Nonlinear Error Correction Model and Multiple-threshold Cointegration

Man Wang[1], Ngai Hang Chan[2,3] and Chun Yip Yau[3]

[1] *Dong Hua University*, [2] *Southwestern University of Finance and Economics and* [3] *The Chinese University of Hong Kong*

Supplementary Material

S1 Proof of Theorems

Before we prove Theorem 1, we first establish the consistence of $\hat{\theta}$. For simplicity, we consider a two-threshold cointegration in a three-regime TVECM and the proof can be easily extended to a multiple-regime case. Let

$$\Delta x_t = A^{'} X_{t-1}(\beta) + D^{'} X_{t-1}(\beta) I(\beta, \gamma_1) + E^{'} X_{t-1}(\beta) I(\beta, \gamma_2) + u_t, \quad t = l+1, \ldots, n, \quad \text{(S1.1)}$$

where $\gamma_1 < \gamma_2$, $I(\beta, \gamma_1) = 1\{z_{t-1}(\beta) < \gamma_1\}$ and $I(\beta, \gamma_2) = 1\{z_{t-1}(\beta) \geq \gamma_2\}$. The coefficients in the three regimes are respectively $A + D$, A and $A + E$.

Lemma 1. *Under Assumption 1, $\hat{\theta} - \theta^0$ is $o_p(1)$, and furthermore, $\sqrt{n}(\hat{\beta} - \beta^0)$ is $o_p(1)$.*

Proof of Lemma 1 consists of two steps in the framework of the proof of Theorem 1 in Seo (2011). In the first step, we define a sequence of r_n and prove that

$$\hat{\beta} - \beta^0 = o_p(r_n^{-1}). \quad \text{(S1.2)}$$

In the second step, we show $\sqrt{n}(\hat{\beta} - \beta^0) = O_p(1)$ by Lemma 1 of Seo (2011) and further obtain the consistency of $\hat{\theta}$ and $\sqrt{n}(\hat{\beta} - \beta^0) = o_p(1)$ by Theorem 1 of de Jong (2002). It is obvious that the indicator function in model (S1.1) satisfies the assumption of Theorem 1 in Seo (2011), so the second step can be proved in the same way as Seo (2011) and is omitted. The proof of the first step can be demonstrated in similar way to Seo (2011), except that some additional terms need to be taken care of.

Let r_n be a sequence of real numbers such that $\sqrt{n} \geq r_n \to \infty$ and $r_n/\sqrt{n} \to 0$ as $n \to \infty$. Let $\Theta_{r_n,\delta} = \{\theta \in \Theta : r_n|\beta - \beta^0| > \delta\}$ be a subspace of Θ, we take all the following supremums and infimums in this proof over $\Theta_{r_n,\delta}$ unless stated otherwise.

To prove (S1.2), it is sufficient to show that

$$\Pr\{\inf_{\theta} \left(S_n(\theta) - S_n(\theta^0)\right)/n > 0\} \to 1, \forall \delta > 0. \quad \text{(S1.3)}$$

Let $\eta = \sqrt{n}(\beta - \beta^0)$, then $|\eta| \geq \delta\sqrt{n}/r_n \to \infty$, and $z_t(\beta) = z_t + \frac{\eta^{'} x_{2,t}}{\sqrt{n}}$. For simplicity, let $X_{\beta,\gamma}$=$(X(\beta), X_{\gamma_1}(\beta), X_{\gamma_2}(\beta)) \otimes I_p$, where $X_{\gamma_j}(\beta)$ stacks $X_{t-1}(\beta)^{'}\ I(z_{t-1}(\beta), \gamma_j)$, $j = 1, 2$,

$X(\beta)$ stacks $X_{t-1}(\beta)^{'}$ and $\lambda = \mathrm{vec}((A^{'}, D^{'}, E^{'})^{'})$ and $\lambda_z = \mathrm{vec}((A_z^{'}, D_z^{'}, E_z^{'})^{'})$. Then

$$S_n(\theta)/n - S_n(\theta^0)/n = \left[y^{'}y - S_n(\theta^0) - 2y^{'}X_{\beta,\gamma}\lambda + \lambda^{'}X_{\beta,\gamma}^{'}X_{\beta,\gamma}\lambda\right]/n$$
$$= \frac{y^{'}y - u^{'}u}{n} - 2\frac{1}{n}\frac{y^{'}X_{\beta,\gamma}}{|\eta|}\lambda|\eta| + |\eta|^2\frac{1}{n}\lambda^{'}\frac{X_{\beta,\gamma}^{'}X_{\beta,\gamma}}{|\eta|^2}\lambda. \tag{S1.4}$$

We analyze the three terms on the r.h.s of (S1.4) separately. Since u_t is independent of $X_{t-1}(\beta)$, based on (2) and standard arguments in linear regression analysis, we can show that $\frac{y^{'}y-u^{'}u}{n}$ converges in probability to the limit of $x'x/n$, where $x = \left[\left(\tilde{X}_1(\beta,\gamma), \ldots, \tilde{X}_m(\beta,\gamma)\right)\otimes I_p\right]\lambda$. This limit is a positive constant in view of Assumptions 1.2 and 1.3. By similar argument of Seo (2011), it can be proved that $\sup_{\theta}\left|\frac{1}{n}\frac{y^{'}X_{\beta,\gamma}}{|\eta|}\lambda\right| = O_p(1)$. Therefore, with $|\eta| \to \infty$, (S1.3) will hold if we can prove

$$\inf_{\theta}\frac{1}{n}\lambda^{'}\frac{X_{\beta,\gamma}^{'}X_{\beta,\gamma}}{|\eta|^2}\lambda \overset{d}{\Longrightarrow} G, \tag{S1.5}$$

where G is a random variable that is positive with probability 1.

