}_{n,j}\left[\sum_{t=1}^{n}\varepsilon_t\chi_t\left(I_t^{(j)}(\gamma_1,\gamma_2) - I_t^{(j)}(\gamma_1^0,\gamma_2^0)\right)\right] \\
&- 2\sum_{j=2}^{4}\widehat{\kappa}_{n,j}^T\left[\sum_{t=1}^{n}\Delta_{n_t}\chi_t\left(I_t^{(j)}(\gamma_1,\gamma_2) - I_t^{(j)}(\gamma_1^0,\gamma_2^0)\right)\right] \\
&+ \sum_{t=1}^{n}\Big\{ [(\widehat{\kappa}_{n,2} - \widehat{\kappa}_{n,4}) - (\kappa_{n,2} - \kappa_{n,4})]^T \chi_t\chi_t^T [(\widehat{\kappa}_{n,2} - \widehat{\kappa}_{n,4}) + (\kappa_{n,2} - \kappa_{n,4})] I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2) \\
&+ [\widehat{\kappa}_{n,2} - \kappa_{n,2}]^T \chi_t\chi_t^T [\widehat{\kappa}_{n,2} + \kappa_{n,2}] I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2) \\
&+ [\widehat{\kappa}_{n,3} - \kappa_{n,3}]^T \chi_t\chi_t^T [\widehat{\kappa}_{n,3} + \kappa_{n,3}] I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} \le \gamma_2) \\
&+ [(\widehat{\kappa}_{n,3} - \widehat{\kappa}_{n,4}) - (\kappa_{n,3} - \kappa_{n,4})]^T \chi_t\chi_t^T [(\widehat{\kappa}_{n,3} - \widehat{\kappa}_{n,4}) + (\kappa_{n,3} - \kappa_{n,4})] I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Big\} \\
&= S_{n,1}^*(\gamma) - 2S_{n,2}^*(\gamma) - 2S_{n,3}^*(\gamma) + S_{n,4}^*(\gamma),
\end{aligned} \tag{D.19}
$$

where Δ_{n_t} is the t^{th} element of Δ_n.

To examine the limiting behavior of $S_n(\gamma) - S_n(\gamma_0)$, we study the asymptotics of $S_{n,i}^*(\gamma)$, for $i = 1, 2, 3, 4$.

First, closely following the proof of Theorem 2 in Kourtellos et al. (2022) and Lemma A.9 in Hansen (2000), we can show that,

$$a_n(\widehat{\gamma} - \gamma_0) = O_p(1),$$

where $a_n = n^{1-2min(\alpha,\varrho)}$.

Next, we consider $S_{n,1}^*(\gamma)$. Applying simple calculation, we have
$\mathrm{S}_{n,1}^*(\gamma) = \sum_{t=1}^{n}\Big\{ \left[(\kappa_{n,2} - \kappa_{n,4})^T \chi_t^*\right]^2 I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2) + \left(\kappa_{n,2}^T\chi_t^*\right)^2 I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} <$

$$\gamma_2) + \left(\kappa_{n,3}^T \chi_t^*\right)^2 I(\gamma_1^0 \le q_{1t} < \gamma_{1,2t} \le \gamma_2) + \left[(\kappa_{n,3} - \kappa_{n,4})^T \chi_t^*\right]^2 I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Bigg\}$$

$$+ 2\sum_{t=1}^n \Bigg\{ \left[\kappa_{n,2} - \kappa_{n,4}\right]^T (\chi_t - \chi_t^*)\chi_t^{*^T} \left[\kappa_{n,2} - \kappa_{n,4}\right] I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2)$$

$$+ \kappa_{n,2}^T (\chi_t - \chi_t^*) \chi_t^{*^T} \kappa_{n,2} I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2) + \kappa_{n,3}^T (\chi_t - \chi_t^*) \chi_t^{*^T} \kappa_{n,3} I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} \le \gamma_2)$$

$$+ \left[\kappa_{n,3} - \kappa_{n,4}\right]^T (\chi_t - \chi_t^*) \chi_t^{*^T} \left[\kappa_{n,3} - \kappa_{n,4}\right] I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Bigg\}$$

$$+ \sum_{t=1}^n \Bigg\{ \left[(\kappa_{n,2} - \kappa_{n,4})^T (\chi_t - \chi_t^*)\right]^2 I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2) + \left[\kappa_{n,2}^T (\chi_t - \chi_t^*)\right]^2 I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)$$

$$+ \left[\kappa_{n,3}^T (\chi_t - \chi_t^*)\right]^2 I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} \le \gamma_2) + \left[(\kappa_{n,3} - \kappa_{n,4})^T (\chi_t - \chi_t^*)\right]^2 I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Bigg\}$$

$$= \sum_{j=1}^4 G_n^j(\gamma) + 2n^{-2\varrho} \sum_{t=1}^n \Bigg\{ (\alpha_{L_n,0,2} - \alpha_{L_n,0,4})^T \left[\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2})\right] \times \left(\eta_0^{*(2)} - \eta_0^{*(4)}\right) I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2)$$

$$- \alpha_{L_n,0,2}^T (\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2})) \eta_0^{*(2)} I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)$$

$$+ \alpha_{L_n,0,2}^T \left[\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2})\right] \eta_0^{*(2)} I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)$$

$$+ \alpha_{L_n,0,3}^T \left[\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2})\right] \eta_0^{*(3)} I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} \le \gamma_2)$$

$$+ (\alpha_{L_n,0,3} - \alpha_{L_n,0,4})^T \left[\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2})\right] \left(\eta_0^{*(3)} - \eta_0^{*(4)}\right) I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Bigg\}$$

$$+ n^{-2\varrho} \sum_{t=1}^n \Bigg\{ \left[(\alpha_{L_n,0,2} - \alpha_{L_n,0,4})^T (\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2}))\right]^2 I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} > \gamma_2)$$

$$+ \left[\alpha_{L_n,0,2}^T (\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2}))\right]^2 I(q_{1t} \le \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)$$

$$+ \left[\alpha_{L_n,0,3}^T (\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2}))\right]^2 I(\gamma_1^0 \le q_{1t} < \gamma_1, q_{2t} \le \gamma_2)$$

$$+ \left[(\alpha_{L_n,0,3} - \alpha_{L_n,0,4})^T (\Phi_{L_n}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \Phi_{L_n}(v_{q_1}, v_{q_2}))\right]^2 I(q_{1t} > \gamma_1^0, \gamma_2^0 \le q_{2t} < \gamma_2)\Bigg\}$$

$$= \sum_{j=1}^4 G_n^j(\gamma) + \sum_{j=1}^4 A_{1,n}^{(j)}(\gamma) + \sum_{j=1}^4 A_{2,n}^{(j)}(\gamma) = \sum_{j=1}^4 G_n^j(\gamma) + \sum_{j=1}^4 A_n^{(j)}(\gamma).$$

For any $v \in (0, \bar{v}]$ and $w \in (0, \bar{w}]$, applying the first order Taylor expansion, we have

$$E\left[\sum_{j=1}^4 G_n^j(v, w)\right] = \Bigg\{ v n^{2min(\alpha,\varrho)} \left(E\left[\left((\kappa_{n,2} - \kappa_{n,4})^T \chi_t^*\right)^2 + \left(\kappa_{n,3}^T \chi_t^*\right)^2 \Big| q_t = \gamma_0 \right] \right) f_1^0$$

$$+ w n^{2min(\alpha,\varrho)} \left(E\left[\left(\kappa_{n,2}^T \chi_t^*\right)^2 + \left((\kappa_{n,3} - \kappa_{n,4})^T \chi_t^*\right)^2 \Big| q_t = \gamma_0 \right] \right) f_2^0 \Bigg\} (1 + o(1)),$$

where $f_i^0 = f_i(\gamma_1^0, \gamma_2^2)$ is defined in Assumption 2 (c) for $i = 1, 2$. Closely following the proof of Lemma 5 of Kourtellos et al. (2022), we can show

$$E\left\{\sum_{j=1}^4 G_n^j(v, w) - E\left[\sum_{j=1}^4 G_n^j(v, w)\right]\right\}^2 = o(1). \tag{D.20}$$

Hence, we have

$$\sum_{j=1}^{4} G_n^j(v,w) = v\mu_1 + w\mu_2 + o_p\left(1\right), \tag{D.21}$$

where

$$\begin{aligned}\mu_1 &= \left[d^T D^* d - 2d_2^T D^{(2,4)} d_4\right] f_1^0, \\ \mu_2 &= \left[d^T D^* d - 2d_3^T D^{(3,4)} d_4\right] f_2^0,\end{aligned} \tag{D.22}$$

$d = \left[d_2^T, d_3^T, d_4^T\right]^T$ with $d_2 = \left[\delta_0^{(2)T}, 1\right]^T$, $d_3 = \left[\delta_0^{(3)T}, 1\right]^T$, $d_4 = \left[\delta_0^{(4)T}, 1\right]^T$, $D^* = diag\left\{D^{(2,2)}, D^{(3,3)}, D^{(4,4)}\right\}$, and

$$D^{(i,j)} = \begin{bmatrix} E\left(x_t x_t^T | q_t = \gamma_0\right) I(\alpha \le \varrho) & E(\left(x_t \eta_0^{(i)} | q_t = \gamma_0\right) I(\alpha = \varrho) \\ E\left(\eta_0^{(j)} x_t^T | q_t = \gamma_0\right) I(\alpha = \varrho) & E\left(\eta_0^{(i)} \eta_0^{(j)} | q_t = \gamma_0\right) I(\alpha \ge \varrho) \end{bmatrix}, \tag{D.23}$$

and $\eta_0^{(j)} = \eta_0^{(j)}(v_{q_1}, v_{q_2})$, $\eta_0^{(i)} = \eta_0^{(i)}(v_{q_1}, v_{q_2})$ for $j = 2,3,4$ and $i = 2,3,4$.

