Online Supplement for
"Inference in Partially Identified Panel Data Models with Interactive Fixed Effects"

Shengjie Hong[a], Liangjun Su[b], Yaqi Wang[c]

[a] School of Economics, Central University of Finance and Economics

[b] School of Economics and Management, Tsinghua University

[c] School of Finance, Central University of Finance and Economics

This online supplement contains the proofs of the technical lemmas in Appendix A, as well as a discussion on the sufficient conditions for Assumption 3.4 to hold.

B Proofs of Lemmas A.1–A.6

Proof of Lemma A.1. The proof is mainly based on the definitions of MDD_o and MDD as specified in Shao and Zhang (2014) and Su and Zheng (2017), respectively, as well as their properties as shown in these two papers. Let d_Z denote the dimension of Z.

We first prove the first claim. Note that $\text{MDD}_o(W_1|Z)^2 = 0$ if and only if $\mathbb{E}(W_1|Z) = \mathbb{E}(W_1)$, which, in turn, implies that for any given constant vector $s \in \mathbb{R}^{d_Z}$,

$$\begin{aligned}
\text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right) &= \mathbb{E}\left[W_1 \exp\left(\mathbf{i}s'Z\right)\right] - \mathbb{E}\left[W_1\right]\mathbb{E}\left[\exp\left(\mathbf{i}s'Z\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(W_1|Z\right) \exp\left(\mathbf{i}s'Z\right)\right] - \mathbb{E}\left[W_1\right]\mathbb{E}\left[\exp\left(\mathbf{i}s'Z\right)\right] \\
&= \mathbb{E}\left[W_1\right]\mathbb{E}\left[\exp\left(\mathbf{i}s'Z\right)\right] - \mathbb{E}\left[W_1\right]\mathbb{E}\left[\exp\left(\mathbf{i}s'Z\right)\right] \\
&= 0,
\end{aligned}$$

where the second equality follows from the law of iterated expectations. Then by equation (2.4) in Su and Zheng (2017), we have

$$\begin{aligned}
\text{MDD}_o\left(W_2 - W_1|Z\right)^2 &= \int_{\mathbb{R}^{d_Z}} \left[\text{Cov}\left(W_2 - W_1, \exp\left(\mathbf{i}s'Z\right)\right)\right]^2 \cdot q(s)\, ds \\
&= \int_{\mathbb{R}^{d_Z}} \left[\text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right) - \text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right)\right]^2 \cdot q(s)\, ds \\
&= \int_{\mathbb{R}^{d_Z}} \left[\text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)\right]^2 \cdot q(s)\, ds \\
&= \text{MDD}_o\left(W_2|Z\right)^2, \qquad \text{(B.1)}
\end{aligned}$$

where $\mathbf{i} \equiv \sqrt{-1}$, $q(s) \equiv 1/\left[c\,|s|^{(1+d_Z)}\right]$, $c \equiv \pi^{(1+d_Z)/2}/\Gamma\left(\frac{1+d_Z}{2}\right)$, and $\Gamma(\cdot)$ is the complete gamma function: $\Gamma(z) \equiv \int_0^\infty t^{(z-1)} \exp(-t)\, dt$.

To proceed, note that for a generic real-valued random variable W,

$$\mathrm{MDD}\left(W|Z\right)^2 = \mathrm{MDD}_o\left(W|Z\right)^2 + [\mathbb{E}\left(W\right)]^2 \mathbb{E}\left(\left|Z - Z^{\dagger}\right|\right) \tag{B.2}$$

which follows directly from the definitions of MDD_o and MDD. Also note that $\mathrm{MDD}\left(W_1|Z\right)^2 = 0$ if and only if $\mathbb{E}\left(W_1|Z\right) = 0$. It follows that $\mathrm{MDD}\left(W_1|Z\right)^2 = 0$ if and only if

$$\mathrm{MDD}_o\left(W_1|Z\right)^2 = 0 \text{ and } \mathbb{E}\left(W_1\right) = 0. \tag{B.3}$$

This, in conjunction with the given condition that $\mathrm{MDD}\left(W_1|Z\right)^2 = 0$, implies that

$$\begin{aligned} \mathrm{MDD}\left(W_2 - W_1|Z\right)^2 &= \mathrm{MDD}_o\left(W_2 - W_1|Z\right)^2 + [\mathbb{E}\left(W_2 - W_1\right)]^2 \mathbb{E}\left(\left|Z - Z^{\dagger}\right|\right) \\ &= \mathrm{MDD}_o\left(W_2 - W_1|Z\right)^2 + [\mathbb{E}\left(W_2\right)]^2 \mathbb{E}\left(\left|Z - Z^{\dagger}\right|\right) \\ &= \mathrm{MDD}_o\left(W_2|Z\right)^2 + [\mathbb{E}\left(W_2\right)]^2 \mathbb{E}\left(\left|Z - Z^{\dagger}\right|\right) \\ &= \mathrm{MDD}\left(W_2|Z\right)^2, \end{aligned}$$

where the first and last equalities follow from (B.2), the second equality holds by (B.3), and the third equality holds by (B.1). This proves the first claim.

Now, we prove the second claim. By Su and Zheng (2017),

$$\begin{aligned} \mathrm{MDD}\left(W_2 - W_1|Z\right)^2 &= -\mathbb{E}\left[\left(W_2 - W_1\right)\left(W_2^{\dagger} - W_1^{\dagger}\right)\left|Z - Z^{\dagger}\right|\right] \\ &\quad + 2\mathbb{E}\left[\left(W_2 - W_1\right)\left|Z - Z^{\dagger}\right|\right]\mathbb{E}\left[W_2^{\dagger} - W_1^{\dagger}\right] \\ &= -\mathbb{E}\left[\left(W_2 - W_1\right)\left(W_2^{\dagger} - W_1^{\dagger}\right)\left|Z - Z^{\dagger}\right|\right] \\ &\quad + 2\mathbb{E}\left[W_2\left|Z - Z^{\dagger}\right|\right]\mathbb{E}\left[W_2^{\dagger}\right] - 2\mathbb{E}\left[W_1\left|Z - Z^{\dagger}\right|\right]\mathbb{E}\left[W_2^{\dagger} - W_1^{\dagger}\right] \\ &\quad - 2\mathbb{E}\left[W_2\left|Z - Z^{\dagger}\right|\right]\mathbb{E}\left[W_1^{\dagger}\right]. \end{aligned}$$

