Supplement to "On uniform asymptotic risk of averaging GMM estimators"

(*Quantitative Economics*, Vol. 10, No. 3, July 2019, 931–979)

Xu Cheng
Department of Economics, University of Pennsylvania

Zhipeng Liao
Department of Economics, UCLA

Ruoyao Shi
Department of Economics, UC Riverside

In this Supplemental Appendix, we present supporting materials for Cheng, Liao, and Shi (2019) (cited as CLS hereafter in this Appendix):

- Section D provides primitive conditions for Assumptions 3.1, 3.2, and 3.3 and the proof of Lemma 3.1 of CLS.
- Section E provides the proof of (4.3) in Section 4 and the proof of some Lemmas in Appendix B.1 of CLS. The proof of Lemma A.1 in Appendix A of CLS is also included in this section.
- Section F studies the bounds of asymptotic risk difference of the pre-test GMM estimator.
- Section G contains simulation results under the truncated risk for the simulation designs in Section 6 of CLS.
- Section H includes extra simulation studies.

Appendix D: Primitive conditions for Assumptions 3.1, 3.2, and 3.3 and proof of Lemma 3.1 of CLS

In this section, we provide primitive conditions for Assumptions 3.1, 3.2, and 3.3 in the linear IV model presented in Example 3.1 of CLS.

We first provide a set of sufficient conditions without imposing the normal distribution assumption on $(X', Z_1', V', U)'$ in Lemma D.1. Then we impose the normality assumptions and show that these conditions can be simplified to those in Lemma 3.1 of CLS under normality.

Xu Cheng: xucheng@upenn.edu
Zhipeng Liao: zhipeng.liao@econ.ucla.edu
Ruoyao Shi: ruoyao.shi@ucr.edu

Available at http://qeconomics.org. https://doi.org/10.3982/QE711

For ease of notation, we define $\Gamma_{z_1vu^2} \equiv \mathbb{E}_{F^*}[Z_1V'U^2]$, $\Omega_{z_1z_1u^2} \equiv \mathbb{E}_{F^*}[Z_1Z_1'U^2]$ and $\Omega_{vvu^2} \equiv \mathbb{E}_{F^*}[VV'U^2]$. The Jacobian matrices are

$$G_{1,F} = -\mathbb{E}_F\big[Z_1X'\big] \quad \text{and} \quad G_{2,F} = \begin{pmatrix} -\mathbb{E}_F\big[Z_1X'\big] \\ -\mathbb{E}_F\big[Z^*X'\big] \end{pmatrix}. \tag{D.1}$$

Let $Z_2 = (Z_1', Z^{*\prime})'$. The variance–covariance matrix of the moment conditions is

$$\Omega_{2,F} = \mathbb{E}_F\big[Z_2Z_2'(Y - X'\theta_0)^2\big] - \mathbb{E}_F\big[(Y - X'\theta_0)Z_2\big]\mathbb{E}_F\big[(Y - X'\theta_0)Z_2'\big]. \tag{D.2}$$

By definition, $\Omega_{1,F}$ is the leading $r_1 \times r_1$ submatrix of $\Omega_{2,F}$.

Let F denote the joint distribution of $W = (Y, Z_1', Z^{*\prime}, X')'$ induced by θ_0, δ_0, and F^*. By definition, we can write

$$\delta_F = \Omega_{uu}\delta_0, G_{2,F} = \begin{pmatrix} -\Gamma_{z_1x} \\ -\delta_0\Gamma_{ux} - \Gamma_{vx} \end{pmatrix}, \qquad \Omega_{2,F} = \begin{pmatrix} \Omega_{z_1z_1u^2} & \Omega_{2,1r,F} \\ \Omega_{2,r1,F} & \Omega_{2,rr,F} \end{pmatrix}, \tag{D.3}$$

where

$$\begin{aligned} \Omega_{2,1r,F} &= \Gamma_{z_1u^3}\delta_0' + \Gamma_{z_1vu^2} = \Omega_{2,r1,F}', \quad \text{and} \\ \Omega_{2,rr,F} &= \Omega_{u^2u^2}\delta_0\delta_0' + \delta_0\Gamma_{u^3v} + \Gamma_{vu^3}\delta_0' + \Omega_{vvu^2}. \end{aligned} \tag{D.4}$$

Therefore, the parameter v_F defined in (3.4) depends on F through F^* and δ_0, and its dependence on F^* is through v_{*,F^*}, where

$$v_{*,F^*} = \begin{pmatrix} \Omega_{u^2u^2}, \Omega_{uu}, \mathrm{vec}(\Gamma_{z_1x})', \mathrm{vec}(\Gamma_{ux})', \mathrm{vec}(\Gamma_{vx})', \mathrm{vech}(\Omega_{z_1z_1u^2})', \\ \mathrm{vec}(\Gamma_{z_1u^3})', \mathrm{vec}(\Gamma_{z_1vu^2})', \mathrm{vec}(\Gamma_{u^3v})', \mathrm{vech}(\Omega_{vvu^2})' \end{pmatrix}. \tag{D.5}$$

Define

$$\begin{aligned} \rho_{2,\max} &\equiv \max\Big\{\sup_{F\in\mathcal{F}} \rho_{\max}(\Omega_{2,F}), \sup_{F\in\mathcal{F}} \rho_{\max}\big(G_{2,F}G_{2,F}'\big)\Big\}, \\ \rho_{2,\min} &\equiv \min\Big\{\inf_{F\in\mathcal{F}} \rho_{\min}(\Omega_{2,F}), \inf_{F\in\mathcal{F}} \rho_{\min}\big(G_{2,F}G_{2,F}'\big)\Big\}, \\ C_W &\equiv \sup_{F^*\in\mathcal{F}^*} \mathbb{E}_{F^*}\big[\|(X', Z_1', V', U)\|^2\big] \quad \text{and} \quad C_\Delta \equiv \sup_{\delta_0\in\Delta_\delta} \|\delta_0\|^2. \end{aligned} \tag{D.6}$$

In the proof of Lemma D.1 below, we show that $\rho_{2,\max} < \infty$ (see (D.14) and (D.18)). Moreover, we have $\rho_{2,\min} > 0$, $C_W < \infty$ and $C_\Delta < \infty$ by Assumptions D.1(iii), D.1(ii) and D.1(vii) respectively. Define

$$B_{\rho_2}^c \equiv \big\{\delta \in \mathbb{R}^{r^*} : \|\delta\| \geq \rho_{2,\min}\rho_{2,\max}^{-1}C_\Delta^{-1/2}\big\}. \tag{D.7}$$

Let Θ_0 be a nonempty set in $\mathbb{R}^{d_\theta}$. Define

$$B_{\Theta_0} \equiv \big\{\theta \in \mathbb{R}^{d_\theta} : \|\theta - \theta_0\| \leq \rho_{2,\min}^{-4}\rho_{2,\max}^3 C_\Delta C_W^2 \text{ for any } \theta_0 \in \Theta_0\big\}. \tag{D.8}$$

Let $\{c_{j,\Delta}, C_{j,\Delta}\}_{j=1}^{r^*}$ be a set of finite constants. We next provide the low-level sufficient conditions for Assumptions 3.1, 3.2, and 3.3.

