A Specification Test for Dynamic Conditional Distribution Models with Function-Valued Parameters

Appendix

A.1. Preliminary Results

In this subsection, we provide preliminary results used in the proofs of the propositions. Let $\mathcal{G}$ be a permissible class of functions in such a way that the following holds: (a) $\mathcal{T}$ is a Suslin metric space (a Hausdorff topological space that is the continuous image of a Polish space) with Borel σ-field $\mathcal{B}(\mathcal{T}, \Theta)$, and (b) $g(\cdot, \cdot, \cdot)$ is a $\mathcal{B}(\mathcal{T}, \Theta)$-measurable function from $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^K$ to $\mathbb{R}$ (see Kosorok, 2007, Section 11.6). Let $E_Q g = \int g(W_t, \theta, \tau) dQ(W_t, \theta, \tau)$, for $g \in \mathcal{G}$, with $\mathcal{G} := \{W_t \mapsto g(W_t, \theta, \tau) : \theta \in \Theta, \tau \in \mathcal{T}\}$. We assume that the $\mathcal{G}$ class of functions forms a so-called Vapnik-Chervonenkis subgraph (VC-subgraph) class of functions (see Dudley, 1978). The VC-subgraph class is an extension of the class of indicator functions and is useful for most statistical applications (Arcones and Yu, 1994; Radulović, 1996). If $\mathcal{G}$ is a VC-subgraph class, then for any given $1 \leq p < \infty$, there are constants a and b satisfying

$$N(\varepsilon, \mathcal{G}, \|\cdot\|) \leq a \left(\frac{(E_Q |\mathbb{F}|^p)^{1/p}}{\varepsilon} \right)^b,$$

for all $\varepsilon > 0$ and all probability measures Q, with $E_Q|\mathbb{F}|^p < \infty$, where $N(\varepsilon, \mathcal{G}, \|\cdot\|)$ is the covering number of $\mathcal{G}$ with respect to $\|\cdot\|$, i.e., the minimal number of $L_2(Q)$-balls of radius ε needed to cover $\mathcal{G}$, where a $L_2(Q)$-ball of radius ε around a function $g \in L_2(Q)$ is the set $\{h \in L_2(Q) : \|h - g\| < \varepsilon\}$ (see Pollard, 1984). Moreover, the class of functions $\mathcal{G}$ has a finite and integrable envelope function $\mathbb{F} := \sup_{g \in \mathcal{G}} |g(W_t, \theta, \tau)|$, and it can be covered by a finite number of elements, not necessarily in $\mathcal{G}$.

The following result establishes a central limit theorem for strong mixing processes for the empirical distribution, $\hat{Z}_T(y, x)$, under the null and the alternative hypothesis.

Lemma A.1. If Assumption 1 holds, under H_0 of (2.1) or H_A of (2.2),

$$v_T(y, x) := \sqrt{T}(\hat{Z}_T(y, x) - F_{YX}(y, x)) \Rightarrow \mathbb{H}_1(y, x), \text{ in } \ell^\infty(\mathcal{W}),$$

where $\mathbb{H}_1$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$ with covariance function

$$\mathrm{Cov}(\mathbb{H}_1(y, x), \mathbb{H}_1(y', x')) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(\mathbb{1}\{Y_0 \leq y\}\mathbb{1}\{X_0 \leq x\}, \mathbb{1}\{Y_k \leq y'\}\mathbb{1}\{X_k \leq x'\}).$$

Proof: Parts (a) and (b) of Assumption 1 imply the conditions (2.3) and (2.4) in Theorem 2.1 in Arcones and Yu (1994), respectively. Then the results follow from a direct application of Theorem 2.1 in Arcones and Yu (1994). $\square$

The following result establishes a functional delta method for the empirical analog $\hat{G}(\theta, \tau)$ of the moment conditions in (2.4) and for a consistent estimator of the function-valued parameter $\hat{\theta}_T(\cdot)$.