Let $\dot{\eta} = \eta/|\eta|$. We divide $\frac{1}{n}\lambda^{'}\frac{X_{\beta,\gamma}^{'}X_{\beta,\gamma}}{|\eta|^2}\lambda$ into two parts as:

$$\frac{1}{n}\lambda^{'}\frac{X_{\beta,\gamma}^{'}X_{\beta,\gamma}}{|\eta|^2}\lambda = \lambda_z^{'}(\Xi_n(\beta,\gamma)\otimes I_p)\lambda_z + R_n(\theta), \tag{S1.6}$$

where

$$\Xi_n(\beta,\gamma) = \begin{pmatrix} \frac{1}{n^2}\sum_t(x_{2,t-1}^{'}\dot{\eta})^2 & \mathrm{P}_1 & \mathrm{P}_2 \\ \mathrm{P}_1^{'} & \mathrm{P}_1 & 0 \\ \mathrm{P}_2{'} & 0 & \mathrm{P}_2 \end{pmatrix},$$

with $\mathrm{P}_1 = \frac{1}{n^2}\sum(x_{2,t-1}^{'}\dot{\eta})^2 I_{t-1}(\beta,\gamma_1)$ and $\mathrm{P}_2 = \frac{1}{n^2}\sum(x_{2,t-1}^{'}\dot{\eta})^2 I_{t-1}(\beta,\gamma_2)$. It can be proved that $\sup_{\theta}|R_n(\theta)| = O_p(|\frac{r_n}{\sqrt{n}}|)$, because all the elements in $X_{\beta,\gamma}$ are stationary except for $z_{t-1}(\beta)$ and $\frac{1}{n|\eta|^2}\sum_t|z_{t-1}(\beta)\Delta x_t^{'}|$ is $O_p(\frac{1}{|\eta|})$. Since $\frac{r_n}{\sqrt{n}} \to 0$, the first part of r.h.s of (S1.6) is the dominating one. Consequently, (S1.5) will hold if we can prove

$$\Xi_n(\beta,\gamma) \overset{d}{\Rightarrow} (\dot{\eta}^{'}\Omega\dot{\eta})M. \tag{S1.7}$$

Herein

$$M = \begin{pmatrix} \int_0^1 W^2 & \int_0^1 W^2 1\{W<0\} & \int_0^1 W^2 1\{W>0\} \\ \int_0^1 W^2 1\{W<0\} & \int_0^1 W^2 1\{W<0\} & 0 \\ \int_0^1 W^2 1\{W>0\} & 0 & \int_0^1 W^2 1\{W>0\} \end{pmatrix}, \tag{S1.8}$$

and $\int_0^1 W^2$ is short for $\int_0^1 W(s)^2 ds$, where W is a standard Brownian motion. Notice that $\dot{\eta}^{'}\Omega\dot{\eta}$ is a positive constant bounded away from zero by Assumption 1.2. Further, M is nonnegative definite with a single zero eigenvalue, whose corresponding eigenvector is $(c, -c, -c)$ for any non-zero constant c. However, $\lambda_z = (c, -c, -c)$ means coefficients in front of $z_{t-1}(\beta)$ are (0,c,0) in the three regimes, which is excluded by Assumption 1.3. Therefore, $\lambda_z^{'}(\dot{\eta}^{'}\Omega\dot{\eta})M \otimes I_p\lambda_z$ is positive with probability 1.

(S1.7) can be proved similar to Theorem 1 in Seo (2011), consequently, the proof of Lemma 1 is completed.

Proof of Theorem 1.

Having proved Lemma 5, we only need to show the proof of this theorem on a subspace $\Theta_c = \{\theta : |\theta - \theta^0| < c\}$ for some $c > 0$. Notice that if the extremely fast convergence rates of $\hat{\beta}$ and $\hat{\gamma}$ are proved, then the limiting distribution of $\hat{\lambda}$ is easy to get: by following similar analysis as in the ordinary LSE case, as argued in Seo (2011). Therefore, the proof of this part is omitted and we focus on the convergence rate of $\hat{\beta}$ and $\hat{\gamma}$.

We consider η instead of β and separate the parameters into two groups as $\theta_1 = (\eta, \gamma_1 - \gamma_1^0, \gamma_2 - \gamma_2^0)$ and λ, then the true value of θ_1 is 0. Taking the same strategy used by Chan (1993) and Seo (2011), it is sufficient to show that: for any $\epsilon > 0$, there exist $c > 0$ and $K > 0$ such that

$$\liminf_{n\to\infty} \Pr\left\{\inf_{\theta\in\Theta_{c,K}} [S_n(\theta) - S_n(0,\lambda)] > 0\right\} > 1 - \epsilon, \tag{S1.9}$$

where $\Theta_{c,K} = \Theta_c \cap \{\theta : |\theta_1| > K/n\}$.

Let $z_{t-1}(\beta) = z_{t-1} + \eta' \frac{x_{2,t-1}}{\sqrt{n}}$, $\gamma_{1,t} = \gamma_1 - \eta' \frac{x_{2,t-1}}{\sqrt{n}}$ and $\gamma_{2,t} = \gamma_2 - \eta' \frac{x_{2,t-1}}{\sqrt{n}}$ where $\gamma_1 < \gamma_2$, and write the residual of model (S1.1) when θ is plugged in as:

$$\begin{aligned}
&u_t(\theta)\\
=&u_t - (A - A_0)' X_{t-1}\\
&- (D - D^0)' X_{t-1} 1\{z_{t-1} < \gamma_{1,t}\} - D^{0'} X_{t-1}(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\})\\
&- (E - E^0)' X_{t-1} 1\{z_{t-1} > \gamma_{2,t}\} - E^{0'} X_{t-1}(1\{z_{t-1} > \gamma_{2,t}\} - 1\{z_{t-1} > \gamma_2^0\})\\
&- (A_z + D_z 1\{z_{t-1} < \gamma_{1,t}\} + E_z 1\{z_{t-1} > \gamma_{2,t}\})\eta' \frac{x_{2,t-1}}{\sqrt{n}}\\
=&u_{1t}(\theta) + u_{2t}(\theta).
\end{aligned}$$