ASSUMPTION D.1. *The following conditions hold*:

(i) $\mathbb{E}_{F^*}[V] = 0$, $\mathbb{E}_{F^*}[U] = 0$, $\mathbb{E}_{F^*}[Z_1 U] = 0_{r_1 \times 1}$ *and* $\mathbb{E}_{F^*}[VU] = 0_{r^* \times 1}$ *for any* $F^* \in \mathcal{F}^*$;

(ii) $\sup_{F^* \in \mathcal{F}^*} \mathbb{E}_{F^*}[\|X\|^{4+\gamma} + \|Z_1\|^{4+\gamma} + \|V\|^{4+\gamma} + U^6] < \infty$ *for some* $\gamma > 0$;

(iii) $\inf_{F^* \in \mathcal{F}^*} \mathbb{E}_{F^*}[U^2] > 0$, $\inf_{F^* \in \mathcal{F}^*} \rho_{\min}(\Gamma_{xz_1}\Gamma_{z_1 x}) > 0$ *and* $\inf_{F \in \mathcal{F}} \rho_{\min}(\Omega_{2,F}) > 0$;

(iv) $\inf_{F^* \in \mathcal{F}^*} \inf_{\delta \in B^c_{\rho_2}} \|\delta\|^{-1} \|(\Gamma_{xz_1}\Omega^{-1}_{z_1 z_1 u^2}\Gamma_{z_1 v u^2} - \Gamma_{xv})\delta + \Gamma_{xz_1}\Omega^{-1}_{z_1 z_1 u^2}\Gamma_{z_1 u^3} - \Gamma_{xu}\| > 0$;

(v) *the set* $\{v_{*,F^*} : F^* \in \mathcal{F}^*\}$ *is closed*;

(vi) $\theta_0 \in \Theta_0$, $B_{\Theta_0} \subset \mathrm{int}(\Theta)$ *and* Θ *is compact*;

(vii) $\Delta_\delta = [c_{1,\Delta}, C_{1,\Delta}] \times \cdots \times [c_{r^*,\Delta}, C_{r^*,\Delta}]$ *where* $c_{j,\Delta} < 0 < C_{j,\Delta}$ *for* $j = 1, \ldots, r^*$.

LEMMA D.1. *Suppose that* $\{W_i\}_{i=1}^n$ *are i.i.d. and generated by the linear model* (3.6) *and* (3.8) *in CLS. Then under Assumption* D.1, $\mathcal{F}$ *satisfies Assumptions* 3.1, 3.2, *and* 3.3.

For the linear IV model, Lemma D.1 provides simple conditions on θ_0, δ_0 and $\mathcal{F}^*$ on which uniformity results are subsequently established.

PROOF OF LEMMA D.1. By Assumption D.1(i) and the definition of $G_{1,F}$,

$$\mathbb{E}_F\big[g_1(W,\theta)\big] = \mathbb{E}_{F^*}\big[Z_1\big(U - X'(\theta-\theta_0)\big)\big] = G_{1,F}(\theta-\theta_0), \tag{D.9}$$

which together with Assumption D.1(iii) implies that $\theta_F = \theta_0$ and hence $\mathbb{E}_F[g_1(W,\theta_F)] = 0_{r_1 \times 1}$. Also $\theta_F \in \mathrm{int}(\Theta)$ holds by $\theta_F = \theta_0$ and Assumption D.1(vi). This verifies Assumption 3.1(i).

By (D.9) for any $\theta \in \Theta$ with $\|\theta - \theta_F\| \geq \varepsilon$ and any $F \in \mathcal{F}$,

$$\big\|\mathbb{E}_F\big[g_1(W,\theta)\big]\big\| \geq \rho^{1/2}_{\min}\big(G'_{1,F}G_{1,F}\big)\|\theta_F - \theta\| \geq \varepsilon\rho^{1/2}_{\min}\big(G'_{1,F}G_{1,F}\big), \tag{D.10}$$

which combined with Assumption D.1(iii) and $G_{1,F} = -\Gamma'_{xz_1,F^*}$ implies that

$$\inf_{F \in \mathcal{F}} \inf_{\theta \in B^c_\varepsilon(\theta_F)} \big\|\mathbb{E}_F\big[g_1(W,\theta)\big]\big\| > 0. \tag{D.11}$$

This verifies Assumption 3.1(ii).

Next, we show Assumption 3.1(iii). Let $Z_2 \equiv (Z_1', Z^{*\prime})'$. By the Lyapunov inequality, Assumptions D.1(i)–(ii) and D.1(vii),

$$\begin{aligned}\sup_{F \in \mathcal{F}} \mathbb{E}_F\big[\|Z_2\|^2\big] \leq & \sup_{F^* \in \mathcal{F}^*} \mathbb{E}_{F^*}\big[\|Z_1\|^2\big] + 2\sup_{F^* \in \mathcal{F}^*} \mathbb{E}_{F^*}\big[\|V\|^2\big] \\ & + 2\sup_{\delta_0 \in \Delta_\delta}\|\delta_0\|^2 \sup_{F^* \in \mathcal{F}^*} \mathbb{E}_{F^*}\big[U^2\big] < \infty. \end{aligned} \tag{D.12}$$

By (D.12), the Hölder inequality, the Lyapunov inequality and Assumption D.1(ii),

$$\sup_{F \in \mathcal{F}} \|G_{2,F}\| = \sup_{F \in \mathcal{F}} \big\|\mathbb{E}_F\big[Z_2 X'\big]\big\| \leq \sup_{F \in \mathcal{F}} \big(\mathbb{E}_F\big[\|Z_2\|^2\big]\big)^{1/2} \sup_{F^* \in \mathcal{F}^*} \big(\mathbb{E}_{F^*}\big[\|X\|^2\big]\big)^{1/2} < \infty, \tag{D.13}$$

which together with the definition of $G_{2,F}$ and the Cauchy–Schwarz inequality implies that

$$\sup_{F\in\mathcal{F}}\left\|G'_{2,F}G_{2,F}\right\|<\infty. \tag{D.14}$$

Similarly, by the Cauchy–Schwarz inequality, the Lyapunov inequality, Assumptions D.1(ii) and D.1(vii), we have

$$\begin{aligned}\sup_{F\in\mathcal{F}}\mathbb{E}_F\big[\|Z_2\|^4\big] &= \sup_{F\in\mathcal{F}}\mathbb{E}_F\big[\big(\|Z_1\|^2+\|Z^*\|^2\big)^2\big]\\ &\le 2\sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[\|Z_1\|^4\big]+2\sup_{F\in\mathcal{F}}\mathbb{E}_F\big[\|Z^*\|^4\big]\\ &\le 2\sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[\|Z_1\|^4\big]+8\sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[\|V\|^4\big]\\ &\quad +8\sup_{\delta_0\in\Delta_\delta}\|\delta_0\|^4\sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[U^4\big]<\infty. \end{aligned}\tag{D.15}$$