Noting that $\mathbb{E}\left(W_1|Z\right) = 0$, we have $\mathbb{E}\left(W_1\right) = \mathbb{E}(W_1^{\dagger}) = 0$ and $\mathbb{E}\left[W_1\left|Z - Z^{\dagger}\right|\right] = 0$ by the law of iterated expectations and the independence between (W_1, Z) and $Z^{\dagger}$. Then we have

$$\mathrm{MDD}\left(W_2 - W_1|Z\right)^2 = -\mathbb{E}\left[\left(W_2 - W_1\right)\left(W_2^{\dagger} - W_1^{\dagger}\right)\left|Z - Z^{\dagger}\right|\right] + 2\mathbb{E}\left[W_2\left|Z - Z^{\dagger}\right|\right]\mathbb{E}\left[W_2^{\dagger}\right]. \tag{B.4}$$

And we also know that

$$\mathrm{MDD}\left(W_2|Z\right)^2 = -\mathbb{E}\left[W_2 W_2^{\dagger}\left|Z - Z^{\dagger}\right|\right] + 2\mathbb{E}\left[W_2\left|Z - Z^{\dagger}\right|\right]\mathbb{E}\left[W_2^{\dagger}\right]. \tag{B.5}$$

(B.4)–(B.5), in conjunction with the fact that $\mathrm{MDD}\left(W_2 - W_1|Z\right)^2 = \mathrm{MDD}\left(W_2|Z\right)^2$, implies that

$$\mathbb{E}\left[\left(W_2 - W_1\right)\left(W_2^{\dagger} - W_1^{\dagger}\right)\left|Z - Z^{\dagger}\right|\right] = \mathbb{E}\left[W_2 W_2^{\dagger}\left|Z - Z^{\dagger}\right|\right]. \blacksquare$$

Proof of Lemma A.2. Note that for any real-valued random variable W, it holds that

$$\text{MDD}\,(W|Z)^2 = \text{MDD}_o\,(W|Z)^2 + [\mathbb{E}\,(W)]^2\mathbb{E}\left(\left|Z - Z^\dagger\right|\right). \tag{B.6}$$

We prove the first and second inequalities in Parts I and II, respectively.

Part I. We organize Part I into three subparts. In Part I (i), we show the existence of a finite constant b_1 s.t. for any pair $\{W_1, W_2\} \subseteq \mathcal{W}$, it holds that

$$\left|\text{MDD}_o\,(W_1|Z)^2 - \text{MDD}_o\,(W_2|Z)^2\right| \leq b_1 \cdot \left[\text{MDD}_o\,(W_1 - W_2|Z)^2\right]^{1/2}.$$

In Part I (ii), we show the existence of a finite constant b_2 s.t. for any pair $\{W_1, W_2\} \subseteq \mathcal{W}$, it holds that

$$\left|\mathbb{E}\,(W_1)^2\,\mathbb{E}\left(\left|Z - Z^\dagger\right|\right) - \mathbb{E}\,(W_2)^2\,\mathbb{E}\left(\left|Z - Z^\dagger\right|\right)\right| \leq b_2 \cdot \left[\mathbb{E}\,(W_1 - W_2)^2\,\mathbb{E}\left(\left|Z - Z^\dagger\right|\right)\right]^{1/2}.$$

And in Part I (iii), we combine the results from Part I (i) and Part I (ii) via B.6 to prove the first inequality.

Part I (i). By the definition of $\mathcal{W}$,

$$\sup_{W\in\mathcal{W}} \text{Var}\,(W) = \sup_{W\in\mathcal{W}} \left[\mathbb{E}\left(W^2\right) - \mathbb{E}\,(W)^2\right] \leq \sup_{W\in\mathcal{W}} \mathbb{E}\left(W^2\right) \equiv b_3 < \infty. \tag{B.7}$$

Denote by $\varphi_Z\,(s) \equiv \mathbb{E}\,[\exp\,(\mathbf{i}s'Z)]$, the characteristic function of Z. It holds that

$$\begin{aligned} \left|\text{Var}\left(\exp\left(\mathbf{i}s'Z\right)\right)\right| &= \left|\mathbb{E}\left[\exp\left(\mathbf{i}s'Z\right)^2\right] - \mathbb{E}\left[\exp\left(\mathbf{i}s'Z\right)\right]^2\right| \\ &= \left|\varphi_Z\,(2s) - [\varphi_Z\,(s)]^2\right| \leq |\varphi_Z\,(2s)| + |\varphi_Z\,(s)|^2 \leq 2, \end{aligned} \tag{B.8}$$

where the last inequality follows from the fact that $|\varphi_Z\,(\cdot)| \leq 1$. By equation (2.4) in Su and Zheng (2017), we have that for any pair $\{W_1, W_2\} \subseteq \mathcal{W}$,

$$\begin{aligned} & \left|\text{MDD}_o\,(W_1|Z)^2 - \text{MDD}_o\,(W_2|Z)^2\right| \\ = & \left|\int_{\mathbb{R}^{d_Z}} \left[\text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right)^2 - \text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)^2\right] q\,(s)\,ds\right| \\ \leq & \int_{\mathbb{R}^{d_Z}} \left|\text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right)^2 - \text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)^2\right| q\,(s)\,ds \\ \leq & \int_{\mathbb{R}^{d_Z}} \left|\text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right) - \text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)\right| \\ & \quad \times \left[\left|\text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right)\right| + \left|\text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)\right|\right] q\,(s)\,ds \\ \leq & \int_{\mathbb{R}^{d_Z}} \left|\text{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right) - \text{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)\right| \\ & \quad \times \left[\left|\text{Var}\,(W_1)^{1/2}\,\text{Var}\left(\exp\left(\mathbf{i}s'Z\right)\right)^{1/2}\right| + \left|\text{Var}\,(W_2)^{1/2}\,\text{Var}\left(\exp\left(\mathbf{i}s'Z\right)\right)^{1/2}\right|\right] q\,(s)\,ds \end{aligned}$$

$$
\begin{aligned}
&\leq 2\sqrt{2b_3}\int_{\mathbb{R}^{d_Z}}\left|\mathrm{Cov}\left(W_1,\exp\left(\mathbf{i}s'Z\right)\right)-\mathrm{Cov}\left(W_2,\exp\left(\mathbf{i}s'Z\right)\right)\right|q\left(s\right)ds\\
&= 2\sqrt{2b_3}\int_{\mathbb{R}^{d_Z}}\left|\mathrm{Cov}\left(W_1-W_2,\exp\left(\mathbf{i}s'Z\right)\right)\right|q\left(s\right)ds\\
&\leq 2\sqrt{2b_3}\left[\int_{\mathbb{R}^{d_Z}}\left|\mathrm{Cov}\left(W_1-W_2,\exp\left(\mathbf{i}s'Z\right)\right)\right|^2 q\left(s\right)ds\right]^{1/2}\\
&= 2\sqrt{2b_3}\left[\mathrm{MDD}_o\left(W_1-W_2|Z\right)^2\right]^{1/2}
\end{aligned}
$$