Denote $I_t^{(j)}(v,w) = I_t^{(j)}(\gamma_1^0 + \frac{v}{a_n}, \gamma_2^0 + \frac{w}{a_n})$ for $j = 1,2,3,4$. Following $\sum_{t=1}^n |I_t^{(j)}(v,w) - I_t^{(j)}(\gamma_0)| = O_p(n^{2min(\alpha,\varrho)})$, $\eta_0^{*(j)}(\widehat{v}_{q_1}, \widehat{v}_{q_2}) - \eta_0^{*(j)}(v_{q_1}, v_{q_2}) = O_p(\|\Phi_{L_n})\|_1 n^{-1/2})$, and $\eta_0^{*(j)}(v_{q_1}, v_{q_2})$ is bounded by a constant M for $t = 1, \ldots, n$ and $j = 1,2,3,4$ under Assumptions 1-2, we have

$$\begin{aligned}&\sum_{j=1}^{4} A_n^{(j)}(v,w) = O_p\left(n^{-1/2+2(min(\alpha,\varrho)-\varrho)} \|\Phi_{L_n}\|_1\right) + O_p\left(n^{-1+2(min(\alpha,\varrho)-\varrho)} \|\Phi_{L_n}\|_1^2\right) \\ &= o_p(1),\end{aligned} \tag{D.24}$$

where the last equality holds under Assumption 2 (f).

Closely following the interval split method used in the proof of Lemma 1 in Kourtellos et al. (2022), we can show that $\sum_{j=1}^4 A_n^{(j)}(v,w) = o_p(1)$ uniformly holds for all $v \in (0, \bar{v}]$ and $w \in (0, \bar{w}]$. Therefore, we have

$$S_{n,1}^*(v,w) = v\mu_1 + w\mu_2 + o_p(1).$$

uniformly over $v \in (0, \bar{v}]$ and $w \in (0, \bar{w}]$

Next, we consider $S^*_{n,2}(v,w)$. Given any $v \in (0,\bar{v}]$ and $w \in (0,\bar{w}]$, we have

$$
\begin{aligned}
S^*_{n,2}(v,w) &= \sum_{j=2}^{4}\Bigg(R_n^{(j)}(v,w) + \widehat{\kappa}_{n,j}^T \sum_{t=1}^{n}(\chi_t - \chi_t^*)\,\varepsilon_t\left(I_t^{(j)}(v,w) - I_t^{(j)}(\gamma_0)\right) \\
&+ (\widehat{\kappa}_{n,j} - \kappa_{n,j})^T \sum_{t=1}^{n}\chi_t^*\varepsilon_t\left(I_t^{(j)}(v,w) - I_t^{(j)}(\gamma_0)\right)\Bigg) \\
&= \sum_{j=2}^{4} R_n^{(j)}(v,w)\,[1+o_p(1)] + O_p\left(n^{1/2+min(\alpha,\rho)-\rho}\|\Phi_{L_n}\|_1\right),
\end{aligned}
$$

where

$$
R_n^{(j)} = \sum_{t=1}^{n}\varepsilon_t\kappa_{n,j}^T\chi_t^*\left[I_t^{(j)}(v,w) - I_t^{(j)}(\gamma_0)\right].
$$

Specifically, similar to the Lemma 5 in Kourtellos et al. (2022), for $j = 2,3,4$, under Assumptions 1 - 2, we obtain

$$
\begin{aligned}
&\sum_{t=1}^{n}\varepsilon_t\kappa_{n,2}^T\chi_t^*\left[I_t^{(2)}(v,w) - I_t^{(2)}(\gamma_0)\right] \\
&= \sum_{t=1}^{n}\varepsilon_t\kappa_{n,2}^T\chi_t^*\left[I_t^{(2)}(\gamma_1,\gamma_2) - I_t^{(2)}(\gamma_1^0,\gamma_2)\right] - \sum_{t=1}^{n}\varepsilon_t\kappa_{n,2}^T\chi_t^*\left[I_t^{(2)}(\gamma_1^0,\gamma_2^0) - I_t^{(2)}(\gamma_1^0,\gamma_2)\right] \\
&\Longrightarrow [B_{4,1}^{(2)}(v) - B_{2,2}^{(2)}(w)]d_2,
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{t=1}^{n}\varepsilon_t\kappa_{n,3}^T\chi_t^*\left[I_t^{(3)}(v,w) - I_t^{(3)}(\gamma_0)\right] \\
&= -\sum_{t=1}^{n}\varepsilon_t\kappa_{n,3}^T\chi_t^*\left[I_t^{(3)}(\gamma_1^0,\gamma_2) - I_t^{(3)}(\gamma_1,\gamma_2)\right] + \sum_{t=1}^{n}\varepsilon_t\kappa_{n,3}^T\chi_t^*\left[I_t^{(3)}(\gamma_1^0,\gamma_2) - I_t^{(3)}(\gamma_1^0,\gamma_2^0)\right] \\
&\Longrightarrow [-B_{3,1}^{(3)}(v) + B_{4,2}^{(3)}(w)]d_3,
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{t=1}^{n}\varepsilon_t\kappa_{n,4}^T\chi_t^*\left[I_t^{(4)}(v,w) - I_t^{(4)}(\gamma_0)\right] \\
&= -\sum_{t=1}^{n}\varepsilon_t\kappa_{n,4}^T\chi_t^*\left[I_t^{(4)}(\gamma_1^0,\gamma_2) - I_t^{(4)}(\gamma_1,\gamma_2)\right] - \sum_{t=1}^{n}\varepsilon_t\kappa_{n,4}^T\chi_t^*\left[I_t^{(4)}(\gamma_1^0,\gamma_2^0) - I_t^{(4)}(\gamma_1^0,\gamma_2)\right] \\
&= -\sum_{t=1}^{n}\varepsilon_t\kappa_{n,4}^T\chi_t^*\left[I_t^{(2)}(\gamma_1,\gamma_2) - I_t^{(2)}(\gamma_1^0,\gamma_2)\right] - \sum_{t=1}^{n}\varepsilon_t\kappa_{n,4}^T\chi_t^*[I_t^{(3)}(\gamma_1^0,\gamma_2) - I_t^{(3)}(\gamma_1^0,\gamma_2^0)] \\
&\Longrightarrow \left[-B_{4,1}^{(4)}(v) - B_{4,2}^{(4)}(w)\right]d_4
\end{aligned}
$$

where

$$B_{4,1}^{(4)}(v) = \left(V_4^{(4)}\right)^{1/2} \left(V_4^{(2)}\right)^{-1/2} B_{4,1}^{(2)}(v),$$
$$B_{4,2}^{(4)}(v) = \left(V_4^{(4)}\right)^{1/2} \left(V_4^{(3)}\right)^{-1/2} B_{4,2}^{(3)}(v),$$

$B_{4,1}^{(2)}(\cdot)$, $B_{2,2}^{(2)}(\cdot)$, $B_{3,1}^{(3)}(\cdot)$, and $B_{4,2}^{(3)}(\cdot)$ are four mutually independent Brownian motion vectors corresponding to the four disjointed regions with the covariance matrix

$$E\left[B_{i,1}^{(j)}(1) B_{i,1}^{(j)T}(1)\right] = V_i^{(j)} f_1^0,$$
$$E\left[B_{i,2}^{(j)}(1) B_{i,2}^{(j)T}(1)\right] = V_i^{(j)} f_2^0,$$

$$V_i^{(j)} = \begin{bmatrix} E\left[x_t x_t^T \varepsilon_{it}^2 | q_t = \gamma_0\right] I(\alpha \le \rho) & E\left[x_t \eta_0^{(j)} \varepsilon_{it}^2 | q_t = \gamma_0\right] I(\alpha = \rho) \\ E\left[\eta_0^{(j)} x_t^T \varepsilon_{it}^2 | q_t = \gamma_0\right] I(\alpha = \rho) & E\left[(\eta_0^{(j)})^2 \varepsilon_{it}^2 | q_t = \gamma_0\right] I(\alpha \ge \rho) \end{bmatrix}, \tag{D.25}$$

and i represents the regime heteroskedasticity.