Where

$$\begin{aligned}
&u_{1t}(\theta)\\
=&u_t - (A - A_0)' X_{t-1}\\
&- (D - D^0)' X_{t-1} 1\{z_{t-1} < \gamma_{1,t}\} - D^{0'} X_{t-1}(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\})\\
&- (E - E^0)' X_{t-1} 1\{z_{t-1} > \gamma_{2,t}\} - E^{0'} X_{t-1}(1\{z_{t-1} > \gamma_{2,t}\} - 1\{z_{t-1} > \gamma_2^0\}),
\end{aligned}$$

and $u_{2t}(\theta) = -(A_z + D_z 1\{z_{t-1} < \gamma_{1,t}\} + E_z 1\{z_{t-1} > \gamma_{2,t}\})\eta' \frac{x_{2,t-1}}{\sqrt{n}}$.

In this way, $\frac{1}{n}(S_n(\theta) - S_n(0,\lambda)) = D_{1n} + D_{2n}$ with

$$\begin{aligned}
D_{1n} =& \frac{1}{n}\sum_t [u_{1t}(\theta)' u_{1t}(\theta) - u_{1t}(0,\lambda)' u_{1t}(0,\lambda)],\\
D_{2n} =& \frac{1}{n}\sum_t [u_{2t}(\theta)' u_{2t}(\theta) + 2u_{1t}(\theta)' u_{2t}(\theta)].
\end{aligned}$$

First analyze D_{2n}. We compare D_{2n} with the corresponding quantity appearing in a one threshold cointegration and consider the additional terms arising in the two thresholds model as follows:

$$
\begin{aligned}
&\frac{1}{n}\sum_t |X_{t-1}^{'}(1\{z_{t-1}<\gamma_{1,t}\}-1\{z_{t-1}<\gamma_1^0\})1\{z_{t-1}>\gamma_{2,t}\}D_0\eta^{'}\frac{x_{2,t-1}}{\sqrt{n}}| \\
\leq &\sup_{1\leq t\leq n}|\frac{x_{2,t-1}}{\sqrt{n}}|(\frac{1}{n}\sum_t |X_{t-1}|^2)^{1/2} \\
&\left(\frac{1}{n}\sum_t 1\{\min\{|\gamma_1^0|,|\gamma_{1,t}|\}\leq |z_{t-1}|\leq \max\{|\gamma_1^0|,|\gamma_{1,t}|\}\}\right)^{1/2} O(|\eta|) \\
=&O_p(1)O_p(c)O(|\eta|)=O_p(c|\eta|),
\end{aligned}
$$

by the uniform law of large numbers and Assumption 2.1 that the density function of $f_Z(z)$ is bounded.

The remaining terms can be analyzed by the same reasoning as in the proof of Theorem 2 of Seo (2011). As a result,

$$
\begin{aligned}
|\frac{1}{n}\sum_t u_{2t}(\theta)^{'}u_{2t}(\theta)| &= O_p(c|\eta|), \\
|\frac{1}{n}\sum_t u_{1t}(\theta)^{'}u_{2t}(\theta)]| &\leq o_p(|\eta|)+O_p(c|\eta|).
\end{aligned}
$$

and we conclude that: for any $m_1, \epsilon > 0$, there is $c > 0$ such that:

$$
\liminf_{n\to\infty} \Pr\left\{\sup_{\theta\in\Theta_c}[|D_{2n}(\theta)|-m_1|\theta_1|]\leq 0\right\} > 1-\epsilon. \tag{S1.10}
$$

Next, we prove: $\forall\ \epsilon > 0$ and $c > 0$, there exists $K > 0$, if n is large enough, there exists some constant m_2 such that,

$$
\Pr\{D_{1n}(\theta) > m_2|\theta_1|\} > 1-\epsilon, \forall\theta\in\Theta_{c,K}. \tag{S1.11}
$$

Since m_1 in (S1.10) is arbitrary, we can decrease m_1 such that $m_1 < m_2$, in which case, c will be smaller and (S1.11) still hold. Therefore, the combination of (S1.10) and (S1.11) completes the proof of Theorem 1.

We now prove (S1.10). We compare D_{1n} with the corresponding quantity appearing in a one threshold cointegration model and denote the difference as $\mathrm{tr}(Add)$, then,

$$
\begin{aligned}
Add =&\frac{1}{n}\sum_t[(D-D^0)^{'}X_{t-1}1\{z_{t-1}<\gamma_{1,t}\}-D^{0^{'}}X_{t-1}(1\{z_{t-1}<\gamma_{1,t}\}-1\{z_{t-1}<\gamma_1^0\})] \\
&[(E-E^0)^{'}X_{t-1}1\{z_{t-1}>\gamma_{2,t}\}-E^{0^{'}}X_{t-1}(1\{z_{t-1}>\gamma_{2,t}\}-1\{z_{t-1}>\gamma_2^0\})]^{'} \\
=&-\frac{1}{n}\sum_t[(D-D^0)^{'}X_{t-1}]1\{z_{t-1}<\gamma_{1,t}\}X_{t-1}^{'}E^0(1\{z_{t-1}>\gamma_{2,t}\}-1\{z_{t-1}>\gamma_2^0\}) \\
&-\frac{1}{n}\sum_t[(E-E^0)^{'}X_{t-1}]1\{z_{t-1}>\gamma_{2,t}\}X_{t-1}^{'}D^0(1\{z_{t-1}<\gamma_{1,t}\}-1\{z_{t-1}<\gamma_1^0\}) \\
&+\frac{1}{n}\sum_t[D^{0^{'}}X_{t-1}]^{'}E^{0^{'}}X_{t-1}(1\{z_{t-1}>\gamma_{2,t}\}-1\{z_{t-1}>\gamma_2^0\})
\end{aligned}
$$