where $\mathbf{i}\equiv\sqrt{-1}$, $q\left(s\right)$ is as defined in the proof of Lemma A.1, the fourth inequality follows from (B.7)-(B.8), and the last inequality follows from the Hölder inequality. Consequently, we have that for any pair $\{W_1,W_2\}\subseteq\mathcal{W}$,

$$
\left|\mathrm{MDD}_o\left(W_1|Z\right)^2-\mathrm{MDD}_o\left(W_2|Z\right)^2\right|\leq b_1\left[\mathrm{MDD}_o\left(W_1-W_2|Z\right)^2\right]^{1/2} \tag{B.9}
$$

where $b_1\equiv 2\sqrt{2b_3}$.

Part I (ii). By Jensen inequality and (B.7), we have

$$
\sup_{W\in\mathcal{W}}\mathbb{E}\left|W\right|\leq\sup_{W\in\mathcal{W}}\left[\mathbb{E}\left(W^2\right)\right]^{1/2} = \sqrt{b_3}. \tag{B.10}
$$

For any pair $\{W_1,W_2\}\subseteq\mathcal{W}$,

$$
\begin{aligned}
&\left|\left[\mathbb{E}\left(W_1\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|-\left[\mathbb{E}\left(W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right|\\
=\ &\left|\mathbb{E}\left(W_1+W_2\right)\mathbb{E}\left|Z-Z^{\dagger}\right|^{1/2}\right|\left|\mathbb{E}\left(W_1-W_2\right)\left(\mathbb{E}\left|Z-Z^{\dagger}\right|\right)^{1/2}\right|\\
\leq\ &\left[\mathbb{E}\left|W_1\right|+\mathbb{E}\left|W_2\right|\right]\left(\mathbb{E}\left|Z-Z^{\dagger}\right|\right)^{1/2}\left[\left[\mathbb{E}\left(W_1-W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right]^{1/2}\\
\leq\ &2\sqrt{b_3}\left(\mathbb{E}\left|Z-Z^{\dagger}\right|\right)^{1/2}\left[\left[\mathbb{E}\left(W_1-W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right]^{1/2}\\
=\ &b_2\left[\left[\mathbb{E}\left(W_1-W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right]^{1/2}
\end{aligned} \tag{B.11}
$$

where the last inequality follows from (B.9), and $b_2\equiv 2\sqrt{b_3}\cdot\mathbb{E}\left(\left|Z-Z^{\dagger}\right|\right)^{1/2}<\infty$ by the condition that $\mathbb{E}\left|Z-Z^{\dagger}\right|<\infty$ and the fact that $b_3<\infty$.

Part I (iii). For any pair $\{W_1,W_2\}\subseteq\mathcal{W}$, it holds that

$$
\begin{aligned}
&\left|\mathrm{MDD}\left(W_1|Z\right)^2-\mathrm{MDD}\left(W_2|Z\right)^2\right|\\
\leq\ &\left|\mathrm{MDD}_o\left(W_1|Z\right)^2-\mathrm{MDD}_o\left(W_2|Z\right)^2\right|+\left|\left\{\left[\mathbb{E}\left(W_1\right)\right]^2-\left[\mathbb{E}\left(W_2\right)\right]^2\right\}\mathbb{E}\left(\left|Z-Z^{\dagger}\right|\right)\right|\\
\leq\ &b_1\left[\mathrm{MDD}_o\left(W_1-W_2|Z\right)^2\right]^{1/2}+b_2\left[\left[\mathbb{E}\left(W_1-W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right]^{1/2}\\
\leq\ &\max\left\{b_1,b_2\right\}\left\{\left[\mathrm{MDD}_o\left(W_1-W_2|Z\right)^2\right]^{1/2}+\left[\left[\mathbb{E}\left(W_1-W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right]^{1/2}\right\}\\
\leq\ &\sqrt{2}\max\left\{b_1,b_2\right\}\left\{\mathrm{MDD}_o\left(W_1-W_2|Z\right)^2+\left[\mathbb{E}\left(W_1-W_2\right)\right]^2\mathbb{E}\left|Z-Z^{\dagger}\right|\right\}^{1/2}
\end{aligned}
$$

$$= b\left[\mathrm{MDD}\left(W_1 - W_2|Z\right)^2\right]^{1/2},$$

where $b \equiv \sqrt{2}\max\{b_1, b_2\} < \infty$, the first inequality follows from the triangle inequality, and the second inequality holds by (B.9) and (B.11).

Part II. It is easy to see that

$$\mathbb{E}\left(W_1 - W_2\right)^2 \mathbb{E}\left|Z - Z^{\dagger}\right| \leq 2\left[\mathbb{E}\left(W_1\right)^2 \mathbb{E}\left|Z - Z^{\dagger}\right| + \mathbb{E}\left(W_2\right)^2 \mathbb{E}\left|Z - Z^{\dagger}\right|\right]. \quad \text{(B.12)}$$

Next, note that

$$\begin{aligned}
&\mathrm{MDD}_o\left(W_1 - W_2|Z\right)^2 \\
= & \int_{\mathbb{R}^{d_Z}} \mathrm{Cov}\left(W_1 - W_2, \exp\left(\mathbf{i}s'Z\right)\right)^2 q(s)\, ds \\
= & \int_{\mathbb{R}^{d_Z}} \left[\mathrm{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right) - \mathrm{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)\right]^2 q(s)\, ds \\
\leq & \; 2\int_{\mathbb{R}^{d_Z}} \mathrm{Cov}\left(W_1, \exp\left(\mathbf{i}s'Z\right)\right)^2 q(s)\, ds + 2\int_{\mathbb{R}^{d_Z}} \mathrm{Cov}\left(W_2, \exp\left(\mathbf{i}s'Z\right)\right)^2 q(s)\, ds \\
= & \; 2\left[\mathrm{MDD}_o\left(W_1|Z\right)^2 + \mathrm{MDD}_o\left(W_2|Z\right)^2\right]. \qquad \text{(B.13)}
\end{aligned}$$