Lemma A.2. Let $v_T(y, x) := \sqrt{T}(\hat{Z}_T(y, x) - F_{YX}(y, x))$ be the empirical process of Lemma A.1 and define the empirical process $r_T(\theta, \tau) := \sqrt{T}(\hat{G}(\theta, \tau) - G(\theta, \tau))$. If Assumption 1 is satisfied, under H_0 of (2.1) or H_A of (2.2),

$$
\begin{aligned}
(v_T(y,x), r_T(\theta,\tau)) &\Rightarrow (\mathbb{H}_1(y,x), \tilde{\mathbb{H}}_2(\theta,\tau)), \text{ in } \ell^\infty(\mathcal{W} \times \Theta \times \mathcal{T}), \\
\sqrt{T}(\hat{\theta}_T(\cdot) - \theta_0(\cdot)) &\Rightarrow -\dot{G}^{-1}_{\theta_0,\cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot),\cdot)) \text{ in } \ell^\infty(\mathcal{T}),
\end{aligned}
$$

where $\tilde{\mathbb{H}}_2$ is a tight mean zero Gaussian process with covariance function

$$
\mathrm{Cov}(\tilde{\mathbb{H}}_2(\theta,\tau), \tilde{\mathbb{H}}_2(\theta',\tau')) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(g(W_0,\theta,\tau), g(W_k,\theta',\tau')).
$$

Proof: By Lemma E.1 in Chernozhukov et al. (2013), Assumption 1 implies that (a) the inverse of $G(\cdot,\tau)$ defined as $G^{-1}(x,\tau) := \{\theta \in \Theta : G(\theta,\tau) = x\}$ is continuous at $x = 0$ uniformly in $\tau \in \mathcal{T}$ with respect to the Hausdorff distance, (b) there exists $\dot{G}_{\theta_0,\tau}$ such that

$$
\lim_{t\to 0} \sup_{\tau\in\mathcal{T}, \|h\|=1} |t^{-1}(G(\theta_0(\tau)+th,\tau) - G(\theta_0(\tau),\tau)) - \dot{G}_{\theta_0,\tau} h| = 0,
$$

where $\inf_{\tau\in\mathcal{T}} \inf_{\|h\|=1} \|\dot{G}_{\theta_0,\tau} h\| > 0$, (c) the maps $\tau \mapsto \theta_0(\tau)$ and $\tau \mapsto \dot{G}_{\theta_0,\tau}$ are continuous, and (d) the mapping $\tau \mapsto \theta_0(\tau)$ is continuously differentiable. Under the previous conditions, Lemma E.2 in Chernozhukov et al. (2013) holds, and the process $r_T(\theta,\tau)$ weakly converges to $\tilde{\mathbb{H}}_2(\theta,\tau)$ in $\ell^\infty(\Theta \times \mathcal{T})$ and the map $\theta \mapsto G(\theta,\cdot)$ is Hadamard differentiable at θ_0 with continuously invertible derivative $\dot{G}_{\theta_0,\cdot}$. By Hadamard differentiability of the map $\theta \mapsto G(\theta,\cdot)$, it follows the weak convergence of the process $\sqrt{T}(\hat{\theta}_T(\cdot) - \theta_0(\cdot))$ in $\ell^\infty(\mathcal{T})$.

□

Lemma A.3. Let $v_T(y,x) := \sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x))$ be the empirical process of Lemma A.1 and define the empirical process $v_T^{\theta_0}(y,x) := \sqrt{T}(\hat{F}_T(y,x,\hat{\theta}_T) - F(y,x,\theta_0))$. If Assumption 1 holds, under H_0 of (2.1) or H_A of (2.2),

$$
(v_T(y,x), v_T^{\theta_0}(y,x)) \Rightarrow (\mathbb{H}_1(y,x), \mathbb{H}_2(y,x)) \text{ in } \ell^\infty(\mathcal{W} \times \mathcal{W}),
$$

where $\mathbb{H}_2$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$.