$$(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\}).$$

The first and second terms in *Add* are

$$O_p(c)\frac{1}{n}\sum_t X_{t-1}X_{t-1}^{'}1\{\gamma_2^0 < z_{t-1} < \gamma_{2,t} \text{ or } \gamma_{2,t} < z_{t-1} < \gamma_2^0\}E^0.$$

and

$$O_p(c)\frac{1}{n}\sum_t X_{t-1}X_{t-1}^{'}1\{\gamma_1^0 < z_{t-1} < \gamma_{1,t} \text{ or } \gamma_{1,t} < z_{t-1} < \gamma_1^0\}D^0.$$

Since $(1\{z_{t-1} > \gamma_{2,t}\} - 1\{z_{t-1} > \gamma_2^0\})(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\})$ is always negative or zero, the third term satisfies

$$\begin{aligned}
&[D^{0'}X_{t-1}]^{'}[E^{0'}X_{t-1}](1\{z_{t-1} > \gamma_{2,t}\} - 1\{z_{t-1} > \gamma_2^0\})\\
&(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\})\\
\leq& 1/2\left\{X_{t-1}^{'}D^0D^{0'}X_{t-1}(1\{\gamma_1^0 < z_{t-1} < \gamma_{1,t} \text{ or } \gamma_{1,t} < z_{t-1} < \gamma_1^0\})\right.\\
&\left.+X_{t-1}^{'}E^0E^{0'}X_{t-1}(1\{\gamma_2^0 < z_{t-1} < \gamma_{2,t} \text{ or } \gamma_{2,t} < z_{t-1} < \gamma_2^0\})\right\},
\end{aligned}$$

because of Assumption 1.3. Therefore, D_{1n} becomes:

$$\begin{aligned}
&D_{1n}\\
=&tr\left((D^{0'} + O_p(c))\frac{1}{n}\sum_t X_{t-1}X_{t-1}^{'}(1\{\gamma_1^0 < z_{t-1} < \gamma_{1,t} \text{ or } \gamma_{1,t} < z_{t-1} < \gamma_1^0\})D^0\right)\\
&+tr\left((E^{0'} + O_p(c))\frac{1}{n}\sum_t X_{t-1}X_{t-1}^{'}(1\{\gamma_2^0 < z_{t-1} < \gamma_{2,t} \text{ or } \gamma_{2,t} < z_{t-1} < \gamma_2^0\})E^0\right)\\
&+tr\left(\frac{1}{n}\sum_t[D^{0'}X_{t-1}]^{'}E^{0'}X_{t-1}(1\{z_{t-1} > \gamma_{2,t}\} - 1\{z_{t-1} > \gamma_2^0\})\right.\\
&\left.(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\})\right)\\
&+tr\left((D^{0'} + O(c))\frac{1}{n}\sum_t X_{t-1}u_t^{'}(1\{\gamma_1^0 < z_{t-1} < \gamma_{1,t} \text{ or } \gamma_{1,t} < z_{t-1} < \gamma_1^0\})\right)\\
&+tr\left((E^{0'} + O(c))\frac{1}{n}\sum_t X_{t-1}u_t^{'}(1\{\gamma_2^0 < z_{t-1} < \gamma_{2,t} \text{ or } \gamma_{2,t} < z_{t-1} < \gamma_2^0\})\right)\\
&+tr\left(O_p(c^2)\right). \qquad \text{(S1.12)}
\end{aligned}$$

By the same reasoning as in Seo (2011), the last three quantities on the r.h.s of (S1.12) have the same conclusion of (S1.10) as D_{2n}. Further,

$$\begin{aligned}
&tr[(D^{0'} + O_p(c))\frac{1}{n}\sum_t X_{t-1}X_{t-1}^{'}(1\{\gamma_1^0 < z_{t-1} < \gamma_{1,t} \text{ or } \gamma_{1,t} < z_{t-1} < \gamma_1^0\})D^0\\
&+(E^{0'} + O_p(c))\frac{1}{n}\sum_t X_{t-1}X_{t-1}^{'}(1\{\gamma_2^0 < z_{t-1} < \gamma_{2,t} \text{ or } \gamma_{2,t} < z_{t-1} < \gamma_2^0\})E^0
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{n}\sum_t [D^{0'}X_{t-1}]' E^{0'} X_{t-1}(1\{z_{t-1} > \gamma_{2,t}\} - 1\{z_{t-1} > \gamma_2^0\}) \\
&(1\{z_{t-1} < \gamma_{1,t}\} - 1\{z_{t-1} < \gamma_1^0\})] \\
\geq & tr[(1/2D^{0'} + O_p(c))\frac{1}{n}\sum_t X_{t-1}X_{t-1}'(1\{\gamma_1^0 < z_{t-1} < \gamma_{1,t} \text{ or } \gamma_{1,t} < z_{t-1} < \gamma_1^0\})D^0 \\
&+ (1/2E^{0'} + O_p(c))\frac{1}{n}\sum_t X_{t-1}X_{t-1}'(1\{\gamma_2^0 < z_{t-1} < \gamma_{2,t} \text{ or } \gamma_{2,t} < z_{t-1} < \gamma_2^0\})E^0], \quad \text{(S1.13)}
\end{aligned}
$$

and we can analyze the r.h.s of (S1.13) with the same reasoning as in Seo (2011) and conclude that (S1.11) hold. Therefore, the proof is completed.

Proof of Theorem 3.