Next, let $B_1^*(v) = \left[B_{4,1}^{(2)}(v), -B_{3,1}^{(3)}(v)\right]$ and $B_2^*(v) = \left[-B_{2,2}^{(2)}(v), B_{4,2}^{(3)}(v)\right]$. Thus, $B_1^*(v)$ and $B_2^*(v)$ are two independent Brownian motion vectors with covariance matrix

$$E\left[B_1^*(1) B_1^*(1)^T\right] = diag\left(V_4^{(2)}, V_3^{(3)}\right) f_1^0,$$
$$E\left[B_2^*(1) B_2^*(1)^T\right] = diag\left(V_2^{(2)}, V_4^{(3)}\right) f_2^0.$$

Hence, closely following Chen et al. (2012), we have

$$\begin{aligned} &\sum_{j=2}^{4} R_n^{(j)}(v, w) \\ &\Longrightarrow -2\left\{B_1^*(v) D_{11} + B_2^*(v) D_{21}\right\} \\ &= -2\left\{\sqrt{D_{11}^T V_{11}^* D_{11} f_1^0}\; W_1(v) + \sqrt{D_{21}^T V_{21}^* D_{21} f_2^0}\; W_2(w)\right\}, \end{aligned}$$

where $D_{11} = \left[d_2^T - \left((V_4^{(4)})^{1/2}(V_4^{(2)})^{-1/2} d_4\right)^T, d_3^T\right]^T$, $D_{21} = \left[d_2^T, d_3^T - \left((V_4^{(4)})^{1/2}(V_4^{(3)})^{-1/2} d_4\right)^T\right]^T$, $W_1(v)$ and $W_2(w)$ are independent standard Brownian motions, $V_{11}^* = diag\left\{V_4^{(2)}, V_3^{(3)}\right\}$, $V_{21}^* = diag\left\{V_2^{(2)}, V_4^{(3)}\right\}$.

Similarly, for all four cases ($v \in [\underline{v}, \overline{v}]$, $w \in [\underline{w}, \overline{w}]$), we have

$$\sum_{j=2}^{4} R_n^{(j)}(v,w) \Longrightarrow -2\Bigg\{ \left[\sqrt{D_{11}^T V_{11}^* D_{11} f_1^0}\ I(v \geq 0) + \sqrt{D_{12}^T V_{12}^* D_{12} f_1^0}\ I(v < 0)\right] W_1(v) \\ + \left[\sqrt{D_{21}^T V_{21}^* D_{21} f_2^0}\ I(w \geq 0) + \sqrt{D_{22}^T V_{22}^* D_{22} f_2^0}\ I(w < 0)\right] W_2(w)\Bigg\}, \tag{D.26}$$

where $W_i(x)$ is two-sided Brownian motion on the real line defined as

$$W_i(x) = \begin{cases} \Lambda_{i1}(-x), & \text{if } x < 0 \\ 0, & \text{if } x = 0 \\ \Lambda_{i2}(x), & \text{if } x > 0 \end{cases}$$

and $\Lambda_{ij}(x)$, for $i = 1, 2$, $j = 1, 2$ are four independent standard Brownian motions on $[0, \infty)$, and

$$V_{12}^* = diag\left\{V_2^{(2)}, V_1^{(3)}\right\},\ V_{22}^* = diag\left\{V_1^{(2)}, V_3^{(3)}\right\},$$
$$D_{12} = \left[d_2^T - \left((V_4^{(4)})^{1/2}(V_2^{(2)})^{-1/2} d_4\right)^T, d_3^T\right]^T,\ D_{22} = \left[d_2^T, d_3^T - \left((V_4^{(4)})^{1/2}(V_3^{(3)})^{-1/2} d_4\right)^T\right]^T.$$

Next, note that $S_{n,4}^*(v,w) = o_p\left(S_{n,1}^*(v,w)\right)$ and $S_{n,3}^*(v,w) = o_p\left(S_{n,1}^*(v,w)\right)$. Hence, to sum up, for all four cases, we have

$$S_n(\gamma_1^0 + \frac{v}{a_n}, \gamma_2^0 + \frac{w}{a_n}) - S_n(\gamma_0) = Q_n(v,w) \Longrightarrow Q(v,w) = |v|\mu_1 + |w|\mu_2 - 2\left[\lambda_1 W_1(v) + \lambda_2 W_2(w)\right], \tag{D.27}$$

where

$$\lambda_1 = \lambda_{11}\ I(v > 0) + \lambda_{12}\ I(v \leq 0), \tag{D.28}$$

$$\lambda_2 = \lambda_{21}\ I(w > 0) + \lambda_{22}\ I(w \leq 0), \tag{D.29}$$

$$\lambda_{11} = \sqrt{\sigma_{11}^2 f_1^0},\ \lambda_{12} = \sqrt{\sigma_{12}^2 f_1^0},\ \lambda_{21} = \sqrt{\sigma_{21}^2 f_2^0},\ \lambda_{22} = \sqrt{\sigma_{22}^2 f_2^0}, \tag{D.30}$$

$$\sigma_{11}^2 = D_{11}^T V_{1,1}^* D_{11},\ \sigma_{12}^2 = D_{12}^T V_{1,2}^* D_{12},\ \sigma_{21}^2 = D_{21}^T V_{2,1}^* D_{21},\ \sigma_{22}^2 = D_{22}^T V_{2,2}^* D_{22}, \tag{D.31}$$

and $\mu_1,\ \mu_2$ are defined in (D.22).

Let $v = \frac{\lambda_{11}^2}{\mu_1^2} r_1$ and $w = \frac{\lambda_{21}^2}{\mu_2^2} r_1$. Applying the change of variables, we complete the proof of this theorem.

Proof of Theorem 3: Under the null of $\gamma = \gamma_0$ and closely following the proof of Theorem 2 and the proof of Theorem 2 of Kourtellos et al. (2022), we have

$$n^{-1}S_n\left(\widehat{\theta}(\widehat{\gamma}),\widehat{\gamma}\right)LR_n(\gamma_0) - Q_n\left(\widehat{\gamma}\right) = o_p(1),$$

where $S_n(\theta,\gamma) = \sum_{t=1}^{n}\left(y_t - \chi_{t,\gamma}^T\widehat{\theta}(\gamma)\right)^2$, $n^{-1}S_n\left(\widehat{\theta}(\widehat{\gamma}),\widehat{\gamma}\right) \xrightarrow{p} \bar{\sigma}_\varepsilon^2$ following the proof of Theorem 1, and $Q_n\left(\widehat{\gamma}\right) = Q_n(\widehat{v},\widehat{w})$ is defined in (D.27).

Hence, by continuous mapping theorem, we have

$$LR_n(\gamma_0) = \frac{\underset{v,w}{sup}\ Q_n(v,w)}{\bar{\sigma}_\varepsilon^2} + o_p(1) \xrightarrow{d} \frac{\underset{v,w}{sup}\ Q(v,w)}{\bar{\sigma}_\varepsilon^2},$$

where $Q(v,w)$ is defined in (D.27).

Let $v = \frac{\lambda_{11}^2}{\mu_1^2}s_1$, $w = \frac{\lambda_{21}^2}{\mu_2^2}s_2$, and apply the change of variables, we have

$$\frac{\underset{v,w}{sup}\ Q(v,w)}{\bar{\sigma}_\varepsilon^2} = \frac{\lambda_{11}^2}{\bar{\sigma}_\varepsilon^2\mu_1}\xi,$$

where $\xi = 2max\left\{\xi_1,\xi_2,\xi_3,\xi_4\right\}$,

$$\begin{aligned}
\xi_1 &= \underset{s_1>0,\ s_2>0}{sup}\left[-\frac{|s_1|}{2} + W_1(s_1) + \psi\left(-\frac{|s_2|}{2} + W_2(s_2)\right)\right], \\
\xi_2 &= \underset{s_1>0,\ s_2\le 0}{sup}\left[-\frac{|s_1|}{2} + W_1(s_1) + \psi\left(-\frac{|s_2|}{2} + \sqrt{\varphi_2}W_2(s_2)\right)\right], \\
\xi_3 &= \underset{s_1\le 0,\ s_2>0}{sup}\left[-\frac{|s_1|}{2} + \sqrt{\varphi_1}W_1(s_1) + \psi\left(-\frac{|s_2|}{2} + W_2(s_2)\right)\right], \\
\xi_4 &= \underset{s_1\le 0,\ s_2\le 0}{sup}\left[-\frac{|s_1|}{2} + \sqrt{\varphi_1}W_1(s_1) + \psi\left(-\frac{|s_2|}{2} + \sqrt{\varphi_2}W_2(s_2)\right)\right].
\end{aligned}$$

Note that, for any two iid exponential random variables, X_1, X_2, with distribution function $P(X_1 \le x) = 1 - e^{-\frac{x}{a}}$ and $P(X_2 \le x) = 1 - e^{-\frac{x}{b}}$ respectively, where a and b are two positive constants, by convolution theorem, the distribution function of $Z = X_1 + cX_2$ is $P(Z \le x) = 1 - \frac{a}{a-bc}e^{-\frac{x}{a}} + \frac{bc}{a-bc}e^{-\frac{x}{bc}}$, where c is a positive constant. In particular, if $a = b = c = 1$, we have $P(Z \le x) = 1 - e^{-x} - xe^{-x}$.