Combining (B.6), (B.12) and (B.13) yields $\mathrm{MDD}\left(W_1 - W_2|Z\right)^2 \leq 2\left[\mathrm{MDD}\left(W_1|Z\right)^2 + \mathrm{MDD}\left(W_2|Z\right)^2\right]$. It follows that

$$\begin{aligned}
\left[\mathrm{MDD}\left(W_1 - W_2|Z\right)^2\right]^{1/2} & \leq \sqrt{2}\left[\mathrm{MDD}\left(W_1|Z\right)^2 + \mathrm{MDD}\left(W_2|Z\right)^2\right]^{1/2} \\
& \leq 2\left\{\left[\mathrm{MDD}\left(W_1|Z\right)^2\right]^{1/2} + \left[\mathrm{MDD}\left(W_2|Z\right)^2\right]^{1/2}\right\}
\end{aligned}$$

where the second inequality follows from the fact that $\left(a^2 + b^2\right)^{1/2} \leq \sqrt{2}\left(a + b\right)$ for any $a, b \geq 0$. This completes the proof. ■

Proof of Lemma A.3. To prove the first claim, note that $\left(\mathbb{R}^{(T-R)R} \times \mathcal{W}^s\left(\mathcal{X}\right), \|\cdot\|_c\right)$ forms a metric space, in which compactness is equivalent to sequential compactness. So it is sufficient to show that $(\Theta, \|\cdot\|_c)$ is sequentially compact.

Recall that $\Theta = \mathbf{\Phi} \times \mathcal{G}$, where $\mathbf{\Phi}$ is compact by assumption. By Lemma A.2 in Santos (2012), $(\mathcal{G}, \|\cdot\|_c)$ is also compact despite the difference in notations. Specifically, $\lfloor \frac{d}{2} \rfloor$ is equivalent to Santos's (2012) m, and $d - \lfloor \frac{d}{2} \rfloor$ is equivalent to his m_0. In addition, the condition $d \geq d_x + 2$ required by Assumption 2.1(ii) guarantees 'min $\{m_0, m\} > \frac{d_x}{2}$', which is required in Santos (2012). Consequently, both $\mathbf{\Phi}$ and $\mathcal{G}$ are sequentially compact. Then for any given sequence $\{\theta_N = \left(\phi_N', g_N\right)'\}$ in Θ, there is subsequence $\{\theta_{M_N} = \left(\phi_{M_N}', g_{M_N}\right)'\}$ s.t. $\phi_{M_N} \to \phi$ under $|\cdot|$ for some $\phi \in \mathbf{\Phi}$ as $M_N \to \infty$, and there is also a subsubsequence $\{\theta_{L_{M_N}} = (\phi_{L_{M_N}}', g_{L_{M_N}})'\}$ in $\{\theta_{M_N} = \left(\phi_{M_N}', g_{M_N}\right)'\}$ s.t. $g_{L_{M_N}} \to g$

under $\|\cdot\|_c$ for some $g \in \mathcal{G}$ as $L_{M_N} \to \infty$. In short, for any given sequence $\{\theta_N = (\phi_N', g_N)'\}$ in Θ, we are able to find a subsequence $\{\theta_{L_{M_N}} = (\phi_{L_{M_N}}', g_{L_{M_N}})'\}$ s.t. $\theta_{L_{M_N}} \to (\phi', g)' \in \Theta$ under $\|\cdot\|_c$ as $L_{M_N} \to \infty$. This shows that $(\Theta, \|\cdot\|_c)$ is sequentially compact.

To prove the second claim, note that $(\mathcal{W}^s(\mathcal{X}), \|\cdot\|_c)$ is a metric space. The compactness of its subset $(\mathcal{G}, \|\cdot\|_c)$ as proven above implies total boundedness of $\mathcal{G}$ under $\|\cdot\|_c$, which in turn implies boundedness of $\mathcal{G}$ under $\|\cdot\|_c$. So there exists a constant B_c s.t. $\sup_{g\in\mathcal{G}} \|g\|_c \leq B_c$. As a result, for any given vector of non-negative integer λ with $\langle\lambda\rangle \leq \frac{d}{2}$, it hods that

$$
\begin{aligned}
\sup_{x\in\mathcal{X}} \left|D^\lambda g(x)\right| &\leq \sup_{x\in\mathcal{X}} \left|D^\lambda g(x)\right| (1+x'x)^{\zeta/2} \leq \max_{\langle\lambda\rangle\leq\frac{d}{2}} \left[\sup_{x\in\mathcal{X}} \left|D^\lambda g(x)\right| (1+x'x)^{\zeta/2}\right] \\
&= \|g\|_c \leq \sup_{g\in\mathcal{G}} \|g\|_c \leq B_c,
\end{aligned}
$$

where the first inequality follows from the fact that $(1+x'x)^{\zeta/2} \geq 1$ under the requirement that $\zeta > \left(\frac{d_x}{2}\cdot\lfloor\frac{d}{2}\rfloor\right) / \left(\lfloor\frac{d}{2}\rfloor - \frac{d_x}{2}\right) > 0$ as specified in Assumption 2.1(ii). ∎

Proof of Lemma A.4. Here we follow the same notations for various bounds as we have adopted in the previous proofs. Specifically, the compactness of $\mathbf{\Phi}$ according to Assumption 2.1(i) implies that $B_{\mathbf{\Phi}} \equiv \sup_{\phi\in\mathbf{\Phi}} |\phi| < \infty$. By Lemma A.3, for all $g \in \mathcal{G}$, it holds that $\sup_{x\in\mathcal{X}} |g(x)| \leq B_c < \infty$. Then for any $\theta \in \Theta$ and $s = 1, ..., T-R$, we have

$$
\begin{aligned}
|m_s(Y, X, \theta)| &= \left|[y_s - g(x_s)] + \sum_{r=1}^{R} \phi_{s,r}[y_{T-R+r} - g(x_{T-R+r})]\right| \\
&\leq [|y_s| + |g(x_s)|] + \sum_{r=1}^{R} \left|\phi_{s,r}\right| [|y_{T-R+r}| + |g(x_{T-R+r})|] \\
&\leq |Y| + B_c + B_{\mathbf{\Phi}} R[|Y| + B_c] = (B_{\mathbf{\Phi}} R + 1)[|Y| + B_c]
\end{aligned}
$$

where $Y = (y_1, ..., y_T)'$, $X = (x_1, ..., x_T)'$, $\phi_{s,r}$ denote the rth element in ϕ_s, and the first inequality holds by the triangle inequality. It follows that for any $\theta \in \Theta$ and $s = 1, ..., T-R$,

$$
\mathbb{E}\left\{[m_s(Y, X, \theta)]^2\right\} \leq 2(B_{\mathbf{\Phi}} R + 1)^2 \left[\mathbb{E}(|Y|^2) + B_c^2\right] < \infty
$$

where the last inequality follows from the fact that $\mathbb{E}(|Y|^2) < \infty$ under Assumption 3.1(ii). In addition, by the triangle inequality, Jensen inequality, and Assumption 3.1(ii), $\mathbb{E}\left|\underline{z}_s - \underline{z}_s^\dagger\right| \leq \mathbb{E}\left|Z - Z^\dagger\right| \leq 2\left[\mathbb{E}(|Z|^2)\right]^{1/2} < \infty$. Therefore, the result in Lemma A.2 is applicable here.