By (D.12), (D.15), Assumption D.1(ii), the Lyapunov inequality, and the Hölder inequality, we have

$$\begin{aligned}&\sup_{F\in\mathcal{F}}\left\|\mathbb{E}_F\big[Z_2Z_2'(Y-X'\theta_0)^2\big]\right\|\\ &\le\sup_{F\in\mathcal{F}}\mathbb{E}_F\big[\|Z_2\|^2(Y-X'\theta_0)^2\big]\\ &\le\sup_{F\in\mathcal{F}}\big(\mathbb{E}_F\big[\|Z_2\|^4\big]\big)^{1/2}\sup_{F^*\in\mathcal{F}^*}\big(\mathbb{E}_{F^*}\big[U^4\big]\big)^{1/2}<\infty,\end{aligned}\tag{D.16}$$

and

$$\sup_{F\in\mathcal{F}}\left\|\mathbb{E}_F\big[(Y-X'\theta_0)Z_2\big]\right\|\le\sup_{F\in\mathcal{F}}\big(\mathbb{E}_F\big[\|Z_2\|^2\big]\big)^{1/2}\sup_{F^*\in\mathcal{F}^*}\big(\mathbb{E}_{F^*}\big[U^2\big]\big)^{1/2}<\infty. \tag{D.17}$$

By the definition of $\Omega_{2,F}$, the triangle inequality, the Cauchy–Schwarz inequality and the results in (D.16) and (D.17),

$$\sup_{F\in\mathcal{F}}\|\Omega_{2,F}\|<\infty. \tag{D.18}$$

We then show that $\theta_F^*\in\mathrm{int}(\Theta)$. By the triangle inequality, the Cauchy–Schwarz inequality and the Hölder inequality,

$$\begin{aligned}\|G_{2,F}\| &\le \|\Gamma_{xz_1}\|+\|\delta_0\|\|\Gamma_{xu}\|+\|\Gamma_{xv}\|\\ &\le\big(\mathbb{E}_{F^*}\big[\|X\|^2\big]\big)^{1/2}\big(\mathbb{E}_{F^*}\big[\|Z_1\|^2\big]\big)^{1/2}\\ &\quad+\|\delta_0\|\big(\mathbb{E}_{F^*}\big[\|X\|^2\big]\big)^{1/2}\big(\mathbb{E}_{F^*}\big[U^2\big]\big)^{1/2}\\ &\quad+\big(\mathbb{E}_{F^*}\big[\|X\|^2\big]\big)^{1/2}\big(\mathbb{E}_{F^*}\big[\|V\|^2\big]\big)^{1/2}\\ &\le C_W\big(2+C_\Delta^{1/2}\big),\end{aligned}\tag{D.19}$$

for any $F \in \mathcal{F}$, where $C_W < \infty$ by Assumptions D.1(ii) and (vii). Since $G'_{2,F} = (G'_{1,F}, G'_{r^*,F})$ where $G_{r^*,F} = -\delta_0 \mathbb{E}_{F^*}[UX'] - \mathbb{E}_{F^*}[VX']$, we have

$$G'_{2,F} G_{2,F} = G'_{1,F} G_{1,F} + G'_{r^*,F} G_{r^*,F}, \tag{D.20}$$

which implies that for any $F \in \mathcal{F}$,

$$\rho_{\min}\big(G'_{2,F} G_{2,F}\big) \geq \rho_{\min}\big(G'_{1,F} G_{1,F}\big). \tag{D.21}$$

To show Assumption 3.1(iii), we write

$$\begin{aligned} Q_F(\theta) &= \mathbb{E}_F\big[Z_2\big(Y - X'\theta\big)\big]' \Omega^{-1}_{2,F} \mathbb{E}_F\big[Z_2\big(Y - X'\theta\big)\big] \\ &= \theta' G'_{2,F} \Omega^{-1}_{2,F} G_{2,F} \theta + 2\theta' G'_{2,F} \Omega^{-1}_{2,F} C_F + C'_F \Omega^{-1}_{2,F} C_F, \end{aligned} \tag{D.22}$$

where $C_F = \mathbb{E}_F[Z_2 Y]$. Since $G'_{2,F}\Omega^{-1}_{2,F} G_{2,F}$ is nonsingular by (D.18), (D.21), and Assumption D.1(iii), $Q_F(\theta)$ is minimized at $\theta^*_F = -(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F})^{-1} G'_{2,F}\Omega^{-1}_{2,F} C_F$ for any $F \in \mathcal{F}$. Therefore,

$$\begin{aligned} \|\theta^*_F - \theta_0\|^2 &= \big\|\big(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F}\big)^{-1} G'_{2,F}\Omega^{-1}_{2,F}\mathbb{E}_F[Z_2 U]\big\|^2 \\ &\leq \frac{\rho^2_{\max}(\Omega_{2,F})}{\rho^2_{\min}\big(G'_{2,F}G_{2,F}\big)} \mathbb{E}_F\big[U Z'_2\big]\Omega^{-1}_{2,F} G_{2,F} G'_{2,F} \Omega^{-1}_{2,F} \mathbb{E}_F[Z_2 U] \\ &\leq \frac{\rho^2_{\max}(\Omega_{2,F})\rho_{\max}\big(G'_{2,F}G_{2,F}\big)\Gamma^2_{uu}}{\rho^2_{\min}(\Omega_{2,F})\rho^2_{\min}\big(G'_{2,F}G_{2,F}\big)} \|\delta_0\|^2 \\ &\leq \rho^{-4}_{2,\min}\rho^3_{2,\max} C_\Delta C^2_W \end{aligned} \tag{D.23}$$

for any $F \in \mathcal{F}$. By Assumption D.1(vi), $\theta^*_F \in \operatorname{int}(\Theta)$. Moreover, for any $\theta \in \Theta$ with $\|\theta - \theta^*_F\| \geq \varepsilon$,

$$\begin{aligned} Q_F(\theta) - Q_F\big(\theta^*_F\big) &\geq \rho_{\min}\big(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F}\big)\|\theta - \theta^*_F\|^2 \\ &\geq \varepsilon^2 \rho_{\min}\big(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F}\big) \\ &\geq \varepsilon^2 \rho^{-1}_{\max}(\Omega_{2,F})\rho_{\min}\big(G'_{2,F}G_{2,F}\big), \end{aligned} \tag{D.24}$$

which together with (D.18), (D.21), and Assumption D.1(iii) implies that

$$\inf_{F\in\mathcal{F}} \inf_{\theta \in B^c_\varepsilon(\theta^*_F)} \big[Q_F(\theta) - Q_F\big(\theta^*_F\big)\big] > 0. \tag{D.25}$$

This verifies Assumption 3.1(iii).