Proof: By Lemma A.2, $\sqrt{T}(\hat{\theta}_T(\cdot) - \theta_0(\cdot)) \Rightarrow -\dot{G}^{-1}_{\theta_0,\cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot), \cdot))$ in $\ell^\infty(\mathcal{T})$, where $\tilde{\mathbb{H}}_2$ is a tight mean zero Gaussian process. Similarly to Lemma A.1, under H_0 of (2.1) or H_A of (2.2), if parts (a)-(b) of Assumption 1 hold, then $\sqrt{T}(\hat{F}_X(x^*) - F_X(x^*))$ weakly converges to a tight mean zero Gaussian process. Now, let the measurable functions $\Gamma : \mathcal{W} \mapsto [0,1]$ be defined by $(y,x) \mapsto \Gamma(y,x)$, and the bounded maps $\Pi : \mathcal{H} \mapsto \mathbb{R}$ be defined by $f \mapsto \int f d\Pi$. Then it follows from Lemma D.1 in Chernozhukov et al. (2013) that the mapping $(\Gamma, \Pi) \mapsto \int \Gamma(\cdot, x) d\Pi(x)$, with $\Gamma(\cdot, x) = \mathbb{1}\{x^* \leq x\} F(\cdot|x)$ and $\Pi = F_X(\cdot)$, is well defined and Hadamard differentiable at (Γ, Π). Under H_0 of (2.1) or H_A of (2.2), we can write $\hat{F}_T(y, x, \hat{\theta}_T) = \int F(y|x^*, \hat{\theta}_T) \mathbb{1}\{x^* \leq x\} d\hat{F}_X(x^*)$ and $F(y, x, \theta_0) = \int F(y|x^*) \mathbb{1}\{x^* \leq x\} dF_X(x^*)$. Then, by the functional delta method from Lemma B.1 of Chernozhukov et al. (2013), it follows that

$$
\begin{aligned}
\sqrt{T}(\hat{F}_T(y, x, \hat{\theta}_T) - F(y, x, \theta_0)) &= \int \sqrt{T} \left[F(y|x^*, \hat{\theta}_T) - F(y|x^*) \right] \mathbb{1}\{x^* \leq x\} dF_X(x^*) \\
&+ \int F(y|x^*) \mathbb{1}\{x^* \leq x\} \sqrt{T} d\left(\hat{F}_X(x^*) - F_X(x^*) \right) + o_p(1).
\end{aligned}
$$

Using the same arguments of the Proof of Lemma A.2, we can show that the map $\theta \mapsto F(\cdot|\cdot, \theta(\cdot))$ is Hadamard differentiable. Thus, we apply the functional delta method, for fixed y and x, as follows:

$$
\begin{aligned}
\sqrt{T}\left(F(y|x, \hat{\theta}_T) - F(y|x) \right) &\Rightarrow -\dot{F}(y|x, \theta_0)(-\dot{G}^{-1}_{\theta_0,\cdot}(\tilde{\mathbb{H}}_2(\theta_0(\cdot), \cdot))) \\
&:= \mathbb{H}_2^*(y, x) \text{ in } \ell^\infty(\mathcal{W}).
\end{aligned}
$$

Given the Hadamard differentiability of the mapping $(\Gamma, \Pi) \mapsto \int \Gamma(\cdot, x) d\Pi(x)$, the result follows from an application of the functional delta method, where the Gaussian process $\mathbb{H}_2$

is given by

$$\mathbb{H}_2(y,x) := \int \mathbb{H}_2^*(y,x^*)\mathbb{1}\{x^* \le x\} dF_X(x^*) + \int F(y|x^*)\mathbb{1}\{x^* \le x\} d\mathbb{H}_1(\infty, x^*),$$

where $\mathbb{H}_1$ is the tight mean zero Gaussian process defined in Lemma A.1. $\square$

Lemma A.4. Under the local alternatives $H_{A,T}$ in (2.7) and Assumptions 1-2, let $F_T^A(y,x) = \int F_T(y|x^*)\mathbb{1}\{x^* \le x\} dF_X(x^*)$ and $G_{F_T}(\theta,\tau) = E_{F_T}[g(W_t,\theta,\tau)]$, then:

$$\begin{pmatrix} \sqrt{T}\left(\hat{Z}_T(y,x) - F_T^A(y,x)\right) \\ \sqrt{T}\left(\hat{G}(\theta,\tau) - G_{F_T}(\theta,\tau)\right) \end{pmatrix} \Rightarrow \begin{pmatrix} \mathbb{H}_1(y,x) \\ \tilde{\mathbb{H}}_2(\theta,\tau) \end{pmatrix} \text{ in } \ell^\infty(\mathcal{W}\times\Theta\times\mathcal{T}),$$

where $(\mathbb{H}_1, \tilde{\mathbb{H}}_2)$ are the tight mean zero Gaussian processes defined in Lemmas A.1-A.2.