By Lemma A.2, we have that for any $\theta_1 \in \Theta$ and $\theta_2 \in \Theta$,

$$
\begin{aligned}
\left|\mathrm{MDD}[m_s(Y, X, \theta_1) \mid \underline{z}_s]^2 - \mathrm{MDD}[m_s(Y, X, \theta_2) \mid \underline{z}_s]^2\right| &\leq b_s \left\{\mathrm{MDD}[m_s(Y, X, \theta_1) - m_s(Y, X, \theta_2) \mid \underline{z}_s]^2\right\}^{1/2} \\
&\leq b_s c \|\theta_1 - \theta_2\|_{L^2}
\end{aligned}
$$

for some finite constants b_s and c, where the first and second inequalities hold by Lemmas A.2 and 3.1, respectively. As a result,

$$\begin{aligned} |\mathcal{Q}(\theta_1) - \mathcal{Q}(\theta_2)| &= \left| \sum_{s=1}^{T-R} \mathrm{MDD}\left[m_s(Y, X, \theta_1) \mid \underline{z}_s\right]^2 - \sum_{s=1}^{T-R} \mathrm{MDD}\left[m_s(Y, X, \theta_2) \mid \underline{z}_s\right]^2 \right| \\ &\le \sum_{s=1}^{T-R} \left| \mathrm{MDD}\left[m_s(Y, X, \theta_1) \mid \underline{z}_s\right]^2 - \mathrm{MDD}\left[m_s(Y, X, \theta_2) \mid \underline{z}_s\right]^2 \right| \\ &\le \left[\sum_{s=1}^{T-R} b_s \right] c \left\| \theta_1 - \theta_2 \right\|_{L^2}. \end{aligned}$$

This completes the proof of the lemma. ■

Proof of Lemma A.5. We refer to Conditions (i) – (iv) listed in the statement of Lemma A.5 as C(i) – C(iv), respectively. By C(iii), $\forall\, \varepsilon > 0$, $\exists$ a constant $L_\varepsilon < \infty$ and a positive integer N_ε s.t.

$$\Pr\left(\sup_{\theta \in \Theta_N} |Q_N(\theta) - Q(\theta)| < L_\varepsilon b_N \right) \ge 1 - \varepsilon \tag{B.14}$$

for all $N \ge N_\varepsilon$. For any given $\theta^0 \in \Theta_I$, it follows from C(ii) that $\exists$ a sequence $\left\{\theta_N^0\right\}$ with $\theta_N^0 \in \Theta_N$ s.t. $\mathrm{d}\left(\theta_N^0, \theta^0\right) \le \sigma_N$.

Let $A_\varepsilon \equiv \max\left\{2\sqrt{L_\varepsilon/a_1}, \sqrt{2a_2/a_1}\right\} < \infty$ and $\rho_{N,\varepsilon} \equiv A_\varepsilon \max\{\sigma_N, b_N^{1/2}\}$. Let $\Theta_I^{\rho_{N,\varepsilon}}$ be the open $\rho_{N,\varepsilon}$ enlargement of Θ_I under $\mathrm{d}(\cdot,\cdot)$. By C(i) (compactness of Θ) and C(iv), it holds that

$$\Delta_{N,\varepsilon} \equiv \inf_{\theta \in \left(\Theta_I^{\rho_{N,\varepsilon}}\right)^c \cap \Theta} Q(\theta) \ge a_1 \rho_{N,\varepsilon}^2,$$

where A^c denotes the complement of set A. Note that

$$\begin{aligned} Q\left(\hat{\theta}_N\right) - Q\left(\theta_N^0\right) &= \left[Q\left(\hat{\theta}_N\right) - Q_N\left(\hat{\theta}_N\right)\right] + \left[Q_N\left(\hat{\theta}_N\right) - Q_N\left(\theta_N^0\right)\right] + \left[Q_N\left(\theta_N^0\right) - Q\left(\theta_N^0\right)\right] \\ &\le \left[Q\left(\hat{\theta}_N\right) - Q_N\left(\hat{\theta}_N\right)\right] + \left[Q_N\left(\theta_N^0\right) - Q\left(\theta_N^0\right)\right] \\ &\le \left|Q\left(\hat{\theta}_N\right) - Q_N\left(\hat{\theta}_N\right)\right| + \left|Q_N\left(\theta_N^0\right) - Q\left(\theta_N^0\right)\right|, \end{aligned} \tag{B.15}$$

where the first inequality holds because $Q_N\left(\hat{\theta}_N\right) - Q_N\left(\theta_N^0\right) \le 0$ by the definition of $\hat{\theta}_N$. Then for $N \ge N_\varepsilon$, we have

$$\begin{aligned} \Pr\left(Q\left(\hat{\theta}_N\right) < Q\left(\theta_N^0\right) + \frac{\Delta_{N,\varepsilon}}{2}\right) &= \Pr\left(Q\left(\hat{\theta}_N\right) - Q\left(\theta_N^0\right) < \frac{\Delta_{N,\varepsilon}}{2}\right) \\ &\ge \Pr\left(Q\left(\hat{\theta}_N\right) - Q\left(\theta_N^0\right) < \frac{a_1 \rho_{N,\varepsilon}^2}{2}\right) \\ &\ge \Pr\left(\left|Q\left(\hat{\theta}_N\right) - Q_N\left(\hat{\theta}_N\right)\right| + \left|Q_N\left(\theta_N^0\right) - Q\left(\theta_N^0\right)\right| < 2L_\varepsilon b_N\right) \\ &\ge 1 - \varepsilon \end{aligned} \tag{B.16}$$

where the second inequality follows from (B.15) and the fact that $a_1\rho_{N,\varepsilon}^2/2 \geq 2L_\varepsilon b_N$, and the last inequality follows from (B.14).