Hence, closely following the proof of Theorem 2 of Hansen (2000) and the proof of Theorem 2 of Kourtellos et al. (2022) and applying the results above, we can show the distribution functions of ξ_1, ξ_2, ξ_3, and ξ_4 as

$$P(\xi_1 \le x) = \left(1 - \frac{1}{1-\psi}e^{-x} + \frac{\psi}{1-\psi}e^{-\frac{x}{\psi}}\right),$$
$$P(\xi_2 \le x) = \left(1 - \frac{1}{1-\varphi_2\psi}e^{-x} + \frac{\varphi_2\psi}{1-\varphi_2\psi}e^{-\frac{x}{\varphi_2\psi}}\right),$$
$$P(\xi_3 \le x) = \left(1 - \frac{\varphi_1}{\varphi_1-\psi}e^{-\frac{x}{\varphi_1}} + \frac{\psi}{\varphi_1-\psi}e^{-\frac{x}{\psi}}\right),$$
$$P(\xi_4 \le x) = \left(1 - \frac{\varphi_1}{\varphi_1-\varphi_2\psi}e^{-\frac{x}{\varphi_1}} + \frac{\varphi_2\psi}{\varphi_1-\varphi_2\psi}e^{-\frac{x}{\varphi_2\psi}}\right).$$

Therefore,

$$P(\xi \le x) = P(2max\{\xi_1, \xi_2, \xi_3, \xi_4\} \le x) = P(\xi_1 \le \frac{x}{2})P(\xi_2 \le \frac{x}{2})P(\xi_3 \le \frac{x}{2})P(\xi_4 \le \frac{x}{2}),$$

which completes our proof of this theorem.

Proof of Theorem 4: Let $\Delta_{v,t} = \Big[\Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})I_t^{(1)}(v,w), \Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})I_t^{(2)}(v,w),$ $\Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})I_t^{(3)}(v,w), \Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})I_t^{(4)}(v,w)\Big]^T$, and $x_{v,t} = \Big[x_t^T I_t^{(1)}(v,w), x_t^T I_t^{(2)}(v,w), x_t^T I_t^{(3)}(v,w),$ $x_t^T I_t^{(4)}(v,w)\Big]^T$. Then, we have

$$\widehat{\beta} - \beta = \left[X_v^T(I_n - P_v)X_v\right]^{-1} X_v^T\left(I_n - P_v\right)\left(y - X_v\beta\right) = A_n(v,w)^{-1}B_n(v,w)$$

where X_v is an $n \times 4k$ matrix and stacks up $x_{v,t}^T$, P_v is a projection matrix formed by Δ_v (an $n \times 4L_n$ matrix stacks up $\Delta_{v,t}^T$), and I_n is an $n \times n$ identity matrix.

Applying the first order Taylor expansion and closely following the proof of Lemma C.1 and the proof of Theorem 3 in Kourtellos et al. (2022), we can show that, for $j = 1, 2, 3, 4$, and $v \in [\underline{v}, \overline{v}]$, $w \in [\underline{w}, \overline{w}]$, under Assumption 2 (f),

$$\frac{1}{n}\sum_{t=1}^{n} \Phi_{L_n}(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})x_t^T I_t^{(j)}(v,w) = \frac{1}{n}\sum_{t=1}^{n} \Phi_{L_n}(v_{q_1t}, v_{q_2t})x_t^T I_t^{(j)}(v,w)) + O_p(\|\Phi_{L_n}\|_1/\sqrt{n}),$$

$$\frac{1}{n}\sum_{t=1}^{n} \Phi_{L_n}(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})\Phi_{L_n}^T(\widehat{v}_{q_1t}, \widehat{v}_{q_2t})I_t^{(j)}(v,w)$$
$$= \frac{1}{n}\sum_{t=1}^{n} \Phi_{L_n}(v_{q_1t}, v_{q_2t})\Phi_{L_n}^T(v_{q_1t}, v_{q_2t})I_t^{(j)}(v,w) + O_p(\|\Phi_{L_n}\|_1^2 L_n/n),$$

and

$$
\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n} x_t x_t^T I_t^{(j)}(v,w) = E\left[x_t x_t^T I_t^{(j)}(\gamma_0)\right][1+O(a_n^{-1})] + o_p(1),\\
&\frac{1}{n}\sum_{t=1}^{n} \Phi_{L_n}(v_{q_1t},v_{q_2t}) x_t^T I_t^{(j)}(v,w) = E\left[\Phi_{L_n}(v_{q_1t},v_{q_2t}) x_t^T I_t^{(j)}(\gamma_0)\right][1+O(a_n^{-1})] + o_p(1),\\
&\frac{1}{n}\sum_{t=1}^{n} \Phi_{L_n}(v_{q_1t},v_{q_2t}) \Phi_{L_n}^T(v_{q_1t},v_{q_2t}) I_t^{(j)}(v,w)\\
&= E\left[\Phi_{L_n}(v_{q_1t},v_{q_2t}) \Phi_{L_n}^T(v_{q_1t},v_{q_2t}) I_t^{(j)}(v,w)][1+O(a_n^{-1})\right] + o_p(1).
\end{aligned}
$$

Hence, we obtain $n^{-1}X_v^T X_v \xrightarrow{p} \sum_{xx^T,\gamma_0}$, $n^{-1}X_v^T \Delta_v \xrightarrow{p} \sum_{x\Phi_{L_n}^T,\gamma_0}$, and $n^{-1}\Delta_v^T \Delta_v \xrightarrow{p} \sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}$, uniformly for all $v \in [\underline{v},\overline{v}],\ w \in [\underline{w},\overline{w}]$, where

$$
\begin{aligned}
&\Sigma_{xx^T,\gamma_0} = diag\left\{E\left[x_t x_t^T I_t^{(1)}(\gamma_0)\right], E\left[x_t x_t^T I_t^{(2)}(\gamma_0)\right], E\left[x_t x_t^T I_t^{(3)}(\gamma_0)\right], E\left[x_t x_t^T I_t^{(4)}(\gamma_0)\right]\right\},\\
&\Sigma_{x\Phi_{L_n}^T,\gamma_0} = diag\left\{E\left[x_t \Phi_{L_n}^T I_t^{(1)}(\gamma_0)\right], ..., E\left[x_t \Phi_{L_n}^T I_t^{(4)}(\gamma_0)\right]\right\},\\
&\Sigma_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0} = diag\left\{E\left[\Phi_{L_n}\Phi_{L_n}^T I_t^{(1)}(\gamma_0)\right], ..., E\left[\Phi_{L_n}\Phi_{L_n}^T I_t^{(4)}(\gamma_0)\right]\right\},
\end{aligned}
$$

and $\Phi_{L_n} = \Phi_{L_n}(v_{q_1}, v_{q_2})$.

Therefore, we have

$$
A_n(v,w) = n^{-1}\left(X_v^T X_v - X_v^T P_v X_v\right) = J_n + o_p(1), \tag{D.32}
$$

where $J_n = \sum_{xx^T,\gamma_0} - \sum_{x\Phi_{L_n}^T,\gamma_0} \sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1} \sum_{\Phi_{L_n}x^T,\gamma_0}$.

Secondly, we consider $B_n(v,w)$. Note that

$$
B_n(v,w) = n^{-1}X_v^T(I_n - P_v)\left[\sum_{i=2}^{4}\left(\mathbf{I}^{(i)}(\gamma_0) - \mathbf{I}^{(i)}(v,w)\right) x\delta_n^{(j)} + \sum_{j=1}^{4}\mathbf{I}^{(j)}(\gamma_0) h_j(v_{q_1}, v_{q_2}) + \varepsilon\right],
$$

where $\mathbf{I}^{(i)}(v,w)$ is an $n \times n$ diagonal matrix with its t^{th} diagonal element being $I_t^{(i)}(v,w)$.

(i) Similar to the proof of Lemma C.1, applying the first order Taylor expansion, we have, for all $i = 2, 3, 4$,

$$
\begin{aligned}
&\left\| n^{-1}X_v^T\left(\mathbf{I}^{(i)}(\gamma_0) - \mathbf{I}^{(i)}(v,w)\right)\right\| = O_p(a_n^{-1}),\\
&\left\| n^{-1}\Delta_v^T\left(\mathbf{I}^{(i)}(\gamma_0) - \mathbf{I}^{(i)}(v,w)\right)\right\| = O_p(a_n^{-1}),
\end{aligned} \tag{D.33}
$$

uniformly for all $v \in [\underline{v},\overline{v}],\ w \in [\underline{w},\overline{w}]$.

Closely following the proof of Theorem 3 in Kourtellos et al. (2022), we can show that, for all $i = 2, 3, 4$,

$$\left\| n^{-1} X_v^T P_v \left[\mathbf{I}^{(i)}(\gamma_0) - \mathbf{I}^{(i)}(v, w) \right] \right\| = O_p(a_n^{-1/2}),$$
$$\left\| n^{-1} \Delta_v^T P_v \left[\mathbf{I}^{(i)}(\gamma_0) - \mathbf{I}^{(i)}(v, w) \right] \right\| = O_p(a_n^{-1/2}), \tag{D.34}$$

uniformly for all $v \in [\underline{v}, \overline{v}],\ w \in [\underline{w}, \overline{w}]$.

Hence, we have, for $i = 2, 3, 4$ and for all $v \in [\underline{v}, \overline{v}],\ w \in [\underline{w}, \overline{w}]$,

$$\left\| n^{-1} X_v^T (I_n - P_v) \left[\mathbf{I}^{(i)}(\gamma_0) - \mathbf{I}^{(i)}(v, w) \right] x \delta_n^{(i)} \right\| = O_p(a_n^{-1/2} n^{-\alpha}) = o_p(n^{-1/2}),$$

where the last equality holds under Assumption 2 (f).