Proof: Under Assumption 2, $F_T^A(y,x)$ is contiguous to $F(y,x,\theta_0)$, then under the sequence of local alternatives $H_{A,T}$ in (2.7) and Assumptions 1-2, $F_T(y|x)$ of (2.7) is a linear combination of two measures that are VC-subgraph class with a p-integrable envelope, for some $2 < p < \infty$. From an application of Lemma 2.8.7 in Van der Vaart and Wellner (2000), we have that $(\sqrt{T}(\hat{Z}_T(y,x) - F_T^A(y,x)), \sqrt{T}(\hat{G}(\theta,\tau) - G_{F_T}(\theta,\tau))) \Rightarrow (\mathbb{H}_1(y,x), \tilde{\mathbb{H}}_2(\theta,\tau))$ in $\ell^\infty(\mathcal{W}\times\Theta\times\mathcal{T})$.
$\square$

Lemma A.5. Under the DGP Size.1_C: $Y_t = 0.2 + 0.3Y_{t-1} + (1 + 0.1Y_{t-1})u_t$, where u_t follows an i.i.d process with distribution $\mathcal{N}(0,1)$, the conditional quantile function of Y_t given Y_{t-1} is given by:

(i) $Q_\tau(Y_t|Y_{t-1}) = (\Phi_u^{-1}(\tau) + 0.2) + (0.1\Phi_u^{-1}(\tau) + 0.3)Y_{t-1}$, when $1 + 0.1Y_{t-1} > 0$;

(ii) $Q_\tau(Y_t|Y_{t-1}) = (\Phi_u^{-1}(1-\tau) + 0.2) + (0.1\Phi_u^{-1}(1-\tau) + 0.3)Y_{t-1}$, when $1 + 0.1Y_{t-1} < 0$;

(iii) $Q_\tau(Y_t|Y_{t-1}) = 0.2 + 0.3Y_{t-1}$, when $1 + 0.1Y_{t-1} = 0$.

Proof: (i) $1 + 0.1Y_{t-1} > 0$:

$$\begin{aligned} F_{Y_t}(y|Y_{t-1}) &= P(Y_t \leq y|Y_{t-1}) \\ &= P(0.2 + 0.3Y_{t-1} + (1 + 0.1Y_{t-1})u_t \leq y|Y_{t-1}) \\ &= P((1 + 0.1Y_{t-1})u_t \leq y - 0.2 - 0.3Y_{t-1}|Y_{t-1}) \\ &= P\left(u_t \leq \frac{y - 0.2 - 0.3Y_{t-1}}{1 + 0.1Y_{t-1}}|Y_{t-1}\right) \\ &= \Phi_u\left(\frac{y - 0.2 - 0.3Y_{t-1}}{1 + 0.1Y_{t-1}}\right) \stackrel{!}{=} \tau \end{aligned}$$

The last line is equivalent to

$$\begin{aligned} \frac{y - 0.2 - 0.3Y_{t-1}}{1 + 0.1Y_{t-1}} &= \Phi_u^{-1}(\tau) \\ \Leftrightarrow y - 0.2 - 0.3Y_{t-1} &= \Phi_u^{-1}(\tau) + 0.1\Phi_u^{-1}(\tau)Y_{t-1} \\ \Leftrightarrow y &= (\Phi_u^{-1}(\tau) + 0.2) + (0.1\Phi_u^{-1}(\tau) + 0.3)Y_{t-1}. \end{aligned}$$

Therefore, $Q_\tau(Y_t|Y_{t-1}) = (\Phi_u^{-1}(\tau) + 0.2) + (0.1\Phi_u^{-1}(\tau) + 0.3)Y_{t-1}$.