Next, we verify Assumption 3.1(iv). Let $\Omega^{(22)}_{2,F} = (\Omega_{2,rr,F} - \Omega'_{2,r1,F}\Omega^{-1}_{z_1 z_1 u^2}\Omega_{2,1r,F})^{-1}$, where $\Omega_{2,1r,F}$ and $\Omega_{2,rr,F}$ are defined in (D.4). Then

$$G'_{2,F}\Omega^{-1}_{2,F}\delta_{2,F}$$

$$
\begin{aligned}
&= -\left(\Gamma_{xz_1}, \Gamma_{xv} + \Gamma_{xu}\delta_0'\right)\begin{pmatrix} -\Omega_{z_1z_1u^2}^{-1}\Omega_{2,1r,F} \\ I_{r^*} \end{pmatrix}\Omega_{2,F}^{(22)}\Omega_{uu}\delta_0 \\
&= \Omega_{uu}\left[\left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1u^3} - \Gamma_{xu}\right)\delta_0' + \Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1vu^2} - \Gamma_{xv}\right]\Omega_{2,F}^{(22)}\delta_0 \\
&= \Omega_{uu}\delta_0'\Omega_{2,F}^{(22)}\delta_0\left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1u^3} - \Gamma_{xu}\right) \\
&\quad + \Omega_{uu}\left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1vu^2} - \Gamma_{xv}\right)\Omega_{2,F}^{(22)}\delta_0, \qquad \text{(D.26)}
\end{aligned}
$$

by the formula of the inverse of partitioned matrix. For any $\delta_0 \in \Delta_\delta$ with $\|\delta_0\| > 0$, we have

$$
\frac{\delta_0'\left(\Omega_{2,F}^{(22)}\right)^2\delta_0}{\left(\delta_0'\Omega_{2,F}^{(22)}\delta_0\right)^2} \geq \frac{\left(\rho_{\min}\left(\Omega_{2,F}^{(22)}\right)\right)^2}{\left(\rho_{\max}\left(\Omega_{2,F}^{(22)}\right)\right)^2}\frac{1}{\delta_0'\delta_0} \geq \frac{\rho_{2,\min}^2}{C_\Delta\rho_{2,\max}^2} \qquad \text{(D.27)}
$$

and

$$
\delta_0'\Omega_{2,F}^{(22)}\delta_0 = \frac{\delta_0'\Omega_{2,F}^{(22)}\delta_0\left(\delta_0'\left(\Omega_{2,F}^{(22)}\right)^2\delta_0\right)^{1/2}}{\left(\delta_0'\left(\Omega_{2,F}^{(22)}\right)^2\delta_0\right)^{1/2}} \geq \frac{\|\delta_0\|}{\rho_{2,\max}}\frac{\delta_0'\Omega_{2,F}^{(22)}\delta_0}{\left(\delta_0'\left(\Omega_{2,F}^{(22)}\right)^2\delta_0\right)^{1/2}}, \qquad \text{(D.28)}
$$

where the last inequality in (D.27) and the inequality in (D.28) are due to

$$
\rho_{\min}\left(\Omega_{2,F}^{(22)}\right) \geq \rho_{\min}\left(\Omega_{2,F}^{-1}\right) = \rho_{2,\max}^{-1}
$$

and

$$
\rho_{\max}\left(\Omega_{2,F}^{(22)}\right) \leq \rho_{\max}\left(\Omega_{2,F}^{-1}\right) = \rho_{2,\min}^{-1}.
$$

Therefore, for any $F \in \mathcal{F}$ with $\delta_{2,F} = \Omega_{uu}(0_{1\times r_1}, \delta_0')'$ and $\|\delta_0\| > 0$,

$$
\begin{aligned}
\frac{\|G_{2,F}'\Omega_{2,F}^{-1}\delta_{2,F}\|}{\|\delta_{2,F}\|} &= \frac{\delta_0'\Omega_{2,F}^{(22)}\delta_0}{\|\delta_0\|}\left\| \begin{matrix} \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1vu^2} - \Gamma_{xv}\right)\dfrac{\Omega_{2,F}^{(22)}\delta_0}{\delta_0'\Omega_{2,F}^{(22)}\delta_0} \\ + \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1u^3} - \Gamma_{xu}\right) \end{matrix} \right\| \\
&\geq \frac{1}{\rho_{2,\max}}\frac{\delta_0'\Omega_{2,F}^{(22)}\delta_0}{\left(\delta_0'\left(\Omega_{2,F}^{(22)}\right)^2\delta_0\right)^{1/2}}\left\| \begin{matrix} \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1vu^2} - \Gamma_{xv}\right)\dfrac{\Omega_{2,F}^{(22)}\delta_0}{\delta_0'\Omega_{2,F}^{(22)}\delta_0} \\ + \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1u^3} - \Gamma_{xu}\right) \end{matrix} \right\| \\
&= \frac{1}{\rho_{2,\max}}\frac{1}{\|\widetilde{\delta}_0\|}\left\| \begin{matrix} \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1vu^2} - \Gamma_{xv}\right)\widetilde{\delta}_0 \\ + \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1u^3} - \Gamma_{xu}\right) \end{matrix} \right\|, \qquad \text{(D.29)}
\end{aligned}
$$

where $\widetilde{\delta}_0 \equiv \Omega_{2,F}^{(22)}\delta_0/\delta_0'\Omega_{2,F}^{(22)}\delta_0$ and the inequality is by (D.28). By (D.28) and the definition of $B_{\rho_2}^c$, $\widetilde{\delta}_0 \in B_{\rho_2}^c$. Therefore, (D.29) implies that

$$
\frac{\|G_{2,F}'\Omega_{2,F}^{-1}\delta_{2,F}\|}{\|\delta_{2,F}\|} \geq \frac{1}{\rho_{2,\max}}\inf_{\delta\in B_{\rho_2}^c}\|\delta\|^{-1}\left\| \begin{matrix} \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1vu^2} - \Gamma_{xv}\right)\delta \\ + \left(\Gamma_{xz_1}\Omega_{z_1z_1u^2}^{-1}\Gamma_{z_1u^3} - \Gamma_{xu}\right) \end{matrix} \right\|. \qquad \text{(D.30)}
$$

Collecting the results in (D.18) and (D.30) and then applying Assumption D.1(iv), we get

$$\inf_{\{F\in\mathcal{F}:\,\|\delta_F\|>0\}} \frac{\|G_{2,F}'\Omega_{2,F}^{-1}\delta_{2,F}\|}{\|\delta_{2,F}\|} > 0, \tag{D.31}$$

which shows Assumption 3.1(iv) with $\tau = 1$.

Assumption 3.1(v) is implied by Assumption D.1(vii). This finishes the verification of Assumption 3.1.

To verify Assumption 3.2, note that $g_2(W,\theta) = Z_2(U - X'(\theta-\theta_0))$, $g_{2,\theta}(W,\theta) = -Z_2X'$ and $g_{2,\theta\theta}(W,\theta) = 0_{(r_2 d_\theta)\times d_\theta}$. Therefore, Assumption 3.2(i) holds automatically. Moreover, Assumption 3.2(ii) is implied by Assumption D.1(ii) and the assumption that Θ is bounded. Assumptions 3.2(iii)–(iv) follow from Assumption D.1(iii).