(ii) Note that $I_n - P_v$ eliminates any linear combination of $\Phi_{L_n}^T (\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) I_t^{(j)}(v, w)$, for $j = 1, 2, 3, 4$. Hence, $(I_n - P_v) \left[\sum_{j=1}^4 \mathbf{I}^{(j)}(\gamma_0) h_j(v_{q_1}, v_{q_2}) \right] = (I_n - P_v) \left[\sum_{j=1}^4 \Big(\mathbf{I}^{(j)}(\gamma_0) h_j(v_{q_1}, v_{q_2}) - \mathbf{I}^{(j)}(v, w) h_j^*(\widehat{v}_{q_1}, \widehat{v}_{q_2}) \Big) \right]$, where, for $j = 1, 2, 3, 4$ and $t = 1, \ldots, n$, we can partition

$$\begin{aligned} & h_j(v_{q_1 t}, v_{q_2 t}) I_t^{(j)}(\gamma_0) - h_j^*(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) I_t^{(j)}(v, w) \\ & = \left[h_j(v_{q_1 t}, v_{q_2 t}) - h_j^*(v_{q_1 t}, v_{q_2 t}) \right] I_t^{(j)}(\gamma_0) + \left[h_j^*(v_{q_1 t}, v_{q_2 t}) - h_j^*(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) \right] I_t^{(j)}(\gamma_0) \\ & - h_j^*(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) \left[I_t^{(j)}(v, w) - I_t^{(j)}(\gamma_0) \right]. \end{aligned}$$

Following (D.3), uniformly for $v \in [\underline{v}, \overline{v}]$, $w \in [\underline{w}, \overline{w}]$, and $j = 1, 2, 3, 4$, we can show

$$\left\| n^{-1} X_v^T \mathbf{I}^{(j)}(\gamma_0) \left[h_j(v_{q_1}, v_{q_2}) - h_j^*(v_{q_1}, v_{q_2}) \right] \right\| = O(L_n^{-p}) \left\| \frac{1}{n} \sum_{t=1}^n x_{v,t}^{(j)}(\gamma_0) \right\| = O_p(L_n^{-p}),$$
$$\left\| n^{-1} \Delta_v^T \mathbf{I}^{(j)}(\gamma_0) \left[h_j(v_{q}1, v_{q}2) - h_j^*(v_{q_1}, v_{q_2}) \right] \right\| = O(L_n^{-p}) \left\| \frac{1}{n} \sum_{t=1}^n \Delta_{v,t} I^{(j)}(\gamma_0) \right\| = O_p(L_n^{-p}),$$

Hence, under Assumption 2 (f), we obtain, for $j = 1, 2, 3, 4$,

$$\left\| n^{-1} X_v^T \left(I_n - P_v \right) \mathbf{I}^{(j)}(\gamma_0) \left[h_j(z_1 \pi_{q_1}, z_2 \pi_{q_2}) - h_j^*(v_{q_1}, v_{q_2}) \right] \right\| = o_p(1).$$

Note that, for $t = 1, ..., n$, $j = 1, 2, 3, 4$ and for all $v \in [\underline{v}, \overline{v}],\ w \in [\underline{w}, \overline{w}]$, we have

$$\begin{aligned} & n^{-1} \sum_{t=1}^n x_t I_t^{(j)}(v, w) = E \left[x_t I_t^{(j)}(\gamma_0) \right] + O_p(a_n^{-1}), \\ & n^{-1} \sum_{t=1}^n \Phi_{L_n}(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) I_t^{(j)}(v, w) = E \left[\Phi_{L_n}(v_{q_1 t}, v_{q_2 t}) I_t^{(j)}(\gamma_0) \right] + O_p \left(a_n^{-1} \| \Phi_{L_n} \|_1 \sqrt{n} \right). \end{aligned}$$

Hence, we have, for $i, j = 1, 2, 3, 4$, and for all $v \in [\underline{v}, \overline{v}],\ w \in [\underline{w}, \overline{w}]$,

$$n^{-1} \sum_{t=1}^{n} \left\{ x_t I_t^{(i)}(v, w) \left[h_j^*(v_{q_1 t}, v_{q_2 t}) - h_j^*(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) \right] I_t^{(j)}(\gamma_0) \right\}$$
$$= \begin{cases} -\Gamma_{x,j} \begin{bmatrix} \widehat{\pi}_{q_1} - \pi_{q_1} \\ \widehat{\pi}_{q_2} - \pi_{q_2} \end{bmatrix} + O_p(L_n^{-p}), \text{ if } i = j, \\ O_p\left(a_n^{-1} \|\Phi_{L_n}\|_1 / \sqrt{n} + L_n^{-p} \right), \text{ if } i \neq j, \end{cases}$$

$$n^{-1} \sum_{t=1}^{n} \left\{ \Phi_{L_n}(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) I_t^{(i)}(v, w) \left[h_j^*(v_{q_1 t}, v_{q_2 t}) - h_j^*(\widehat{v}_{q_1 t}, \widehat{v}_{q_2 t}) \right] I_t^{(j)}(\gamma_0) \right\}$$
$$= \begin{cases} -\Gamma_{\Phi_{L_n},j} \begin{bmatrix} \widehat{\pi}_{q_1} - \pi_{q_1} \\ \widehat{\pi}_{q_2} - \pi_{q_2} \end{bmatrix} + O_p(L_n^{-p}), \text{ if } i = j, \\ O_p\left(a_n^{-1} \|\Phi_{L_n}\|_1 / \sqrt{n} + L_n^{-p} \right), \text{ if } i \neq j, \end{cases}$$

where

$$\Gamma_{x,j} = E\left[x_t I_t^{(j)}(\gamma_0) \mathbb{D} h_j(v_{q_1 t}, v_{q_2 t}) \begin{bmatrix} z_{1t}^T \\ z_{2t}^T \end{bmatrix} \right],$$
$$\Gamma_{\Phi_{L_n},j} = E\left[\Phi_{L_n}(v_{q_1 t}, v_{q_2 t}) I_t^{(j)}(\gamma_0) \mathbb{D} h_j(v_{q_1 t}, v_{q_2 t}) \begin{bmatrix} z_{1t}^T \\ z_{2t}^T \end{bmatrix} \right].$$

To sum up, under Assumption 2 (f), we have

$$\sum_{j=1}^{4} \left\{ n^{-1} X_v^T (I_n - P_v) \mathbf{I}^{(j)}(\gamma_0) \left[h_j^*(v_{q_1}, v_{q_2}) - h_j^*(\widehat{v}_{q_1}, \widehat{v}_{q_2}) \right] \right\}$$
$$= -B_n \begin{bmatrix} \widehat{\pi}_{q_1} - \pi_{q_1} \\ \widehat{\pi}_{q_2} - \pi_{q_2} \end{bmatrix} + o_p(n^{-1/2}), \tag{D.35}$$

where $B_n = \sum_{j=1}^{4} B_{n,j}$, and $B_{n,j} = \Gamma_{x,j} - \sum_{x\Phi_{L_n}^T,\gamma_0} \sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1} \Gamma_{\Phi_{L_n},j}$.

(iii) By (D.8) and (D.34), we have, for $j = 1, 2, 3, 4$,

$$\left\| n^{-1} X_v^T (I_n - P_v) \left[\mathbf{I}^{(j)}(v, w) - \mathbf{I}^{(j)}(\gamma_0) \right] h_j^*(\widehat{v}_{q_1}, \widehat{v}_{q_2}) \right\| = O_p\left(a_n^{-1/2} L_n^{-p} + a_n^{-1/2} \|\Phi_{L_n}\|_1 / \sqrt{n} \right).$$

Therefore, taking (i)-(iii) together, we have

$$B_n(v, w) = -B_n \left[\widehat{\boldsymbol{\pi}}_q - \boldsymbol{\pi}_q \right] + n^{-1} X_v^T \left(I_n - P_v \right) \varepsilon + o_p(n^{-1/2}) \tag{D.36}$$

$$= -B_n\left[\widehat{\boldsymbol{\pi}}_q - \boldsymbol{\pi}_q\right] + n^{-1}\left[X_0^T - \sum_{x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Delta_0^T\right]\varepsilon + o_p(n^{-1/2}),$$

where $\widehat{\boldsymbol{\pi}}_q - \boldsymbol{\pi}_q = \begin{bmatrix}\widehat{\pi}_{q_1} - \pi_{q_1}\\ \widehat{\pi}_{q_2} - \pi_{q_2}\end{bmatrix}$.