It follows from C(iv) that $Q\left(\theta_N^0\right) \leq a_2 \,\mathrm{d}\left(\theta_N^0, \Theta_I\right)^2 \leq a_2 \,\mathrm{d}\left(\theta_N^0, \theta^0\right)^2 \leq a_2\sigma_N^2 \leq \Delta_{N,\varepsilon}/2$ for N large enough, which, together with (B.16), implies that

$$\Pr\left(Q\left(\hat{\theta}_N\right) < \Delta_{N,\varepsilon}\right) \geq \Pr\left(Q\left(\hat{\theta}_N\right) < Q\left(\theta_N^0\right) + \frac{\Delta_{N,\varepsilon}}{2}\right) \geq 1-\varepsilon \tag{B.17}$$

for N large enough.

Note that $Q(\hat{\theta}_N) < \Delta_{N,\varepsilon}$ if and only if $\hat{\theta}_N \in \Theta_I^{\rho_{N,\varepsilon}}$, or, equivalently, $\mathrm{d}(\hat{\theta}_N, \Theta_I) < \rho_{N,\varepsilon} = A_\varepsilon \max\{\sigma_N, b_N^{1/2}\}$. Therefore, we can rewrite (B.17) as

$$\Pr\left(\mathrm{d}\left(\hat{\theta}_N, \Theta_I\right) < A_\varepsilon \max\{\sigma_N, b_N^{1/2}\}\right) \geq 1-\varepsilon$$

for N large enough. This exactly shows that $\mathrm{d}\left(\hat{\theta}_N, \Theta_I\right) = O_p\left(\max\{\sigma_N, b_N^{1/2}\}\right)$. ■

To prove Lemma A.6, we need the following Lemma.

Lemma B.1 *Consider a generic econometric model $Q(\theta) = 0$, the identified set of which is characterized by $\Theta_I \equiv \{\theta \in \Theta : Q(\theta) = 0 \text{ a.s.}\}$. Suppose the following conditions hold: (i) $Q(\cdot) \geq 0$ and Θ is compact under (pseudo-)norm $\|\cdot\|$; (ii) $\Theta_N \subseteq \Theta$ are closed and s.t. $\exists\, \Pi_N\theta \in \Theta_N$ for each $\theta \in \Theta$ s.t. $\sup_{\theta\in\Theta} \|\Pi_N\theta - \theta\| = o(1)$; (iii) $\sup_{\theta\in\Theta_N} |Q_N(\theta) - Q(\theta)| = o_p(1)$; (iv) $Q(\cdot)$ is continuous w.r.t. $\|\cdot\|$ in Θ. Then for $\hat{\theta}_N \in \operatorname*{argmin}_{\theta\in\Theta_N} Q_N(\theta)$, it holds that $\mathrm{d}_{\|\cdot\|}(\hat{\theta}_N, \Theta_I) = o_p(1)$.*

Proof of Lemma B.1. Lemma B.1 is essentially the same as Lemma A.5 in Santos (2012), except that we do not assume the continuity of Q_N w.r.t. $\|\cdot\|$ in Θ_N. A close inspection on the proof of Lemma A.5 in Santos (2012) shows that the continuity of Q_N does not play a role in the proof. In other words, the proof of Lemma A.5 in Santos (2012) works without assuming the continuity of Q_N, and therefore applies directly to proving Lemma B.1 here. ■

Proof of Lemma A.6. We prove parts (i) and (ii) of the Lemma in turn.

Part I. Proof of part (i).

Let $Q(\theta) \equiv \sum_{s=1}^{T-R} \mathrm{MDD}\left[m_s(Y, X, \theta) | \underline{z}_s\right]^2$ and $Q_N(\theta) \equiv \frac{1}{N} S_N(\theta) = \sum_{s=1}^{T-R} \frac{1}{N} S_{Ns}(\theta)$, as in the proof of Theorem 3.2. Our goal is to show that, over the restricted parameter space $\Theta \cap \Theta_R$ under $\|\cdot\|_{L^2}$, $Q(\cdot)$ and $Q_N(\cdot)$ as specified above satisfy Conditions (i)–(iv) in Lemma B.1.

Due to the non-negativity of MDD, $Q(\cdot) \geq 0$. By Lemma A.3, Θ is compact under $\|\cdot\|_c$ and hence is compact under $\|\cdot\|_{L^2}$, which is weaker than $\|\cdot\|_c$. Since Θ_R is closed due to the continuity of $L(\cdot)$ under Assumption 2.2, $\Theta \cap \Theta_R$ is also compact under $\|\cdot\|_{L^2}$. So Condition (i) in Lemma B.1 is satisfied. Assumption 3.3(i) guarantees Condition (ii) in Lemma B.1 to hold; Condition (iii) in Lemma B.1 holds according to Theorem 3.1. Condition (iv) in Lemma B.1 holds according to Lemma A.4. Consequently, the conclusion in part (i) follows from Lemma B.1.

Part II. Proof of part (ii).

This part of the proof is similar to the proof of Lemma A.3 in Hong (2017) and it goes as follows. For any given $\theta^0 \in \Theta_I$,

$$\begin{aligned}\left\|\hat{\theta}_N - \theta^0\right\|_{L^2} &\leq \left\|\hat{\theta}_N - \Pi_N\theta^0\right\|_{L^2} + \left\|\Pi_N\theta^0 - \theta^0\right\|_{L^2} \\ &\leq \varrho_N \, \mathrm{d}_w\left(\hat{\theta}_N, \Pi_N\theta^0\right) + \delta_{s,N} \\ &\leq 2\varrho_N\left[\mathrm{d}_w\left(\hat{\theta}_N, \theta^0\right) + \mathrm{d}_w\left(\Pi_N\theta^0, \theta^0\right)\right] + \delta_{s,N} \\ &\leq 2\varrho_N\left[\mathrm{d}_w\left(\hat{\theta}_N, \theta^0\right) + \delta_{w,N}\right] + \delta_{s,N},\end{aligned}$$

where the first inequality follows from the triangle inequality for $\|\cdot\|_{L^2}$, the second one holds by Definition 3.2 and Assumption 3.3(i), the third one follows from Lemma 3.1, and the last inequality holds by Assumption 3.3(ii). Taking infimum over $\theta^0 \in \Theta_I \cap \Theta_R$ yields