We next verify Assumption 3.3. By definition,

$$v_F = \big(\mathrm{vec}(G_{2,F})', \mathrm{vech}(\Omega_{2,F})', \delta_F\big). \tag{D.32}$$

Let $\Lambda_* = \{v_{*,F^*} : F^* \in \mathcal{F}^*\}$. From the expressions in (D.3), we see that $\Lambda = \{v_F\colon F \in \mathcal{F}\}$ is the image of $\Lambda_* \times \Delta_\delta$ under a continuous mapping. By Assumption D.1(ii) and the Hölder inequality, Λ_* is bounded which together with Assumption D.1(v) implies that Λ_* is compact. Since Δ_δ is also a compact set by Assumption D.1(vii), we know that $\Lambda_* \times \Delta_\delta$ is compact. Therefore, Λ is compact, and hence closed. This verifies Assumption 3.3(ii).

Let $\varepsilon_F = \Omega_{uu}c_\Delta$ where $c_\Delta = \min\{\min_{j\le r^*}|c_{j,\Delta}|, \min_{j\le r^*}|C_{j,\Delta}|\}$. Below we show that for any $\widetilde{\delta} \in \mathbb{R}^{r^*}$ with $0 \le \|\widetilde{\delta}\| \le \varepsilon_F$, there is $\widetilde{F} \in \mathcal{F}$ such that

$$\widetilde{\delta}_{\widetilde{F}} = \widetilde{\delta}, \qquad \|G_{2,\widetilde{F}} - G_{2,F}\| \le C_1\|\widetilde{\delta}_F\|^{1/4} \quad \text{and} \quad \|\Omega_{2,\widetilde{F}} - \Omega_{2,F}\| \le C_2\|\widetilde{\delta}\|^{1/4} \tag{D.33}$$

for some fixed constants C_1 and C_2. This verifies Assumption 3.3(i) with $\kappa = 1/4$.

First, if $\widetilde{\delta} = 0_{r^*\times 1}$, then we set $\widetilde{F}$ to be F which is induced by δ_0, θ_0 and F^* with $\delta_0 = 0_{r^*\times 1}$. By definition, $G_{2,\widetilde{F}} = G_{2,F}$, $\Omega_{2,\widetilde{F}} = \Omega_{2,F}$, and $\widetilde{\delta}_{\widetilde{F}} = \delta_F = \delta_0\Omega_{uu} = 0 = \widetilde{\delta}$ which implies that (D.33) holds.

Second, consider any $\widetilde{\delta} \in \mathbb{R}^{r^*}$ with $0 < \|\widetilde{\delta}\| < \varepsilon_F$. Define $\widetilde{\delta}_0 = \widetilde{\delta}\Omega_{uu}^{-1}$. Since $\|\widetilde{\delta}\| < \varepsilon_F$ and $\varepsilon_F = \Omega_{uu}c_\Delta$,

$$\|\widetilde{\delta}_0\| = \big\|\widetilde{\delta}\Omega_{uu}^{-1}\big\| = \|\widetilde{\delta}\|\Omega_{uu}^{-1} < c_\Delta, \tag{D.34}$$

which combined with the definition of Δ_δ implies that $\widetilde{\delta}_0 \in \Delta_\delta$. Let $\widetilde{F}$ be the joint distribution induced by $\widetilde{\delta}_0$, θ_0, and F^*. By the definition of $\mathcal{F}$, we have $\widetilde{F} \in \mathcal{F}$. Moreover,

$$\widetilde{\delta}_{\widetilde{F}} = \widetilde{\delta}_0\Omega_{uu} = \widetilde{\delta}, \tag{D.35}$$

which verifies the equality in (D.33). By definition,

$$G_{2,\widetilde{F}} = \begin{pmatrix} -\mathbb{E}_{F^*}[Z_1X'] \\ -\widetilde{\delta}_0\mathbb{E}_{F^*}[UX'] - \mathbb{E}_{F^*}[VX'] \end{pmatrix} \quad \text{and} \quad G_{2,F} = \begin{pmatrix} -\mathbb{E}_{F^*}[Z_1X'] \\ -\mathbb{E}_{F^*}[VX'] \end{pmatrix}, \tag{D.36}$$

which together with the Cauchy–Schwarz inequality and the Hölder inequality implies that

$$\|G_{2,\widetilde{F}} - G_{2,F}\| = \big\|\widetilde{\delta}_0\mathbb{E}_{F^*}[UX']\big\| \le \|\widetilde{\delta}_0\|\big(\Omega_{uu}\mathbb{E}_{F^*}[\|X\|^2]\big)^{1/2}$$

$$= \|\widetilde{\delta}_0\|^{3/4}\Omega_{uu}^{1/4}\big(\mathbb{E}_{F^*}\big[\|X\|^2\big]\big)^{1/2}\|\widetilde{\delta}_0\Omega_{uu}\|^{1/4}. \tag{D.37}$$

By Assumption D.1(ii),

$$\sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[\|X\|^2\big] < \infty \quad \text{and} \quad \sup_{F^*\in\mathcal{F}^*}\Omega_{uu} < \infty, \tag{D.38}$$

which together with (D.34), (D.37), and the definition of $\widetilde{\delta}$ implies that

$$\|G_{2,\widetilde{F}} - G_{2,F}\| \le C_1\|\widetilde{\delta}\|^{1/4}, \tag{D.39}$$

where $C_1 = c_\Delta^{3/4}\sup_{F^*\in\mathcal{F}^*}(\mathbb{E}_{F^*}[\|X\|^2])^{1/2}\sup_{F^*\in\mathcal{F}^*}\Omega_{uu}^{1/4}$ is finite.

To show the last inequality in (D.33), note that by definition $\theta_{\widetilde{F}} = \theta_0 = \theta_F$, and hence

$$\mathbb{E}_{\widetilde{F}}\big[Z_1Z_1'\big(Y - X'\theta_{\widetilde{F}}\big)^2\big] = \mathbb{E}_{F^*}\big[Z_1Z_1'U^2\big] = \mathbb{E}_F\big[Z_1Z_1'\big(Y - X'\theta_F\big)^2\big]. \tag{D.40}$$

Under $\widetilde{F}$,

$$\mathbb{E}_{\widetilde{F}}\big[Z_1Z^{*\prime}\big(Y - X'\theta_{\widetilde{F}}\big)^2\big] = \mathbb{E}_{F^*}\big[Z_1(U\widetilde{\delta}_0 + V)'U^2\big] = \mathbb{E}_{F^*}\big[U^3Z_1\big]\widetilde{\delta}_0' + \mathbb{E}_{F^*}\big[U^2Z_1V'\big], \tag{D.41}$$

and

$$\begin{aligned}
&\mathbb{E}_{\widetilde{F}}\big[Z^*Z^{*\prime}\big(Y - X'\theta_{\widetilde{F}}\big)^2\big] \\
&\quad = \mathbb{E}_{F^*}\big[(U\widetilde{\delta}_0 + V)(U\widetilde{\delta}_0 + V)'U^2\big] \\
&\quad = \mathbb{E}_{F^*}\big[U^4\big]\widetilde{\delta}_0\widetilde{\delta}_0' + \widetilde{\delta}_0\mathbb{E}_{F^*}\big[U^3V'\big] + \mathbb{E}_{F^*}\big[U^3V\big]\widetilde{\delta}_0' + \mathbb{E}_{F^*}\big[U^2VV'\big].
\end{aligned} \tag{D.42}$$