(ii) $1 + 0.1Y_{t-1} < 0$:

$$\begin{aligned} F_{Y_t}(y|Y_{t-1}) &= P(Y_t \leq y|Y_{t-1}) \\ &= P(0.2 + 0.3Y_{t-1} + (1 + 0.1Y_{t-1})u_t \leq y|Y_{t-1}) \\ &= P((1 + 0.1Y_{t-1})u_t \leq y - 0.2 - 0.3Y_{t-1}|Y_{t-1}) \\ &= P\left(u_t \geq \frac{y - 0.2 - 0.3Y_{t-1}}{1 + 0.1Y_{t-1}}|Y_{t-1}\right) \\ &= 1 - \Phi_u\left(\frac{y - 0.2 - 0.3Y_{t-1}}{1 + 0.1Y_{t-1}}\right) \stackrel{!}{=} \tau \end{aligned}$$

The last line is equivalent to

$$\begin{aligned}
\frac{y-0.2-0.3Y_{t-1}}{1+0.1Y_{t-1}} &= \Phi_u^{-1}(1-\tau) \\
\Leftrightarrow y-0.2-0.3Y_{t-1} &= \Phi_u^{-1}(1-\tau)+0.1\Phi_u^{-1}(1-\tau)Y_{t-1} \\
\Leftrightarrow y &= (\Phi_u^{-1}(1-\tau)+0.2)+(0.1\Phi_u^{-1}(1-\tau)+0.3)Y_{t-1}.
\end{aligned}$$

Therefore, $Q_\tau(Y_t|Y_{t-1}) = (\Phi_u^{-1}(1-\tau)+0.2)+(0.1\Phi_u^{-1}(1-\tau)+0.3)Y_{t-1}$.

(iii) $1+0.1Y_{t-1}=0$:

$$\begin{aligned}
F_{Y_t}(y|Y_{t-1}) &= P(Y_t \leq y|Y_{t-1}) \\
&= P(0.2+0.3Y_{t-1} \leq y|Y_{t-1}).
\end{aligned}$$

Therefore, $Q_\tau(Y_t|Y_{t-1}) = 0.2+0.3Y_{t-1}$ for $\tau \in (0,1]$ and $Q_\tau(Y_t|Y_{t-1}) = -\infty$ for $\tau = 0$. $\square$

A.2. Proofs of the Propositions

Proof of Proposition 1: To prove part (a), we consider the empirical processes $v_T(y,x) = \sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x))$ and $v_T^{\theta_0}(y,x) = \sqrt{T}(\hat{F}_T(y,x,\hat{\theta}_T) - F(y,x,\theta_0))$ defined in Lemma A.1 and in Lemma A.3, respectively. Under H_0 in (2.1), we have that $F_{YX}(y,x) \equiv F(y,x,\theta_0)$.

Then:

$$
\begin{aligned}
S_T^{CM} &= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T)\right)^2 d\hat{Z}_T(y,x) \\
&= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T) \pm F_{YX}(y,x)\right)^2 d\hat{Z}_T(y,x) \\
&= \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 d\hat{Z}_T(y,x) \\
&= \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 dF_{YX}(y,x) \\
&+ \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 d(\hat{Z}_T(y,x) - F_{YX}(y,x)).
\end{aligned}
$$

By Lemma A.1, $\sqrt{T}(\hat{Z}_T(y,x) - F_{YX}(y,x)) \Rightarrow \mathbb{H}_1(y,x)$, where $\mathbb{H}_1(y,x)$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$. Then, it follows that

$$
S_T^{CM} = \int \left(v_T(y,x) - v_T^{\theta_0}(y,x)\right)^2 dF_{YX}(y,x) + o_p(1).
$$

By Lemma A.3, $(v_T(y,x), v_T^{\theta_0}(y,x)) \Rightarrow (\mathbb{H}_1(y,x), \mathbb{H}_2(y,x))$ in $\ell^\infty(\mathcal{W}\times\mathcal{W})$, where $\mathbb{H}_2(y,x)$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$. Then, the result follows by an application of the continuous mapping theorem.