Hence, following the proof of Theorem 3 in Kourtellos et al. (2022) and applying Wooldridge and White's central limit theorem for strong mixing process (White (2001), Th.5.2, p.130), we obtain

$$\begin{bmatrix} n^{-1/2}Z_q^T V_q \\ n^{-1/2}(X_0^T - \sum_{x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Delta_0^T)\varepsilon\end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix}\Omega_{qq} & \Omega_{q\varepsilon}\\ \Omega_{q\varepsilon}^T & \Omega_{\varepsilon\varepsilon},\end{bmatrix}\right) \tag{D.37}$$

where Z_q stacks up $[q_{1t}, q_{2t}]$ and V_q stacks up $[v_{q_{1t}}, v_{q_{2t}}]$ for $t = 1, .., n$, $\Omega_{qq} = \lim_{n\to\infty} Var(n^{-1/2}Z_q^T V_q)$, $\Omega_{q\varepsilon} = \left[\lim_{n\to\infty}\sum_{t=2}^n\sum_{s=1}^{t-1}E\left(z_{1t}\zeta_s^T v_{q_{1t}}\varepsilon_t\right), \lim_{n\to\infty}\sum_{t=2}^n\sum_{s=1}^{t-1}E\left(z_{2t}\zeta_s^T v_{q_{2t}}\varepsilon_t\right)\right] = O(1)$ under Assumption 1, ζ_s is the s^{th} column elements of $\left(X_0^T - \sum_{x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Delta_0^T\right)$,
and $\Omega_{\varepsilon\varepsilon} = \sum_{\varepsilon x x^T,\gamma_0} - \sum_{x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\varepsilon\Phi_{L_n}x^T,\gamma_0} - \sum_{\varepsilon x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\Phi_{L_n}x^T,\gamma_0}$
$+ \sum_{x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\varepsilon\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\Phi_{L_n}x^T,\gamma_0}$.

Applying the continuous mapping theorem, this completes the proof of this Theorem.

Proof of Theorem 5: With the re-defined χ_t, we can rewrite the model as

$$y_t = \sum_{j=1}^4 \chi_t^T\theta_j I_t^{(j)}(\gamma_1^0, \gamma_2^0) + \tilde{\varepsilon}_t,$$

where

$$\tilde{\varepsilon}_t = \sum_{j=1}^4\left[z_{xt}^T\left(\pi_x - \widehat{\pi}_x\right)^T\beta_j + h_j(v_{q_{1t}}, v_{q_{2t}}) - h_j^*(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}})\right]I_t^{(j)}(\gamma_1^0, \gamma_2^0) + \varepsilon_t,$$

and ε_t is defined in (3.2). Therefore, the objective function is the same as (D.1) with newly defined χ and $\tilde{\varepsilon}$.

Note that $\pi_x - \widehat{\pi}_x = O_p(n^{-1/2})$. Hence, similar the proof of Theorem 1, under Assumption 3 (c) and 4 (c), we have

$$\underset{\gamma\in\theta_\gamma}{max} S_n(\gamma_1, \gamma_2) = \varepsilon^T\varepsilon + o_p(1).$$

Next, by the triangular inequality, we have

$$S_{n2}(\gamma_1, \gamma_2) = \sum_{j=1}^4 |n^{-1}\tilde{\varepsilon}^T\mathbf{I}^{(j)}(\gamma_1^0, \gamma_2^0)\chi\theta_j| \le \sum_{j=1}^4 |n^{-1}\sum_{t=1}^n \chi_{j,t,\gamma}^T\theta_j\left[h_j(v_{q_{1t}}, v_{q_{2t}}) - h_j^*(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}})\right]|$$

$$+ \sum_{j=1}^4 |n^{-1}\sum_{t=1}^n \chi_{j,t,\gamma}^T\theta_j z_{xt}^T\left(\pi_x - \widehat{\pi}_x\right)^T\beta_j| + \sum_{j=1}^4 |n^{-1}\sum_{t=1}^n \chi_{j,t,\gamma}^T\theta_j\varepsilon_t|, \tag{D.38}$$

where $\chi_{j,t,\gamma}$ defines in section 3.1.

Note that, $\chi_{j,t,\gamma}^T\theta_j = \beta_j^T\widehat{\pi}_x z_{xt} I_t^{(j)}(\gamma_1,\gamma_2) + h_j^*(\widehat{v}_{q_1t},\widehat{v}_{q_2t}) = \beta_j^T\pi_x z_{xt} I_t^{(j)}(\gamma_1,\gamma_2) + h_j(v_{q_1t},v_{q_2t}) + O_p(n^{-1/2}+L_n^{-p}+\|\Phi_{L_n}\|_1\sqrt{n})$. Hence, for all $\gamma\in\Theta_\gamma$, under Assumptions 3 and 4, following the proof of Theorem 1, we can show (i) $\sum_{j=1}^4 |n^{-1}\varepsilon^T\chi\theta_j| = O_p(n^{-1/2})$; (ii) $\sum_{j=1}^4 |n^{-1}\sum_{t=1}^n \chi_{j,t,\gamma}^T\theta_j z_x^T(\pi_x - \widehat{\pi}_x)^T\beta_j| = O_p(n^{-1/2}(1+L_n^{-p}+\|\Phi_{L_n}\|_1/\sqrt{n})) = O_p(n^{-1/2})$; (iii) $\sum_{j=1}^4 |n^{-1}\sum_{t=1}^n \chi_{j,t,\gamma}^T\theta_j\left[h_j(v_{q_1t},v_{q_2t}) - h_j^*(\widehat{v}_{q_1t},\widehat{v}_{q_2t})\right]| = O_p(L_n^{-p}+\|\Phi_{L_n}\|_1/\sqrt{n})$. As a result, we have $S_{n2}(\gamma_1,\gamma_2) = O_p(n^{-1/2}+L_n^{-p}+\|\Phi_{L_n}\|_1/\sqrt{n}) = o_p(1)$.

Next, for $j=1,2,3,4$ and all $\gamma_{\gamma_1}\in\Theta_1, \gamma_{\gamma_2}\in\Theta_2$, under Assumption 3 (c) and Assumption 4, we have

$$\left\|n^{-1}\sum_{t=1}^n \chi_{j,t,\gamma}\tilde{\varepsilon}_t\right\| = \left\|n^{-1}\sum_{t=1}^n \pi_x z_{xt}\tilde{\varepsilon}_t\right\| + \left\|n^{-1}\sum_{t=1}^n \Phi_{L_n}(\widehat{v}_{q_1t},\widehat{v}_{q_2t})\tilde{\varepsilon}\right\| + O_p\left(n^{-1/2}(L_n^{-p}+\|\Phi_{L_n}\|_1/\sqrt{n}+n^{-1/2})\right)$$
$$= O_p\left(L_n^{-p}+\|\Phi_{L_n}\|_1/\sqrt{n}+\sqrt{L_n/n}\right) = o_p(1).$$

Then, closely following the proof of Theorem 1, we can show $n^{-1}S_{n1}(\gamma_1,\gamma_2) = S_1(\gamma_1,\gamma_2)+o_p(1)$, where $S_1(\gamma_1,\gamma_2)$ is the same as the proof of Theorem 1 with newly defined $M_j(\gamma_1,\gamma_2)$ in Lemma C.3. Hence, following the proof of Theorem 1, we obtain $\widehat{\gamma}\xrightarrow{p}\gamma_0$.

Next, by the triangular inequality, for all $\gamma\in\Theta_\gamma$, we have

$$\begin{aligned}
\left\|n^{-1}\chi_\gamma(y-\chi_\gamma\theta)\right\| &\le \left\|\sum_{j=1}^4 n^{-1}\sum_{t=1}^n \chi_{t,\gamma}z_{xt}^T\left(\pi_x-\widehat{\pi}\right)^T\beta_j\right\| \\
&+ \left\|\sum_{j=1}^4 n^{-1}\sum_{t=1}^n \chi_{t,\gamma}z_{xt}^T\left(\pi_x-\widehat{\pi}\right)^T\delta_n^j I_t^{(j)}(\gamma_1,\gamma_2)\right\| \\
&+ \left\|\sum_{j=1}^4 n^{-1}\sum_{t=1}^n \chi_{t,\gamma}z_{xt}^T\pi_x^T\delta_n^{(j)}\left[I_t^{(j)}(\gamma_1^0,\gamma_2^0)-I_t^{(j)}(\gamma_1,\gamma_2)\right]\right\| \\
&+ \left\|\sum_{j=2}^4 n^{-1}\sum_{t=1}^n \chi_{t,\gamma}\left[\eta_n^{(j)}(v_{q_1t},v_{q_2t})I_t^{(j)}(\gamma_1^0,\gamma_2^0)-\eta_n^{*(j)}(\widehat{v}_{q_1t},\widehat{v}_{q_2t})I_t^{(j)}(\widehat{\gamma}_1,\widehat{\gamma}_2)\right]\right\| \\
&+ \left\|n^{-1}\sum_{t=1}^n \chi_{t,\gamma}\varepsilon_t\right\|,
\end{aligned}$$

where $\left\|\sum_{j=1}^4 n^{-1}\sum_{t=1}^n \chi_{t,\gamma}z_{xt}^T(\pi_x-\widehat{\pi})^T\beta_j\right\| = O_p(n^{-1/2})$ and $\left\|\sum_{j=1}^4 n^{-1}\sum_{t=1}^n \chi_{t,\gamma}z_{xt}^T(\pi_x-\widehat{\pi})^T\delta_n^j I_t^{(j)}(\gamma_1,\gamma_2)\right\|$ $= O_p(n^{-1/2-\alpha})$.

Then following the proof of Theorem 1, which completes the proof of this Theorem.