Under F,

$$\begin{aligned}
&\mathbb{E}_F\big[Z_1Z^{*\prime}\big(Y - X'\theta_F\big)^2\big] = \mathbb{E}_{F^*}\big[U^2Z_1V'\big] \quad \text{and} \\
&\mathbb{E}_F\big[Z^*Z^{*\prime}\big(Y - X'\theta_F\big)^2\big] = \mathbb{E}_{F^*}\big[U^2VV'\big].
\end{aligned} \tag{D.43}$$

Collecting the results in (D.40), (D.41), (D.42), and (D.43), and applying the triangle inequality, we get

$$\begin{aligned}
&\big\|\mathbb{E}_{\widetilde{F}}\big[Z_2Z_2'\big(Y - X'\theta_{\widetilde{F}}\big)^2\big] - \mathbb{E}_F\big[Z_2Z_2'\big(Y - X'\theta_F\big)^2\big]\big\| \\
&\quad \le \big\|\mathbb{E}_{F^*}\big[U^3Z_1\big]\widetilde{\delta}_0'\big\| + \big\|\mathbb{E}_{F^*}\big[U^4\big]\widetilde{\delta}_0\widetilde{\delta}_0'\big\| \\
&\qquad + \big\|\widetilde{\delta}_0\mathbb{E}_{F^*}\big[U^3V'\big]\big\| + \big\|\mathbb{E}_{F^*}\big[U^3V\big]\widetilde{\delta}_0'\big\|.
\end{aligned} \tag{D.44}$$

By Assumption D.1(ii) and the Lyapunov inequality,

$$\sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[|U|^5\big] < \infty, \qquad \sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[\|Z_1\|^4\big] < \infty \quad \text{and} \quad \sup_{F^*\in\mathcal{F}^*}\mathbb{E}_{F^*}\big[\|V\|^4\big] < \infty. \tag{D.45}$$

By the Hölder inequality,

$$\big\|\mathbb{E}_{F^*}\big[U^3Z_1\big]\big\| \le \big(\mathbb{E}_{F^*}\big[|U|\|Z_1\|^2\big]\mathbb{E}_{F^*}\big[|U|^5\big]\big)^{1/2}$$

$$\leq \big(\mathbb{E}_{F^*}\big[|U|^5\big]\big)^{1/2}\big(\mathbb{E}_{F^*}\big[\|Z_1\|^4\big]\big)^{1/4}\big(\mathbb{E}_{F^*}\big[U^2\big]\big)^{1/4}$$
$$= \Omega_{uu}^{1/4}\big(\mathbb{E}_{F^*}\big[|U|^5\big]\big)^{1/2}\big(\mathbb{E}_{F^*}\big[\|Z_1\|^4\big]\big)^{1/4}. \tag{D.46}$$

Similarly, we can show that

$$\big\|\mathbb{E}_{F^*}\big[U^3V'\big]\big\| \leq \Omega_{uu}^{1/4}\big(\mathbb{E}_{F^*}\big[|U|^5\big]\big)^{1/2}\big(\mathbb{E}_{F^*}\big[\|V\|^4\big]\big)^{1/4} \tag{D.47}$$

and

$$\mathbb{E}_{F^*}\big[U^4\big] \leq \big(\mathbb{E}_{F^*}\big[U^2\big]\mathbb{E}_{F^*}\big[U^6\big]\big)^{1/2} = \Omega_{uu}^{1/4}\sup_{F^*\in\mathcal{F}^*}\big(\mathbb{E}_{F^*}\big[U^6\big]\big)^{1/2}. \tag{D.48}$$

Let $C_{2,0} = \sup_{F^*\in\mathcal{F}^*}\{(\mathbb{E}_{F^*}[|U|^5])^{1/2}[(\mathbb{E}_{F^*}[\|Z_1\|^4])^{1/4} + (\mathbb{E}_{F^*}[\|V\|^4])^{1/4}] + (\mathbb{E}_{F^*}[U^6])^{1/2}\}$. Combining the results in (D.44), (D.46), (D.47), and (D.48), and applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned}
&\big\|\mathbb{E}_{\widetilde{F}}\big[Z_2Z_2'\big(Y - X'\theta_{\widetilde{F}}\big)^2\big] - \mathbb{E}_F\big[Z_2Z_2'\big(Y - X'\theta_F\big)^2\big]\big\| \\
&\quad\leq 3C_{2,0}\Omega_{uu}^{1/4}\|\widetilde{\delta}_0\| + C_{2,0}\Omega_{uu}^{1/4}\|\widetilde{\delta}_0\|^2 \\
&\quad= \big(3C_{2,0}\|\widetilde{\delta}_0\|^{3/4} + C_{2,0}\|\widetilde{\delta}_0\|^{7/4}\big)\Omega_{uu}^{1/4}\|\widetilde{\delta}_0\|^{1/4} \leq C_{2,1}\|\widetilde{\delta}\|^{1/4},
\end{aligned} \tag{D.49}$$

where $C_{2,1} = C_{2,0}(3c_\Delta^{3/4} + c_\Delta^{7/4})$, the second inequality is by (D.34) and the definition of $\widetilde{\delta}$. By (D.45), Assumption D.1(ii) and the definition of c_Δ,

$$C_{2,1} < \infty. \tag{D.50}$$

Next, note that

$$\mathbb{E}_{\widetilde{F}}\big[Z_2\big(Y - X'\theta_{\widetilde{F}}\big)\big] = \begin{pmatrix}\mathbb{E}_{F^*}[Z_1U] \\ \widetilde{\delta}_0\Omega_{uu}\end{pmatrix} \quad\text{and}\quad \mathbb{E}_F\big[Z_2\big(Y - X'\theta_F\big)\big] = \begin{pmatrix}\mathbb{E}_{F^*}[Z_1U] \\ 0_{r^*\times 1}\end{pmatrix}, \tag{D.51}$$

which implies that

$$\begin{aligned}
&\left\|\begin{matrix}\mathbb{E}_{\widetilde{F}}\big[Z_2\big(Y - X'\theta_{\widetilde{F}}\big)\big]\mathbb{E}_{\widetilde{F}}\big[Z_2'\big(Y - X'\theta_{\widetilde{F}}\big)\big] \\ -\,\mathbb{E}_F\big[Z_2\big(Y - X'\theta_F\big)\big]\mathbb{E}_F\big[Z_2'\big(Y - X'\theta_F\big)\big]\end{matrix}\right\| \\
&\quad= \left\|\begin{pmatrix}0_{r_1\times r_1} & \Omega_{uu,F^*}\mathbb{E}_{F^*}[Z_1U]\widetilde{\delta}_0' \\ \widetilde{\delta}_0\mathbb{E}_{F^*}\big[Z_1'U\big]\Omega_{uu} & \widetilde{\delta}_0\widetilde{\delta}_0'\Omega_{uu}^2\end{pmatrix}\right\| \\
&\quad\leq \Omega_{uu}\big\|\mathbb{E}_{F^*}[Z_1U]\widetilde{\delta}_0'\big\| + \Omega_{uu}\big\|\widetilde{\delta}_0\mathbb{E}_{F^*}[Z_1U]\big\| + \Omega_{uu}^2\|\widetilde{\delta}_0\|^2 \\
&\quad\leq 2\Omega_{uu}\|\widetilde{\delta}_0\|\,\big\|\mathbb{E}_{F^*}[Z_1U]\big\| + \Omega_{uu}^2\|\widetilde{\delta}_0\|^2 \\
&\quad\leq \big(2\Omega_{uu}^{5/4}\|\widetilde{\delta}_0\|^{3/4}\big(\mathbb{E}_{F^*}\big[\|Z_1\|^2\big]\big)^{1/2} + \Omega_{uu}^{7/4}\|\widetilde{\delta}_0\|^{7/4}\big)\|\widetilde{\delta}\|^{1/4} \\
&\quad\leq C_{2,2}\|\widetilde{\delta}\|^{1/4},
\end{aligned} \tag{D.52}$$