$$\begin{aligned}\mathrm{d}_{\|\cdot\|_{L^2}}\left(\hat{\theta}_N, \Theta_I \cap \Theta_R\right) &\leq 2\varrho_N\left[\mathrm{d}_w\left(\hat{\theta}_N, \Theta_I \cap \Theta_R\right) + \delta_{w,N}\right] + \delta_{s,N} \\ &= O_p\left(\varrho_N \, \mathrm{d}_w\left(\hat{\theta}_N, \Theta_I \cap \Theta_R\right) + \delta_{s,N}\right)\end{aligned}$$

where the equality holds by the fact that $\delta_{w,N} = o(N^{-1/2})$. ■

C A discussion on Assumption 3.4

Recall that we have already required in Assumption 3.2 that the eigenvalues of $\mathbb{E}\left[p^{k_N}(x_t)\, p^{k_N'}(x_t)\right]$ for $t = 1, ..., T$ are uniformly bounded and uniformly bounded away from zero. Here, we mainly focus on discussing sufficient conditions for ϱ_N to satisfy Assumption 3.4 for point-identification cases. Under point-identification,

$$\varrho_N = \sup_{\theta \in \Theta_{oN} : \theta \neq \Pi_N\theta^0} \frac{\left\|\theta - \Pi_N\theta^0\right\|_{L^2}}{\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right)},$$

with $\Theta_{oN} = \left\{\theta \in \Theta_N : \left\|\theta - \theta^0\right\|_{L^2} \leq \varsigma_N\right\}$ being the $o(1)$ neighborhood of θ^0 under the L^2 norm as $\varsigma_N \downarrow 0$. Since a proper completeness condition is necessary for point-identification, we maintain such a condition for most of the discussion, stated as follows:

> **(Completeness condition)** *For any measurable function* $v(\cdot) : \mathcal{X} \to \mathbb{R}$, $\mathbb{E}\left[v(x_s)\,|\underline{z}_s\right] = 0$ *iff* $v(\cdot) = 0$ *a.s. for* $s = 1, ..., T - R$.

Note that, for any $\theta \in \Theta_N$, we can write $\theta - \Pi_N\theta^0 = \left(\left(\phi - \phi^0\right)', \Delta_N' p^{k_N} \right)'$ for some $\Delta_N \in \mathbb{R}^{k_N}$. By Assumption 3.2, we have

$$\left\| \theta - \Pi_N \theta^0 \right\|_{L^2}^2 \asymp \left| \phi - \phi^0 \right|^2 + \left| \Delta_N \right|^2 . \tag{C.1}$$

C.1 The case of $R = 0$

We consider the special case of $R = 0$ (i.e., no IFEs) and $T = 1$. Generalization to the $R = 0$ and $T > 1$ case is straightforward. Note that, when $R = 0$, $\theta = g(\cdot)$ and $\Theta = \mathcal{G}$. Consequently, $\forall$ $\theta \in \Theta_N$ s.t. $\theta \neq \Pi_N\theta^0$, it holds that $\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right)^2 = \mathrm{MDD}\left[\Delta_N' p^{k_N}(x_{i1}) | z_{i1}\right]^2$ for some $\Delta_N \neq 0$.

For a fixed N, the following three conditions are equivalent: (i) $\mathrm{MDD}\left[\Delta_N' p^{k_N}(x_{i1}) | z_{i1}\right]^2 > 0$ for all $\Delta_N \neq 0$; (ii) $\mathbb{E}\left[\Delta_N' p^{k_N}(x_{i1}) | z_{i1}\right] \neq 0$ for all $\Delta_N \neq 0$; (iii) There is no multicollinearity among the elements of the $p^{k_N}(\cdot)$ vector. Note that the equivalence between (ii) and (iii) follows from the completeness condition. The condition of the eigenvalues of $\mathbb{E}\left[p^{k_N}(x_t)\, p^{k_N'}(x_t)\right]$ being bounded away from zero uniformly over N by Assumption 3.2 guarantees that there is no multicollinearity in $p^{k_N}(\cdot)$ even as $N \to \infty$ and $k_N \to \infty$. Consequently, it can be shown that $\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right) \asymp \left\| \theta - \Pi_N\theta^0 \right\|_{L^2}$ for $\theta \in \Theta_{oN}$. Therefore, when $R = 0$, Assumption 3.4 holds trivially with $\varrho_N = O(1)$ under Assumption 3.2.

C.2 The case of $R \geq 1$

Now we consider the case of $R = 1$ and $T = 2$. Generalization to the $R \geq 1$ and $T = R + 1$ case is straightforward at the cost of more tedious algebra.

In the $o(1)$ neighborhood of θ^0 under the L^2 norm, which is also the $o(1)$ neighborhood of $\Pi_N\theta^0$ under the L^2 norm by Assumption 3.3(i) ($\left|\Pi_N\theta^0 - \theta^0\right| \downarrow 0$ sufficiently fast),

$$\begin{aligned}
\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right)^2 &= \mathrm{MDD}\left[m_1(Y, X, \theta) - m_1\left(Y, X, \Pi_N\theta^0\right) | \underline{z}_s\right]^2 \\
&= \rho_{11}\left(\theta^0, \theta - \Pi_N\theta^0\right) + o(1),
\end{aligned}$$

where

$$\begin{aligned}
& \rho_{11}\left(\theta^0, \theta - \Pi_N\theta^0\right) \\
\equiv\ & -\mathbb{E}\left[\left[\frac{\partial m_1\left(Y, X, \theta^0\right)}{\partial \theta}\left[\theta - \Pi_N\theta^0\right]\right]\left[\frac{\partial m_1\left(Y^\dagger, X^\dagger, \theta^0\right)}{\partial \theta}\left[\theta - \Pi_N\theta^0\right]\right]\left|z_1 - z_1^\dagger\right|\right] \\
& +2\mathbb{E}\left[\left[\frac{\partial m_1\left(Y, X, \theta^0\right)}{\partial \theta}\left[\theta - \Pi_N\theta^0\right]\right]\left|z_1 - z_1^\dagger\right|\right]\mathbb{E}\left[\left[\frac{\partial m_1\left(Y^\dagger, X^\dagger, \theta^0\right)}{\partial \theta}\left[\theta - \Pi_N\theta^0\right]\right]\right].
\end{aligned}$$