Proof of Theorem 6: In the matrix form, we have

$$\mathbf{y}-\chi_\gamma\widehat{\theta}(\gamma)=z_x\pi_x^T\beta_1+\sum_{j=2}^4 z_{x,\gamma_0}^{(j)}\pi_x^T\delta_n^{(j)}+h_1(v_{q_1},v_{q_2})+\sum_{j=2}^4\eta_{n,\gamma_0}^{(j)}(v_{q_1},v_{q_2})+\varepsilon$$
$$-z_x\widehat{\pi}_x^T\widehat{\beta}_1-\sum_{j=2}^4 z_{x,\gamma_0}^{(j)}\widehat{\pi}_x^T\widehat{\delta}_n^{(j)}-h_1^*(\widehat{v}_{q_1},\widehat{v}_{q_2})-\sum_{j=2}^4\eta_{n,\gamma_0}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})$$
$$=\varepsilon+\Delta_n-\sum_{j=2}^4\Delta z_{x,\gamma}^{(j)}\pi_x^T\widehat{\delta}_n-\sum_{j=2}^4\Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})+o_p(1),$$

where, for $j=2,3,4$, $\Delta z_{x,\gamma}^{(j)}=z_{x,\gamma}^{(j)}-z_{x,\gamma_0}^{(j)}$, $\Delta\eta_{n,\gamma_0}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})=\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})-\eta_{n,\gamma_0}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})$, $\Delta_n=z_x\pi_x^T(\beta_1-\widehat{\beta}_1)+z_x(\pi_x-\widehat{\pi}_x)^T\widehat{\beta}_1+\sum_{j=2}^4 z_{x,\gamma_0}^{(j)}\pi_x^T(\delta_n^{(j)}-\widehat{\delta}_n^{(j)})+\sum_{j=2}^4 z_{x,\gamma_0}^{(j)}(\pi_x-\widehat{\pi}_x)^T\widehat{\delta}_n^{(j)}+h_1(v_{q_1},v_{q_2})-h_1^*(\widehat{v}_{q_1},\widehat{v}_{q_2})+\sum_{j=2}^4\left[\eta_{n,\gamma_0}^{(j)}(v_{q_1},v_{q_2})-\eta_{n,\gamma_0}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right]$, and the typical element of $z_{x,\gamma}^{(j)}$ is $z_t^TI_t^{(j)}(\gamma_1,\gamma_2)$.

Therefore, we can show

$$S_n(\gamma)-S_n(\gamma_0)$$
$$=\left[\sum_{j=2}^4\left(\Delta z_{x,\gamma}^{(j)}\pi_x^T\widehat{\delta}_n+\Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right)\right]^T\left[\sum_{j=2}^4\left(\Delta z_{x,\gamma}^{(j)}\pi_x^T\widehat{\delta}_n+\Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right)\right]$$
$$-2\left(\varepsilon+\Delta_n\right)^T\left[\sum_{j=2}^4\Delta z_{x,\gamma}^{(j)}\pi_x^T\widehat{\delta}_n+\sum_{j=2}^4\Delta\eta_{n,\gamma}^{*(j)}(\widehat{v}_{q_1},\widehat{v}_{q_2})\right]+o_p(1)$$
$$=\sum_{t=1}^n\Big\{\left[\kappa_{n,2}-\kappa_{n,4}\right]^T\chi_t'\chi_t'^T\left[\kappa_{n,2}-\kappa_{n,4}\right]I(\gamma_1^0\le q_{1t}<\gamma_1,q_{2t}>\gamma_2)+\kappa_{n,2}^T\chi_t'\chi_t'^T\kappa_{n,2}I(q_{1t}\le\gamma_1^0,\gamma_2^0\le q_{2t}<\gamma_2)$$
$$+\kappa_{n,3}^T\chi_t'\chi_t'^T\kappa_{n,3}I(\gamma_1^0\le q_{1t}<\gamma_1,q_{2t}\le\gamma_2)+\left[\kappa_{n,3}-\kappa_{n,4}\right]^T\chi_t'\chi_t'^T\left[\kappa_{n,3}-\kappa_{n,4}\right]I(q_{1t}>\gamma_1^0,\gamma_2^0\le q_{2t}<\gamma_2)\Big\}$$
$$-2\sum_{j=2}^4\widehat{\kappa}_{n,j}\left[\sum_{t=1}^n\varepsilon_t\chi_t'\left(I_t^{(j)}(\gamma_1,\gamma_2)-I_t^{(j)}(\gamma_1^0,\gamma_2^0)\right)\right]$$
$$-2\sum_{j=2}^4\widehat{\kappa}_{n,j}^T\left[\sum_{t=1}^n\Delta_{n_t}'\chi_t'\left(I_t^{(j)}(\gamma_1,\gamma_2)-I_t^{(j)}(\gamma_1^0,\gamma_2^0)\right)\right]$$
$$+\sum_{t=1}^n\Big\{\left[(\widehat{\kappa}_{n,2}-\widehat{\kappa}_{n,4})-(\kappa_{n,2}-\kappa_{n,4})\right]^T\chi_t'\chi_t'^T\left[(\widehat{\kappa}_{n,2}-\widehat{\kappa}_{n,4})+(\kappa_{n,2}-\kappa_{n,4})\right]I(\gamma_1^0\le q_{1t}<\gamma_1,q_{2t}>\gamma_2)$$
$$+\left[\widehat{\kappa}_{n,2}-\kappa_{n,2}\right]^T\chi_t'\chi_t'^T\left[\widehat{\kappa}_{n,2}+\kappa_{n,2}\right]I(q_{1t}\le\gamma_1^0,\gamma_2^0\le q_{2t}<\gamma_2)$$
$$+\left[\widehat{\kappa}_{n,3}-\kappa_{n,3}\right]^T\chi_t'\chi_t'^T\left[\widehat{\kappa}_{n,3}+\kappa_{n,3}\right]I(\gamma_1^0\le q_{1t}<\gamma_{1,2t}\le\gamma_2)$$
$$+\left[(\widehat{\kappa}_{n,3}-\widehat{\kappa}_{n,4})-(\kappa_{n,3}-\kappa_{n,4})\right]^T\chi_t'\chi_t'^T\left[(\widehat{\kappa}_{n,3}-\widehat{\kappa}_{n,4})+(\kappa_{n,3}-\kappa_{n,4})\right]I(q_{1t}>\gamma_1^0,\gamma_2^0\le q_{2t}<\gamma_2)\Big\}+o_p(1)$$
$$=S_{n,1}^*(\gamma)-2S_{n,2}^*(\gamma)-2S_{n,3}^*(\gamma)+S_{n,4}^*(\gamma)+o_p(1),$$

where $\chi_t' = \left[z_t^T \pi_x^T, \Phi_{L_n}(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}})\right]^T$, and $\Delta_{n_t}' = (\beta_1 - \widehat{\beta}_1)^T \pi_x z_{xt} + \sum_{j=2}^4 (\delta_n^{(j)} - \widehat{\delta}_n^{(j)})^T \pi_x z_{xt} I_t^{(j)}(\gamma_1^0, \gamma_2^0) + h_1(v_{q_{1t}}, v_{q_{2t}}) - h_1^*(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}}) + n^{-\varrho} \sum_{j=2}^4 \left[\eta_0^{(j)}(v_{q_{1t}}, v_{q_{2t}}) - \eta_0^{*(j)}(\widehat{v}_{q_{1t}}, \widehat{v}_{q_{2t}})\right] I_t^{(j)}(\gamma_1^0, \gamma_2^0)$.

Note that, for $j = 1, 2, 3, 4$, $S_{n,j}^*(\gamma)$ are defined the same as (D.19) with newly defined χ_t and $\chi_t^{'}$.

Denote

$$D^{(i,j)} = \begin{bmatrix} E\left[\pi_x z_{xt} z_{xt}^T \pi_x^T | q = \gamma_0\right] I(\alpha \leq \varrho) & E\left[\pi_x z_{xt} \eta_0^{(i)} | q = \gamma_0\right] I(\alpha = \varrho) \\ E\left[\eta_0^{(j)} z_{xt}^T \pi_x^T | q = \gamma_0\right] I(\alpha = \varrho) & E\left[\eta_0^{(i)} \eta_0^{(j)} | q = \gamma_0\right] I(\alpha \geq \varrho) \end{bmatrix}, \quad \text{(D.39)}$$

and

$$V_i^{(j)} = \begin{bmatrix} E\left[\pi_x z_{xt} z_{xt}^T \pi_x^T \varepsilon_{it}^2 | q = \gamma_0\right] I(\alpha \leq \varrho) & E\left[\pi_x z_{xt} \eta_0^{(j)} \varepsilon_{it}^2 | q = \gamma_0\right] I(\alpha = \varrho) \\ E\left[\eta_0^{(j)} z_{xt}^T \pi_x^T \varepsilon_{it}^2 | q = \gamma_0\right] I(\alpha = \varrho) & E\left[(\eta_0^{(j)})^2 \varepsilon_{it}^2 | q = \gamma_0\right] I(\alpha \geq \varrho) \end{bmatrix}. \quad \text{(D.40)}$$

Then, closely following the proof of Theorem 2 with the newly defined $D^{i,j}$ and $V_i^{(j)}$ completes the proof of this Theorem.