We demonstrate the proof based on a general m-regime TVECM:

$$
\Delta x_t = \sum_{j=1}^{m} A_j' X_{t-1}(\beta) 1\{\gamma_{j-1} \leq z_{t-1}(\beta) < \gamma_j\} + u_t, \qquad t = l+1, \ldots, n. \qquad \text{(S1.14)}
$$

$\hat{\theta}^* = \underset{\theta\in\Theta}{\text{argmin}} S_n^*$. As pointed out in Section 3.3, by Taylor expansion, we have:

$$
\sqrt{n}(D_n Q_n(\theta)|_{\tilde{\theta}} D_n) D_n^{-1}(\hat{\theta}^* - \theta^0) = -\sqrt{n} D_n T_n(\theta^0).
$$

As a result, the proof completes if $(D_n Q_n(\theta)|_{\tilde{\theta}} D_n)$ and $\sqrt{n} D_n T_n(\theta^0)$ are proved to converge to the corresponding matrix and vector present in Theorem 3.

We first prove that,

$$
\sqrt{n} D_n T_n(\theta^0) \overset{d}{\Longrightarrow} \begin{pmatrix} -\int B d \, \sum_{j=1}^{m-1} \sigma_{v_j} W_j \\ \sigma_{v_1} W_1(1) \\ \sigma_{v_2} W_2(1) \\ \vdots \\ \sigma_{v_{m-1}} W_{m-1}(1) \\ 0 \end{pmatrix}. \qquad \text{(S1.15)}
$$

Let $e_t(\theta)$ be the residual of model (S1.14) when θ is plugged in, then

$$
\begin{aligned}
& e_t(\theta) \\
=& u_t - \sum_{j=1}^{m} (A_j - A_j^0)' X_{t-1} 1_{t-1,j} - \sum_{j=1}^{m} A_j' X_{t-1}(\mathcal{K}_{t-1}(\eta, \gamma_{j-1}, \gamma_j) - 1_{t-1,j}) \\
& - \sum_{j=1}^{m} A_{j,z}' \mathcal{K}_{t-1}(\eta, \gamma_{j-1}, \gamma_j) \frac{x_{2,t-1}' \eta}{\sqrt{n}},
\end{aligned}
$$

where $1_{t-1,j} = 1\{\gamma_{j-1} \leq z_{t-1}(\beta) < \gamma_j\}$ and $\mathcal{K}_{t-1}(\eta, \gamma_{j-1}, \gamma_j)$=$\mathcal{K}_{t-1,h}(\beta, \gamma_{j-1}, \gamma_j)$ are defined in Section 3. Denote $\phi_{n,t}^{\gamma_j} = \frac{\partial e_t'(\theta^0)}{\partial \gamma_j} e_t(\theta^0)$ for $j = 1, 2, \ldots, m-1$, then $\frac{\partial S_n^*}{2n\partial\gamma_j}|_{\theta_0} = \frac{1}{n}\sum_{t=1}^{n} \phi_{n,t}^{\gamma_j}$.

For simplicity, let $D_{\mathcal{K},I,j} = \mathcal{K}(\frac{\gamma_j^0 - z_{t-1}}{h}) - 1\{z_{t-1} < \gamma_j^0\}$. From Assumption 3.1, it is obvious that

$$\mathcal{K}_{t-1}(\eta, \gamma_{j-1}, \gamma_j) = \mathcal{K}(\frac{\gamma_j - z_{t-1}(\beta)}{h}) - \mathcal{K}(\frac{\gamma_{j-1} - z_{t-1}(\beta)}{h}).$$

Further, since $e_t(\theta^0) = u_t - \sum_{j=1}^{m-1}(A_j^0 - A_{j+1}^0)^{'} X_{t-1} D_{\mathcal{K},I,j}$, and $\frac{\partial e_t(\theta)}{\partial \gamma_j} = -(A_j - A_{j+1})^{'} X_{t-1} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j - z_{t-1}(\beta)}{h})}{h} - A_{j,z}^{'} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j - z_{t-1}(\beta)}{h})}{h} \frac{x_{2,t-1}^{'}\eta}{\sqrt{n}}$, we have

$$\begin{aligned}
&\sqrt{nh}\mathrm{E}(\phi_{n,t}^{\gamma_j}) \\
=&\sqrt{nh}\mathrm{E}\left[-X_{t-1}^{'}(A_j^0 - A_{j+1}^0)u_t \mathcal{K}^{(1)}\left(\frac{\gamma_j^0 - z_{t-1}}{h}\right)/h\right] \\
&+ \sqrt{nh}\mathrm{E}\left[X_{t-1}^{'}(A_j^0 - A_{j+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1} D_{\mathcal{K},I,j} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h}\right] \\
&+ \sqrt{nh} \sum_{k=1, k\neq j}^{m-1} \mathrm{E}\left[X_{t-1}^{'}(A_k^0 - A_{k+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1} D_{\mathcal{K},I,k} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h}\right].
\end{aligned}$$

The third term is $o_p(1)$, as a result of $|\gamma_j^0 - \gamma_k^0|$ being bounded away form 0 and the property of $\mathcal{K}$ (Assumption 3.3), Assumption 1.3 and Assumpitons 3. The first term is $o_p(1)$ because $\mathrm{E}(u_t)$=0 and independence between u_t and X_{t-1}. The second term has limit 0 because $\int_{-\infty}^{\infty}(1\{s > 0\} - \mathcal{K}(s))\mathcal{K}^{'}(s)\ ds = 0$ and

$$\begin{aligned}
&\mathrm{E}\left[X_{t-1}^{'} X_{t-1}\left(\mathcal{K}(\frac{\gamma_j^0 - z_{t-1}}{h}) - 1\{z_{t-1} < \gamma_j^0\}\right)\frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h}\right] \\
=&\int X^{'} X(\mathcal{K}(s) - 1\{s > 0\})\mathcal{K}^{(1)}(s) f_{Z|X_2}(\gamma_j^0 - hs|x_2)\ \mathrm{d}(-s)\mathrm{d}F_{X_2}(x_2) \\
=&h\int X^{'} X(\mathcal{K}(s) - 1\{s > 0\})\mathcal{K}^{(1)}(s)(-s f_{Z|X_2}^{'}(\gamma_j^0|x_2) + O(s^2h^2) + \cdots)\ \mathrm{d}(-s)\ \mathrm{d}F_{X_2}(x_2).
\end{aligned} \tag{S1.16}$$