Recall $\frac{\partial m_1(Y,X,\theta^0)}{\partial\theta}\left[\theta-\Pi_N\theta^0\right]=\left(\phi_1-\phi_1^0\right)\left[y_2-g^0\left(x_2\right)\right]-\left[g\left(x_1\right)-\Pi_N g^0\left(x_1\right)\right]-\phi_1^0\left[g\left(x_2\right)-\Pi_N g^0\left(x_2\right)\right]$. It then follows that

$$\rho_{11}\left(\theta^0,\theta-\Pi_N\theta^0\right)=\left(\phi_1-\phi_1^0,\Delta_N'\right)M_N\begin{pmatrix}\phi_1-\phi_1^0\\ \Delta_N\end{pmatrix},$$

where

$$M_N\equiv-\mathbb{E}\left[W_N\left(X,X^\dagger,Y,Y^\dagger\right)\left|z_1-z_1^\dagger\right|\right]+2\mathbb{E}\left[W_N\left(X,X^\dagger,Y,Y^\dagger\right)\right]\mathbb{E}\left[\left|z_1-z_1^\dagger\right|\right]$$

with

$$W_N\left(X,X^\dagger,Y,Y^\dagger\right)\equiv\begin{pmatrix}y_2-g^0\left(x_2\right)\\ -p^{k_N}\left(x_1\right)-\phi_1^0p^{k_N}\left(x_2\right)\end{pmatrix}\begin{pmatrix}y_2^\dagger-g^0\left(x_2^\dagger\right)\\ -p^{k_N}\left(x_1^\dagger\right)-\phi_1^0p^{k_N}\left(x_2^\dagger\right)\end{pmatrix}'.$$

By the definition of the martingale difference divergence matrix (MDDM) in Lee and Shao (2018) (see their Definition 1 and Lemma 1),[1] we can rewrite M_N above as

$$M_N=\text{MDDM}_o\left[\begin{pmatrix}y_2-g^0\left(x_2\right)\\ -p^{k_N}\left(x_1\right)-\phi_1^0p^{k_N}\left(x_2\right)\end{pmatrix}|z_1\right]+M_{N1}$$

with

$$M_{N1}\equiv\mathbb{E}\left[\begin{pmatrix}y_2-g^0\left(x_2\right)\\ -p^{k_N}\left(x_1\right)-\phi_1^0p^{k_N}\left(x_2\right)\end{pmatrix}\right]\mathbb{E}\left[\begin{pmatrix}y_2-g^0\left(x_2\right)\\ -p^{k_N}\left(x_1\right)-\phi_1^0p^{k_N}\left(x_2\right)\end{pmatrix}\right]'\mathbb{E}\left[\left|z_1-z_1^\dagger\right|\right].$$

By Lemma 1 and Theorem 1 in Lee and Shao (2018), $\text{MDDM}_o\left[\begin{pmatrix}y_2-g^0\left(x_2\right)\\ -p^{k_N}\left(x_1\right)-\phi_1^0p^{k_N}\left(x_2\right)\end{pmatrix}|z_1\right]$ is positive semi-definite (p.s.d.). This, in conjunction with the p.s.d. of M_{N1}, implies that M_N is also p.s.d.

Note that the MDDM_o in Lee and Shao (2018) is defined to examine conditional mean independence. To examine the case of conditional mean zero, we redefine

$$\begin{aligned}\text{MDDM}\left(V|W\right)&\equiv-\mathbb{E}\left(VV^{\dagger\prime}\left|W-W^\dagger\right|\right)+2\mathbb{E}\left(V\right)\mathbb{E}\left(V^\dagger\right)'\mathbb{E}\left|W-W^\dagger\right|\\&=\text{MDDM}_o\left(V|W\right)+\mathbb{E}\left(V\right)\mathbb{E}\left(V^\dagger\right)'\mathbb{E}\left|W-W^\dagger\right|.\end{aligned}$$

Then it is straightforward to conclude, based on Theorem 1 in Lee and Shao (2018), that $\forall\ V\in\mathbb{R}^p$ and $W\in\mathbb{R}^q$ s.t. $\mathbb{E}\left(|V|^2+|W|^2\right)<\infty$, $\exists\ p-s$ linearly independent combinations of V s.t. they

[1] Lee and Shao (2018) extend the $\text{MDD}_o\left(V|W\right)$ concept of Shao and Zhang (2014) for a scalar variable V to $\text{MDDM}_o\left(V|W\right)$ to a vector-valued V. Specifically, for variables V and W, both of which can be vector-valued, s.t. $\mathbb{E}\left(|V|^2+|W|^2\right)<\infty$, Lee and Shao (2018) specify $\text{MDDM}_o\left(V|W\right)=-\mathbb{E}\left[\left(V-\mathbb{E}\left(V\right)\right)\left(V^\dagger-\mathbb{E}\left(V^\dagger\right)\right)'\left|W-W^\dagger\right|\right]$.

are conditionally mean zero w.r.t. W, iff $\text{rank}\left(\text{MDDM}\left(V|W\right)\right) = s$. Consequently, M_N is strictly positive definite if and only if:

$$\text{no element of } \mathbb{E}\left[\begin{pmatrix} y_2 - g^0\left(x_2\right) \\ -p^{k_N}\left(x_1\right) - \phi_1^0 p^{k_N}\left(x_2\right) \end{pmatrix} |z_1\right] \text{ equals zero.}$$

If, in addition, we require that the smallest eigenvalue of M_N be bounded away from zero, a condition similar to Assumption 3.2, then

$$\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right)^2 = \rho_{11}\left(\theta^0, \theta - \Pi_N\theta^0\right) + o_p\left(1\right) \succeq \left|\phi - \phi^0\right|^2 + \left|\Delta_N\right|^2, \tag{C.2}$$

which, in conjunction with (C.1), implies that $\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right)^2 \succeq \left\|\theta - \Pi_N\theta^0\right\|_{L^2}^2$. This, together with Lemma 3.1, implies that $\mathrm{d}_w\left(\theta, \Pi_N\theta^0\right)^2 \asymp \left\|\theta - \Pi_N\theta^0\right\|_{L^2}^2$ in an $o\left(1\right)$ neighborhood of θ^0 under the L^2 norm. Consequently, Assumption 3.4 is satisfied with $\varrho_N = O\left(1\right)$.

REFERENCES

Hong, S., 2017. Inference in semiparametric conditional moment models with partial identification. Journal of Econometrics 196, 156–179.

Lee, C. E., Shao, X., 2018. Martingale difference divergence matrix and its application to dimension reduction for stationary multivariate time series. Journal of the American Statistical Association 113, 216-229.

Santos, A., 2012. Inference in nonparametric instrumental variables with partial identification. Econometrica 80, 213–275.

Shao, X., Zhang, J., 2014. Martingale divergence correlation and its use in high dimensional variable screening. Journal of the American Statistical Association 109, 1302-1318.

Su, L., Zheng, X., 2017. A Martingale-difference-divergence-based test for specification. Economics Letters 156, 162–167.