where $C_{2,2} = \sup_{F^* \in \mathcal{F}^*}\{2\Omega_{uu}^{5/4}(\mathbb{E}_{F^*}[\|Z_1\|^2])^{1/2}c_\Delta^{3/4} + \Omega_{uu}^{7/4}c_\Delta^{7/4}\}$, the second inequality is by the Cauchy–Schwarz inequality, the third inequality is by the Hölder inequality. By Assumption D.1(ii) and the definition of c_Δ,

$$C_{2,2} < \infty. \tag{D.53}$$

By the definition of $\Omega_{2,F}$ in (D.2), we can use the triangle inequality and the results in (D.49) and (D.52) to deduce that

$$\|\Omega_{2,\widetilde{F}} - \Omega_{2,F}\| \leq C_2\|\widetilde{\delta}\|^{1/4}, \tag{D.54}$$

where $C_2 = C_{2,1} + C_{2,2}$ and $C_2 < \infty$ by (D.50) and (D.53), which proves the second inequality in (D.33). This verifies Assumption 3.3(i) with $\kappa = 1/4$. □

Proof of Lemma 3.1. Next, we apply Lemma D.1 to prove Lemma 3.1 in the paper. For convenience, the conditions of Lemma 3.1 are stated here. The proof verifies the conditions of Lemma D.1 with the following conditions in a Gaussian model. Let $\mathcal{F}^*$ denote the set of normal distributions which satisfies:

(i) $\phi_u = 0$, $\Gamma_{z_1 u} = 0_{r_1\times 1}$ and $\Gamma_{vu} = 0_{r^*\times 1}$;

(ii) $\inf_{F^*\in\mathcal{F}^*}\rho_{\min}(\Gamma_{xz_1}\Gamma_{z_1 x}) > 0$, $\sup_{F^*\in\mathcal{F}^*}\|\phi\|^2 < \infty$ and $0 < \inf_{F^*\in\mathcal{F}^*}\rho_{\min}(\Psi) \leq \sup_{F^*\in\mathcal{F}^*}\rho_{\max}(\Psi) < \infty$;

(iii) $\inf_{F^*\in\mathcal{F}^*}\inf_{\{\|\delta\|\geq\varepsilon\}}\|\delta\|^{-1}\|(\Gamma_{xz_1}\Gamma_{z_1z_1}^{-1}\Gamma_{z_1 v} - \Gamma_{xv})\delta - \Gamma_{xu}\| > 0$ for some $\varepsilon > 0$ that is small enough (where ε is given in (B.3) in the Appendix of CLS);

(iv) $\theta_0 \in \mathrm{int}(\Theta)$ and Θ is compact and large enough such that the pseudo-true value $\theta^*(F) \in \mathrm{int}(\Theta)$;

(v) $\Delta_\delta = [c_{1,\Delta}, C_{1,\Delta}] \times \cdots \times [c_{r^*,\Delta}, C_{r^*,\Delta}]$ where $\{c_{j,\Delta}, C_{j,\Delta}\}_{j=1}^{r^*}$ is a set of finite constants with $c_{j,\Delta} < 0 < C_{j,\Delta}$ for $j = 1, \ldots, r^*$.

Specifically, we assume that Condition (ii) of Lemma 3.1 holds with some constants c_ρ and C_ρ such that $c_\rho \leq \rho_{\min}(\Gamma_{xz_1}\Gamma_{z_1x})$, $\|\phi\|^2 \leq C_\rho$, and $c_\rho \leq \rho_{\min}(\Psi) \leq \rho_{\max}(\Psi) \leq C_\rho$; Condition (iii) of Lemma 3.1 holds with

$$\inf_{\delta\in B_\varepsilon^c}\|\delta\|^{-1}\left\|\left(\Gamma_{xz_1}\Gamma_{z_1z_1}^{-1}\Gamma_{z_1v} - \Gamma_{xv}\right)\delta - \Gamma_{xu}\right\| \geq c_\Gamma \tag{D.55}$$

for some positive constant c_Γ and

$$B_\varepsilon^c \equiv \left\{\delta \in \mathbb{R}^{r^*} : \|\delta\| \geq c_{*,\rho}C_{*,\rho}^{-1}C_\Delta^{-1}\right\}, \tag{D.56}$$

where

$$C_{*,W} \equiv 2(d_\theta + r_2 + 1)C_\rho, \qquad c_{*,\rho} \equiv \min\{1, c_\rho^2\} \quad \text{and} \quad C_{*,\rho} \equiv C_{*,W}^2\left(2 + C_\Delta^{1/2}\right)^2 \tag{D.57}$$

and $C_\Delta \equiv \sup_{\delta_0\in\Delta_\delta}\|\delta_0\|^2$.

Assumption D.1(i) holds under Condition (i) of Lemma 3.1. Since $(X', Z_1', V', U)'$ is a normal random vector, Assumption D.1(ii) holds by $\|\phi\|^2 \leq C_\rho$ and $\rho_{\max}(\Psi) \leq$

C_ρ. By $\rho_{\min}(\Psi) \geq c_\rho$ and $\phi_u = 0$, we have $\mathbb{E}_{F^*}[U^2] \geq c_\rho$ for any $F^* \in \mathcal{F}^*$, and hence $\inf_{F^*\in\mathcal{F}^*} \mathbb{E}_{F^*}[U^2] > 0$. Let F denote the distribution of W induced by F^* with mean ϕ and variance–covariance matrix Ψ. By definition, $G_{1,F} = -\mathbb{E}_{F^*}[Z_1 X'] = \Gamma_{z_1 x}$. Therefore,

$$\inf_{F\in\mathcal{F}} \rho_{\min}\big(G'_{1,F} G_{1,F}\big) \geq c_\rho > 0 \tag{D.58}$$

holds by $\rho_{\min}(\Gamma_{xz_1}\Gamma_{z_1x}) \geq c_\rho > 0$ for any $F^* \in \mathcal{F}^*$. Since $\Gamma_{z_1u} = 0_{r_1\times 1}$ and $\Gamma_{vu} = 0_{r^*\times 1}$ for any $F^* \in \mathcal{F}^*$, U is independent with respect to $(Z'_1, V')'$ under the normality assumption. Therefore, by Condition (i) of Lemma 3.1,