Hence, under Assumption 3.3, the r.h.s of (S1.16) is $O(h)$ and further the second term of $\sqrt{nh}\mathrm{E}(\phi_{n,t}^{\gamma_j})$ is $O(\sqrt{nh}h)$. Therefore, by Assumption 3.4,

$$\lim_{n\to\infty} \sqrt{nh}\mathrm{E}\frac{\partial e_t^{'}(\theta^0)}{\partial \gamma_j} e_t(\theta^0) = 0. \tag{S1.17}$$

Further,

$$\begin{aligned}
&h\mathrm{E}[\phi_{n,t}^{\gamma_j}(\theta^0)^2] \\
=&h\mathrm{E}\left\{\left(X_{t-1}^{'}(A_j^0 - A_{j+1}^0)u_t \frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h}\right)^2\right. && (A.18-1) \\
&+ \sum_{k=1}^{m-1}\left[X_{t-1}^{'}(A_k^0 - A_{k+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1} D_{\mathcal{K},I,k} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h}\right]^2 && (A.18-2)
\end{aligned}$$

$$+ 2 \sum_{j_1 \neq j_2 \neq j} \left[X_{t-1}^{'}(A_{j_1}^0 - A_{j_1+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1} D_{\mathcal{K},I,j_1} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h} \right]$$
$$\left[X_{t-1}^{'}(A_{j_2}^0 - A_{j_2+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1} D_{\mathcal{K},I,j_2} \frac{\mathcal{K}^{(1)}(\frac{\gamma_j^0 - z_{t-1}}{h})}{h} \right]$$
$$+ 2 \sum_{k=1}^{m-1} X_{t-1}^{'}(A_{k+1}^0 - A_k^0)(A_{j+1}^0 - A_j^0)^{'} X_{t-1} u_t^{'}(A_{j+1}^0 - A_j^0)^{'} X_{t-1} D_{\mathcal{K},I,k}$$
$$\left.(\frac{\mathcal{K}^{(1)}((\gamma_j^0 - z_{t-1})/h)}{h})^2 \right\}. \tag{S1.18}$$

We now analyze the r.h.s of (S1.18), we can prove that the quantity in (A.18-1) converges to

$$\left|\left|\mathcal{K}^{'}\right|\right|_2^2 \mathrm{E}\left[(X_{t-1}^{'}(A_j^0 - A_{j+1}^0)u_t)^2 | z_{t-1} = \gamma_j^0\right] f_Z(\gamma_j^0),$$

by similar analysis as in (S1.16).

Let $\tilde{\mathcal{K}_1} = (\mathcal{K}(s) - 1\{s > 0\})\,\mathcal{K}^{(1)}(s)$, then similarly, the quantity in (A.18-2) converges to

$$\left|\left|\tilde{\mathcal{K}_1}\right|\right|_2^2 E\left[\left(X_{t-1}^{'}(A_j^0 - A_{j+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1}\right)^2 | z_{t-1} = \gamma_j^0\right] f_Z(\gamma_j^0),$$

and the third and fourth terms of the r.h.s of (S1.18) converge to zero by Assumptions 1 and 3. Consequently, $\lim_{n\to\infty} \mathrm{Var}\left(\sqrt{nh}\frac{\partial S_n^*(\theta^0)}{2n\partial\gamma_j}\right) = \lim_{n\to\infty} \mathrm{Var}(\sqrt{h}\phi_{n,t}^{\gamma_j}) = \sigma_{v_j}^2$.

By a similar reasoning, we prove $\lim_{n\to\infty} h\mathrm{cov}\left[\phi_{n,t}^{\gamma_j}(\theta^0), \phi_{n,t}^{\gamma_l}(\theta^0)\right] = 0$. Further, by Lemmas in Horowitz (1992) and Theorem 2 in de Jong (1997),

$$\sqrt{nh}T_n^{\gamma} \to N(0, V),$$

where T_n^{γ} stands for derivatives with respect to γ, $V = \mathrm{diag}(\sigma_{v_1}^2, \ldots, \sigma_{v_{m-1}}^2)$ and

$$\sigma_{v_j}^2 = \mathrm{E}\left[\left|\left|\mathcal{K}^{(1)}\right|\right|_2^2 [X_{t-1}^{'}(A_j^0 - A_{j+1}^0)u_t]^2 + \left|\left|\tilde{\mathcal{K}_1}\right|\right|_2^2 (X_{t-1}^{'}(A_j^0 - A_{j+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1})^2 | z_{t-1} = \gamma_j^0\right] f_Z(\gamma_j^0).$$

Further,

$$\frac{\partial e_t^{'}(\theta^0)}{\partial\eta} e_t(\theta^0)$$
$$= \frac{x_{2,t-1}}{\sqrt{n}} \left[\sum_{k=1}^{m-1} (A_k^0 - A_{k+1}^0)^{'} X_{t-1} \frac{\mathcal{K}^{(1)}(\frac{\gamma_k^0 - z_{t-1}}{h})}{h} - \sum_{k=1}^{m-1} (A_{k,z}^0 - A_{k+1,z}^0)^{'} \mathcal{K}(\frac{\gamma_k^0 - z_{t-1}}{h})\right]$$
$$\left(u_t - \sum_{k=1}^{m-1} (A_k^0 - A_{k+1}^0)^{'} X_{t-1} \left(\mathcal{K}(\frac{\gamma_k^0 - z_{t-1}}{h}) - 1\{z_{t-1} < \gamma_k^0\}\right)\right).$$