Proof of Theorem 7: The notation is defined the same as in the proof of Theorem 4 unless defined differently. Denote $x_{v,t} = [z_{xt}^T \widehat{\pi}_x^T I_t^{(1)}(v, w), z_x^T \widehat{\pi}_x^T I_t^{(2)}(v, w), z_{xt}^T \widehat{\pi}_x^T I_t^{(3)}(v, w), z_{xt}^T \widehat{\pi}_x^T I_t^{(4)}(v, w)]^T$. Hence, applying Lemma C.3, Lemma C.4, and following the proof of Theorem 4, we have $n^{-1} X_v^T X_v \xrightarrow{p} \sum_{\pi_x z_x z_x^T \pi_x^T, \gamma_0}$, $n^{-1} X_v^T \Delta_v \xrightarrow{p} \sum_{\pi_x z_x \Phi_{L_n}^T, \gamma_0}$, and $n^{-1} \Delta_v^T \Delta_v \xrightarrow{p} \sum_{\Phi_{L_n} \Phi_{L_n}^T, \gamma_0}$ uniformly for all $v \in [\underline{v}, \overline{v}]$, $w \in [\underline{w}, \overline{w}]$, where

$$\begin{aligned} \textstyle\sum_{\pi_x z_x z_x^T \pi_x^T, \gamma_0} &= diag\left(E\left(\pi_x z_x z_x^T \pi_x^T I_t^{(1)}(\gamma_0)\right), ..., E\left(\pi_x z_x z_x^T \pi_x^T I_t^{(4)}(\gamma_0)\right)\right), \\ \textstyle\sum_{\pi_x z_x \Phi_{L_n}^T, \gamma_0} &= diag\left(E\left(\pi_x z_x \Phi_{L_n}^T I_t^{(1)}(\gamma_0)\right), ..., E\left(\pi_x z_x \Phi_{L_n}^T I_t^{(4)}(\gamma_0)\right)\right). \end{aligned}$$

Therefore, we have

$$A_n(v, w) = n^{-1}\left(X_v^T X_v - X_v^T P_v X_v\right) = J_n + o_p(1), \quad \text{(D.41)}$$

where $J_n = \sum_{\pi_x z_x z_x^T \pi_x^T, \gamma_0} - \sum_{\pi_x z_x \Phi_{L_n}^T, \gamma_0} \sum_{\Phi_{L_n} \Phi_{L_n}^T, \gamma_0}^{-1} \sum_{\Phi_{L_n} z_x^T \pi_x^T, \gamma_0}$.

Secondly, we consider $B_n(v, w) = X_v^T (I_n - P_v)(y - X_v \beta)$, where $y_t - x_{v,t}^T \beta = \beta_1^T (\pi_x - \widehat{\pi}_x) z_{xt} + \sum_{j=2}^4 \delta_n^{(j)T} (\pi_x - \widehat{\pi}) z_{xt} I_t^{(j)}(\gamma_0) + \sum_{j=2}^4 \delta_n^{(j)T} \pi_x z_{xt} \left[I_t^{(j)}(\gamma_0 - I_t^{(j)}(v, w))\right] + \sum_{j=1}^4 h_j(v_{q1t}, v_{q2t}) I_t^{(j)}(\gamma_0) + \varepsilon_t + o_p(1)$.

By $\pi_x - \widehat{\pi}_x = O_p(n^{-1/2})$, closely following the proof of Theorem 4, we have

$$n^{-1} X_v^T (I_n - P_v) \left\{ z_x \left(\pi_x - \widehat{\pi}_x\right)^T \beta_1 + \sum_{2}^{4} \mathbf{I}_{\gamma_0}^{(j)} z_x \left(\pi_x - \widehat{\pi}_x\right)^T \delta_n^{(j)} + \sum_{j=2}^{4} \left(\mathbf{I}^{(j)}(\gamma_0) - \mathbf{I}^{(j)}(v, w)\right) z_x \pi_x^T \delta_n^{(j)} \right\}$$
$$= O_p\left(a_n^{-1/2} n^{-\alpha}\right) = o_p(n^{-1/2}).$$

Similarly, we can show

$$\sum_{j=1}^{4}\left\{n^{-1}X_v^T(I_n-P_v)\mathbf{I}^{(j)}(\gamma_0)\left[h_j^*(v_{q_1},v_{q_2})-h_j^*(\widehat{v}_{q_1},\widehat{v}_{q_2})\right]\right\}$$
$$=-B_n\begin{bmatrix}\widehat{\pi}_{q_1}-\pi_{q_1}\\ \widehat{\pi}_{q_2}-\pi_{q_2}\end{bmatrix}+o_p(n^{-1/2}), \tag{D.42}$$

where $B_n=\sum_{j=1}^{4}B_{n,j}$, $B_{n,j}=\Gamma_{\pi_x z_x,j}-\sum_{\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Gamma_{\Phi_{L_n},j}$, and for $j=1,2,3,4$,

$$\Gamma_{\pi_x z_x,j}=E\left[\pi_x z_t I_t^{(j)}(\gamma_0)\mathbb{D}h_j(v_{q1t},v_{q2t})\begin{bmatrix}z_{1t}^T\\ z_{2t}^T\end{bmatrix}\right].$$

This follows

$$B_n(v,w)=-B_n\left[\widehat{\boldsymbol{\pi}}_q-\boldsymbol{\pi}_q\right]+n^{-1}X_v^T\left(I_n-P_v\right)\varepsilon+o_p(n^{-1/2}) \tag{D.43}$$

$$=-B_n\left[\widehat{\boldsymbol{\pi}}_q-\boldsymbol{\pi}_q\right]+n^{-1}\left[X_0^T-\sum\nolimits_{\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum\nolimits_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Delta_0^T\right]\varepsilon+o_p(n^{-1/2}),$$

where $\widehat{\boldsymbol{\pi}}_q-\boldsymbol{\pi}_q=\begin{bmatrix}\widehat{\pi}_{q_1}-\pi_{q_1}\\ \widehat{\pi}_{q_2}-\pi_{q_2}\end{bmatrix}$.

Hence, applying the central limit theorem, we have

$$\begin{bmatrix}n^{-1/2}Z_q^TV_q\\ n^{-1/2}(X_0^T-\sum_{\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Delta_0^T)\varepsilon\end{bmatrix}\xrightarrow{d}N\left(0,\begin{bmatrix}\Omega_{qq} & \Omega_{q\varepsilon}\\ \Omega_{q\varepsilon}^T & \Omega_{\varepsilon\varepsilon},\end{bmatrix}\right) \tag{D.44}$$

where Z_q stacks up $[q_{1t},q_{2t}]$ and V_q stacks up $[v_{q_{1t}},v_{q_{2t}}]$ for $t=1,..,n$, $\Omega_{qq}=\lim_{n\to\infty}Var(n^{-1/2}Z_q^TV_q)$, $\Omega_{q\varepsilon}=[\lim_{n\to\infty}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E(z_{1t}\zeta_s^Tv_{q_{1t}}\varepsilon_t),\lim_{n\to\infty}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E(z_{2t}\zeta_s^Tv_{q_{2t}}\varepsilon_t)]=O(1)$ under Assumption 1, ζ_s is the s^{th} column elements of $(X_0^T-\sum_{\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\Delta_0^T)$,
and $\Omega_{\varepsilon\varepsilon}=\sum_{\varepsilon\pi_x z_x z_x^T\pi_x^T,\gamma_0}-\sum_{\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\varepsilon\Phi_{L_n}z_x^T\pi_x^T,\gamma_0}-\sum_{\varepsilon\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\Phi_{L_n}z_x^T\pi_x^T,\gamma_0}$ $+\sum_{\pi_x z_x\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\varepsilon\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}\sum_{\Phi_{L_n}\Phi_{L_n}^T,\gamma_0}^{-1}\sum_{\Phi_{L_n}z_x^T\pi_x^T,\gamma_0}$.

This completes the proof of this Theorem.

References

Chen, H., Chong, T. T.-L. & Bai, J. (2012), 'Theory and applications of tar model with two threshold variables', *Econometric Reviews* **31**(2), 142–170.

Chen, X. (2007), 'Chapter 76 large sample sieve estimation of semi-nonparametric models', *Handbook of Econometrics* pp. 5549–5632.

Hansen, B. E. (2000), 'Sample splitting and threshold estimation', *Econometrica* **68**, 575–603.

Kourtellos, A., Stengos, T. & Sun, Y. (2022), 'Endogeneity in semiparametric threshold regression', *Econometric Theory* **38**(3), 562–595.

Newey, W. K. (1997), 'Convergence rates and asymptotic normality for series estimators', *Journal of Econometrics* **79**(1), 147–168.

Newey, W. K. & Mcfadden, D. (1994), 'Large sample estimation and hypothesis testing', *Chapter 36 in Handbook of Econometrics* pp. 2111–2245.

Seo, M. H. & Linton, O. (2007), 'A smoothed least squares estimator for threshold regression models', *Journal of Econometrics* **141**(2), 704–735.

White, H. (2001), *Asymptotic Theory for Econometricians*, Academic Press.

Yu, P. (2013), 'Inconsistency of 2sls estimators in threshold regression with endogeneity', *Economics Letters* **120**(3), 532–536.