To prove part (b), under the alternative hypothesis H_A of (2.2), $F_{YX}(y,x) \neq F(y,x,\theta)$ for some $(y,x) \in \mathcal{W}$ and for all $\theta \in \mathcal{B}(\mathcal{T},\Theta)$, and $v_T^{\theta_0}(y,x)$ becomes $v_T^{\theta}(y,x) = \sqrt{T}(\hat{F}_T(y,x,\hat{\theta}_T) - F_T(y,x,\theta))$. Then,

$$
\begin{aligned}
S_T^{CM} &= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T) \pm F_{YX}(y,x) \pm F(y,x,\theta)\right)^2 d\hat{Z}_T(y,x) \\
&= \int \left(v_T(y,x) - v_T^{\theta}(y,x) + \sqrt{T}\left(F_{YX}(y,x) - F(y,x,\theta)\right)\right)^2 dF_{YX}(y,x) + o_P(1).
\end{aligned}
$$

As a corollary of Lemma A.3, $(v_T(y,x), v_T^{\theta}(y,x)) \Rightarrow (\mathbb{H}_1(y,x), \mathbb{H}_2(y,x))$ in $\ell^\infty(\mathcal{W} \times \mathcal{W})$. Therefore, for all fixed constants $C > 0$, we have $\lim_{T\to\infty} P(S_T^{CM} > C) = 1$, and the result follows. $\square$

Proof of Proposition 2: Under the local alternative $H_{A,T}$ in (2.7), consider the empirical processes

$$
\begin{aligned}
v_T^1(y,x) &= \sqrt{T}\Big(\hat{Z}_T(y,x) - \int F(y|x^*, \theta_0)\mathbb{1}\{x^* \le x\} dF_X(x^*)\Big), \text{ and} \\
r_T^1(\theta,\tau) &= \sqrt{T}(\hat{G}(\theta,\tau) - E_F[g(W_t,\theta,\tau)]),
\end{aligned}
$$

where $G_F(\theta,\tau) := E_F[g(W_t,\theta,\tau)]$, with $E_F[\cdot]$ defined as the expectation w.r.t. $F = F(y|x,\theta_0)$ in (2.7). Then

$$
\begin{aligned}
v_T^1(y,x) &= \sqrt{T}\Big(\hat{Z}_T(y,x) - \int F(y|x^*, \theta_0)\mathbb{1}\{x^* \le x\} dF_X(x^*)\Big) \\
&= \sqrt{T}\hat{Z}_T(y,x) \\
&- \sqrt{T}\int \left(F_T(y|x^*) + \frac{\delta}{\sqrt{T}}\left[F(y|x^*,\theta_0) - J(y|x^*)\right]\right) \mathbb{1}\{x^* \le x\} dF_X(x^*) \\
&= \sqrt{T}\left(\hat{Z}_T(y,x) - F_T^A(y,x)\right) + \delta \int [J(y|x^*) - F(y|x^*,\theta_0)]\mathbb{1}\{x^* \le x\} dF_X(x^*).
\end{aligned}
$$

Thus, it follows from Lemma A.4 that

$$
v_T^1(y,x) \Rightarrow \mathbb{H}_1(y,x) + \delta \int [J(y|x^*) - F(y|x^*,\theta_0)]\mathbb{1}\{x^* \le x\} dF_X(x^*),
$$

where $\mathbb{H}_1$ is a tight mean zero Gaussian process in $\ell^\infty(\mathcal{W})$ defined in Lemma A.1. Then

$$\begin{aligned} r_T^1(\theta,\tau) &= \sqrt{T}\left(\hat{G}(\theta,\tau) - E_F[g(W_t,\theta,\tau)]\right) \\ &= \sqrt{T}\left(\hat{G}(\theta,\tau) - (E_{F_T}[g(W_t,\theta,\tau)] + \delta E_F[g(W_t,\theta,\tau)] - \delta E_J[g(W_t,\theta,\tau)])\right) \\ &= \sqrt{T}\left(\hat{G}(\theta,\tau) - G_{F_T}(\theta,\tau) + \delta\left(E_J[g(W_t,\theta,\tau)] - E_F[g(W_t,\theta,\tau)]\right)\right), \end{aligned}$$

where $G_J(\theta,\tau) := E_J[g(W_t,\theta,\tau)]$, with $E_J[\cdot]$ defined as the expectation w.r.t. $J = J(y|x)$ in (2.7). We define the empirical process $v_T^{1\theta_0}(y,x)$ as follows:

$$v_T^{1\theta_0}(y,x) = \sqrt{T}\bigg(\int F(y|x^*,\hat{\theta}_T)\mathbb{1}\{x^* \le x\}d\hat{F}_X(x^*) - \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \le x\}dF_X(x^*)\bigg).$$