Let

$$\nu_t = \sum_{k=1}^{m-1} -(A_{k,z}^0 - A_{k+1,z}^0)^{'} u_t \mathcal{K}(\frac{\gamma_k^0 - z_{t-1}}{h}),$$

$$\nu_{j,t} = (A^0_{j,z} - A^0_{j+1,z})' \mathcal{K}(\frac{\gamma^0_j - z_{t-1}}{h})$$
$$\sum_{k=1}^{m-1} (A^0_k - A^0_{k+1})' X_{t-1} \left(\mathcal{K}(\frac{\gamma^0_k - z_{t-1}}{h}) - 1\{z_{t-1} < \gamma^0_k\} \right),$$
$$\tilde{\nu}_{j,t} = (A^0_j - A^0_{j+1})' \frac{\mathcal{K}^{(1)}(\frac{\gamma^0_j - z_{t-1}}{h})}{h}$$
$$\left(u_t - \sum_{k=1}^{m-1} (A^0_k - A^0_{k+1})' X_{t-1} \left(\mathcal{K}(\frac{\gamma^0_k - z_{t-1}}{h}) - 1\{z_{t-1} < \gamma^0_k\} \right) \right),$$

then

$$\sqrt{nh}\frac{1}{n}\sum_t \frac{\partial e'_t(\theta^0)}{\partial \eta} e_t(\theta^0)$$
$$= \frac{1}{n}\sum_{t=1}^{n} x_{2,t-1} \left(-\sqrt{h}\nu_t + \sum_{j=1}^{m-1} \sqrt{h}\nu_{j,t} + \sum_{j=1}^{m-1} \sqrt{h}\tilde{\nu}_{j,t} \right). \tag{S1.19}$$

It can be shown that the terms of the form

$$\sqrt{\frac{h}{n}} \frac{\mathcal{K}^{(a)}(\frac{\gamma^0_j - z_{t-1}}{h})}{h} \left(\mathcal{K}(\frac{\gamma^0_k - z_{t-1}}{h}) - 1\{z_{t-1} < \gamma^0_k\} \right), a = 0, 1, \tag{S1.20}$$

converges to zero by Assumption 3, for $k \neq j$. Using (S1.20) and the arguments in Seo (2011), we study the quantity on the r.h.s. of (S1.19) as follows. First, the term $\frac{1}{n}\sum_{t=1}^{n} x_{2,t-1} \sum_{j=1}^{m-1} \sqrt{h}\tilde{\nu}_{j,t}$ dominates the r.h.s of (S1.19) and converges to $-\int_0^1 B \left(\sum_{j=1}^{m-1} \sigma_{v_j} \mathrm{d}W_j(s) \right)$, where W_j are independent standard Brownian motions. Consequently, the first $p-1$ entries of $\sqrt{n} D_n T_n(\theta^0)$, $\sqrt{h}/n \sum_{t=1}^{n} \frac{\partial e'_t(\theta^0)}{\partial \beta} e_t(\theta^0)$, converges to $-\int_0^1 B \left(\sum_{j=1}^{m-1} \sigma_{v_j} \mathrm{d}W_j(s) \right)$. Moreover, it can be shown that the $p-1+j$-th entry of $\sqrt{n} D_n T_n(\theta^0)$ converges to $\sigma_{v_j} W_j(1)$ using the preceding established properties of $\phi^{\gamma_j}_{n,t}$. Finally, the convergence of the remaining entries of $\sqrt{n} D_n T_n(\theta^0)$, $\sqrt{n}\frac{\partial S^*_n}{2n\partial\lambda}$, to zero in probability is standard and can be proved with similar arguments as in Seo and Hansen (2007). Therefore, The proof of the convergence of $\sqrt{n} D_n T_n(\theta^0)$ is completed.

Next we prove the convergence of $D_n Q_n(\tilde{\theta}) D_n$. Since

$$D_n Q_n(\tilde{\theta}) D_n = \left[D_n \sum_{t=1}^{n} (\frac{\partial^2 e_t(\theta)}{\partial\theta\partial\theta'}) e_t(\theta) D_n + D_n \sum_{t=1}^{n} \frac{\partial e_t(\theta)}{\partial\theta} \frac{\partial\{e_t(\theta)\}}{\partial\theta'} D_n \right]_{\theta=\tilde{\theta}}$$
$$\triangleq \left[D_n Q_n(\theta)^a D_n + D_n Q_n(\theta)^b D_n \right]_{\theta=\tilde{\theta}},$$

we need to prove that $D_n Q^a_n(\theta) D_n$ and $D_n Q_n(\theta)^b D_n$ converge in distribution to the following random matrices in a small enough neighborhood of θ^0:

$D_n Q_n(\theta)^b D_n$

$$\stackrel{d}{\Rightarrow} \left|\left|\mathcal{K}^{(1)}\right|\right|_2^2 \begin{pmatrix} \tilde{\sigma}_q^2 \int_0^1 B(s)B(s)^{'}\mathrm{d}s & \tilde{\sigma}_{q_1}^2 \int_0^1 B(s)\mathrm{d}s & \cdots & \tilde{\sigma}_{q_{m-1}}^2 \int_0^1 B(s)\mathrm{d}s & 0 \\ \tilde{\sigma}_{q_1}^2 \int_0^1 B(s)^{'}\mathrm{d}s & \tilde{\sigma}_{q_1}^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{\sigma}_{q_{m-1}}^2 \int_0^1 B(s)^{'}\mathrm{d}s & 0 & \cdots & \tilde{\sigma}_{q_{m-1}}^2 & 0 \\ 0 & 0 & \cdots & 0 & N/\left|\left|\mathcal{K}^{(1)}\right|\right|_2^2 \end{pmatrix}, \tag{S1.21}$$

where $\tilde{\sigma}_{q_j}^2 = \mathrm{E}[X_{t-1}^{'}(A_j^0 - A_{j+1}^0)(A_j^0 - A_{j+1}^0)^{'} X_{t-1})|z_{t-1} = \gamma_j^0] f_Z(\gamma_j^0)$, $\tilde{\sigma}_q^2 = \sum_{j=1}^{m-1} \tilde{\sigma}_{q_j}^2$, and

$$N = \left[\mathrm{E} \left(\left(\begin{array}{cccc} I_1 & 0 & \dots & 0 \\ 0 & I_2 & 0 & 0 \\ & & \dots & \\ 0 & 0 & \dots & I_m \end{array} \right) \otimes X_{t-1}X_{t-1}^{'} \right) \right]^{-1} \otimes \Sigma.$$

$$D_n Q_n(\theta)^a D_n$$

$$\stackrel{d}{\Longrightarrow} \int \tilde{\tilde{\mathcal{K}}}(s)\mathrm{d}s \begin{pmatrix} \tilde{\sigma}_q^2 \int_0^1 B(s)B(s)^{'}\mathrm{d}s & \tilde{\sigma}_{q_1}^2 \int_0^1 B(s)\mathrm{d}s & \cdots & \tilde{\sigma}_{q_{m-1}}^2 \int_0^1 B(s)\mathrm{d}s & 0 \\ \tilde{\sigma}_{q_1}^2 \int_0^1 B(s)^{'}\mathrm{d}s & \tilde{\sigma}_{q_1}^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{\sigma}_{q_{m-1}}^2 \int_0^1 B(s)^{'}\mathrm{d}s & 0 & \cdots & \tilde{\sigma}_{q_{m-1}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{S1.22}$$

where $\tilde{\tilde{\mathcal{K}}} = -\mathcal{K}^{(2)}(s)(\mathcal{K}(s) - 1\{s > 0\})$ and $\left|\left|\mathcal{K}^{(1)}\right|\right|_2^2 + \int \tilde{\tilde{\mathcal{K}}} = \mathcal{K}^{(1)}(0)$.

To prove the convergences of $D_n Q_n(\theta)^a D_n$ and $D_n Q_n(\theta)^b D_n$, we only need to prove convergence of the elements that have non-zero limits. The convergence of these elements can be proved similarly, so we show the proof of the following quantity as an example,

$$h \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 e_t(\theta)}{\partial\eta\partial\eta^{'}} e_t(\theta)|_{\tilde{\theta}} \stackrel{d}{\Longrightarrow} \int \tilde{\tilde{\mathcal{K}}}\mathrm{d}s \tilde{\sigma}_q^2 \int_0^1 B(s)B(s)^{'}\mathrm{d}s.$$

Since

$$\begin{aligned} \frac{\partial^2 e_t(\theta)}{\partial\eta\partial\eta^{'}} e_t(\theta) =& \frac{-x_{2,t-1}}{\sqrt{n}h} \frac{x_{2,t-1}^{'}}{\sqrt{n}h} \left(\sum_{j=1}^{m-1} (A_j - A_{j+1}) X_{t-1}(\eta) \mathcal{K}^{(2)} \left(\frac{\gamma_j - z_{t-1}(\eta)}{h} \right) \right. \\ & \left. - \sum_{j=1}^{m-1} (A_{j,z} - A_{j+1,z}) \mathcal{K}^{(1)} \left(\frac{\gamma_j - z_{t-1}(\eta)}{h} \right) h \right) e_t(\theta), \end{aligned}$$

and with similar reasoning as in Lemma 2 of Seo (2011), we obtain $h^{-1}(\hat{\gamma}^* - \gamma^0) = o_p(1)$ and $h^{-1}(\hat{\beta}^* - \beta^0) = o_p(1)$, the dominating part of $\frac{\partial^2 e_t(\theta)}{\partial\eta\partial\eta^{'}} e_t(\theta)$ is

$$\frac{-x_{2,t-1}}{\sqrt{n}h} \frac{x_{2,t-1}^{'}}{\sqrt{n}h} \sum_{j=1}^{m-1} (A_j - A_{j+1}) X_{t-1} \mathcal{K}^{(2)} \left(\frac{\gamma_j - z_{t-1}(\eta)}{h} \right) e_t(\theta). \tag{S1.23}$$

Then, with the same reasoning as the proof of convergence of $\frac{\partial e_t(\theta^0)}{\partial \eta} e_t(\theta^0)$, we obtain

$$\frac{h}{n^2}\sum_{t=1}^{n}\frac{\partial^2 e_t(\theta)}{\partial\beta\partial\beta'}e_t(\theta) = \frac{h}{n}\sum_{t=1}^{n}\frac{\partial^2 e_t(\theta)}{\partial\eta\partial\eta'}e_t(\theta) \overset{d}{\Rightarrow} \int \tilde{\tilde{\mathcal{K}}}(s)\mathrm{d}s\tilde{\sigma}_q^2\int_0^1 B(s)B(s)'\mathrm{d}s,$$

when θ is in between θ^0 and $\hat{\theta}^*$. The convergence of $\frac{\partial e_t(\theta)}{\partial \lambda}\frac{\partial e_t(\theta)}{\partial \lambda'}$ is standard and is omitted. Notice that $B, W_1, \ldots, W_{m-1}$ are all symmetric, $-D_n T_n(\theta^0)$ has the same limit as $D_n T_n(\theta^0)$, therefore, the proof of this theorem is completed.