By Lemmas A.3-A.4,

$$\begin{pmatrix} v_T^{1\theta_0}(y,x) \\ r_T^1(\theta,\tau) \end{pmatrix} \Rightarrow \begin{pmatrix} \mathbb{H}_2(y,x) + \delta\int \dot{F}(y|x^*)[h]\mathbb{1}\{x^* \le x\}dF_X(x^*) \\ \tilde{\mathbb{H}}_2(\theta,\tau) + \delta\left(E_J[g(W_t,\theta,\tau)] - E_F[g(W_t,\theta,\tau)]\right) \end{pmatrix},$$

with $h(\tau) = (\partial G_F(\theta_0,\tau)/\partial\theta)^{-1}G_J(\theta_0,\tau)$, and where $(\mathbb{H}_2,\tilde{\mathbb{H}}_2)$ are the tight mean zero Gaussian processes described in Lemmas A.2-A.3. Therefore, under $H_{A,T}$ in (2.7),

$$\begin{aligned} S_T^{CM} &= \int T\left(\hat{Z}_T(y,x) - \hat{F}_T(y,x,\hat{\theta}_T) \pm \int F(y|x^*,\theta_0)\mathbb{1}\{x^* \le x\}dF_X(x^*)\right)^2 d\hat{Z}_T(y,x) \\ &= \int\left(v_T^1(y,x) - v_T^{1\theta_0}(y,x)\right)^2 d\hat{Z}_T(y,x) \\ &= \int(v_T^1(y,x) - v_T^{1\theta_0}(y,x))^2 dF_{YX}(y,x) + o_p(1), \end{aligned}$$

Then,

$$S_T^{CM} \xrightarrow{d} \int (\mathbb{H}_1(y,x) - \mathbb{H}_2(y,x) + \Delta(y,x))^2 \, dF_{YX}(y,x),$$

with $\Delta(y,x) = \delta \int (J(y|x^*) - F(y|x^*, \theta_0) + \dot{F}(y|x^*, \theta_0)[h]) \mathbb{1}\{x^* \leq x\} dF_X(x^*)$, and h is the function $h(\tau) = (\partial G_F(\theta_0, \tau)/\partial \theta)^{-1} G_J(\theta_0, \tau)$. □

Proof of Proposition 3: Assumption 1 implies Assumptions 1-2 of Whang (2006). Then, parts (a) and (b) follow from an application of Theorems 2 and 3 of Whang (2006) using the convergence results of our Proposition 1. Further, Assumption 2 implies Assumption 2* of Whang (2006). Therefore, part (c) follows the same steps of Theorem 5 of Whang (2006) using the convergence results of our Proposition 1. □

References

Arcones, M. A., Yu, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* 7:47–71.

Chernozhukov, V., Fernández-Val, I., Melly, B. (2013). Inference on counterfactual distributions. *Econometrica* 81:2205–2268.

Dudley, R. M. (1978). Central limit theorems for empirical measures. *The Annals of Probability* 6:899–929.

Kosorok, M. R. (2007). *Introduction to Empirical Processes and Semiparametric Inference.* Springer Science and Business Media.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

Radulović, D. (1996). The bootstrap for empirical processes based on stationary observations. *Stochastic Processes and Their Applications* 65:259–279.

Van der Vaart, A., Wellner, J. (2000). *Weak Convergence and Empirical Processes: With Applications to Statistics.* New York: Springer.

Whang, Y.-J. (2006). Consistent specification testing for quantile regression models. In: D. Corbae, S. N. Durlauf, and B. E. Hansen (Eds.)*Frontiers of Analysis and Applied Research: Essays in Honor of Peter C. B. Phillips.* Cambridge University Press, pp. 288–310.