$$\begin{aligned}\Omega_{2,F} &= \begin{pmatrix} \Omega_{uu}\Gamma_{z_1z_1} & \Omega_{uu}\Gamma_{z_1v} \\ \Omega_{uu}\Gamma'_{z_1v} & 2\Omega_{uu}^2\delta_0\delta_0' + \Omega_{uu}\Gamma_{vv}\end{pmatrix} \\ &= \Omega_{uu}\begin{pmatrix} \Omega_{z_1z_1} & \Omega_{z_1v} \\ \Omega_{vz_1} & \Omega_{vv}\end{pmatrix} + \Omega_{uu}\begin{pmatrix}\phi_{z_1}\\ \phi_v\end{pmatrix}\begin{pmatrix}\phi_{z_1}\\ \phi_v\end{pmatrix}' + \begin{pmatrix} 0_{r_1\times r_1} & 0_{r_1\times r^*} \\ 0_{r^*\times r_1} & 2\Omega_{uu}^2\delta_0\delta_0'\end{pmatrix},\end{aligned} \tag{D.59}$$

which implies that $\rho_{\min}(\Omega_{2,F}) \geq \rho^2_{\min}(\Psi)$ where F is the distribution of W induced by F^* with mean ϕ and variance–covariance matrix Ψ. Since $\rho_{\min}(\Psi) \geq c_\rho > 0$, we have

$$\inf_{F\in\mathcal{F}} \rho_{\min}(\Omega_{2,F}) \geq c_\rho^2 > 0. \tag{D.60}$$

This completes the proof of Assumption D.1(iii).

By (D.59), Conditions (ii), and (v) of Lemma 3.1

$$\sup_{F\in\mathcal{F}} \rho_{\max}(\Omega_{2,F}) \leq \rho^2_{\max}(\Psi) + \rho_{\max}(\Psi)\|\phi\|^2 + 2\rho^2_{\max}(\Psi)C_\Delta \leq 2C_\rho^2(1+C_\Delta). \tag{D.61}$$

By (D.19) in the proof of Lemma D.1,

$$\|G_{2,F}\| \leq 2C_\rho(d_\theta + r_2 + 1)\big(2 + C_\Delta^{1/2}\big),$$

which implies that

$$\sup_{F\in\mathcal{F}} \rho_{\max}\big(G'_{2,F}G_{2,F}\big) \leq 4C_\rho^2(d_\theta + r_2 + 1)^2\big(2 + C_\Delta^{1/2}\big)^2. \tag{D.62}$$

By (D.58) and (D.60),

$$\min\Big\{\inf_{F\in\mathcal{F}} \rho_{\min}(\Omega_{2,F}), \inf_{F\in\mathcal{F}} \rho_{\min}\big(G'_{2,F}G_{2,F}\big)\Big\} \geq \min\{1, c_\rho^2\}. \tag{D.63}$$

By (D.61) and (D.62),

$$\max\Big\{\sup_{F\in\mathcal{F}} \rho_{\max}(\Omega_{2,F}), \sup_{F\in\mathcal{F}} \rho_{\max}\big(G'_{2,F}G_{2,F}\big)\Big\} \leq 4C_\rho^2(d_\theta + r_2 + 1)^2\big(2 + C_\Delta^{1/2}\big)^2. \tag{D.64}$$

From (D.63), (D.64), the definitions of $c_{*,\rho}$, $C_{*,\rho}$, and $B^c_{N,\rho}$, we have $B^c_{\rho_2} \subset B^c_{N,\rho}$ where $B^c_{\rho_2}$ is defined in (D.7). Moreover, by $\phi_u = 0$, the normality assumption and the independence between U and $(Z'_1, V')'$, we have $\Omega_{z_1z_1u^2} = \Omega_{uu}\Gamma_{z_1z_1}$, $\Gamma_{z_1vu^2} = \Omega_{uu}\Gamma_{z_1v}$, and

$\Gamma_{z_1u^3}=0_{r_1\times 1}$, which implies that

$$\begin{aligned}&\left\|\left(\Gamma_{xz_1}\Omega^{-1}_{z_1z_1u^2}\Gamma_{z_1vu^2}-\Gamma_{xv}\right)\delta+\Gamma_{xz_1}\Omega^{-1}_{z_1z_1u^2}\Gamma_{z_1u^3}-\Gamma_{xu}\right\|\\&=\left\|\left(\Gamma_{xz_1}\Gamma^{-1}_{z_1z_1}\Gamma_{z_1v}-\Gamma_{xv}\right)\delta-\Gamma_{xu}\right\|.\end{aligned}\tag{D.65}$$

Assumption D.1(iv) follows by $B^c_{\rho_2}\subset B^c_{N,\rho}$, (D.65) and Condition (iii) of the lemma.

We next show that Assumption D.1(v) holds. Define

$$\overline{v}_{*,F^*}=\begin{pmatrix}\Omega_{uu},\mathrm{vec}(\Gamma_{xz_1})',\mathrm{vec}(\Gamma_{xu})',\mathrm{vec}(\Gamma_{xv})',\\ \mathrm{vec}(\Gamma_{z_1v})',\mathrm{vech}(\Gamma_{z_1z_1})',\mathrm{vech}(\Gamma_{vv})'\end{pmatrix}.$$

Under Condition (i) of Lemma 3.1 and the normality assumption, $\Gamma_{u^2u^2}=3\Omega^2_{uu}$, $\Gamma_{z_1u^3}=0_{r_1\times 1}$, $\Gamma_{vu^3}=0_{r^*\times 1}$, $\Omega_{z_1z_1u^2}=\Omega_{uu}\Gamma_{z_1z_1}$, $\Gamma_{z_1vu^2}=\Omega_{uu}\Gamma_{z_1v}$, and $\Omega_{vvu^2}=\Omega_{uu}\Gamma_{vv}$. Therefore, to verify Assumption D.1(v), it is sufficient to show that the set $\{\overline{v}_{*,F^*}:F^*\in\mathcal{F}^*\}$ is compact because the set $\{v_{*,F^*}:F^*\in\mathcal{F}^*\}$ is the image of the set $\{\overline{v}_{*,F^*}:F^*\in\mathcal{F}^*\}$ under a continuous mapping. Let $\{(\phi_n,\Psi_n)\}_n$ be a convergent sequence where (ϕ_n,Ψ_n) satisfies Conditions (i)–(iii) of Lemma 3.1 for any n. Let $\widetilde{\phi}$ and $\widetilde{\Psi}$ denote the limits of ϕ_n and Ψ_n under the Euclidean norm, respectively. We first show that Conditions (i)–(iii) of Lemma 3.1 hold for $(\widetilde{\phi},\widetilde{\Psi})$. Since $\phi_{u,n}=0$, $\Gamma_{z_1u,n}=0_{r_1\times 1}$ and $\Gamma_{vu,n}=0_{r^*\times 1}$ for any n, we have $\widetilde{\phi}_u=0$, $\widetilde{\Gamma}_{z_1u}=0_{r_1\times 1}$ and $\widetilde{\Gamma}_{vu}=0_{r^*\times 1}$ which shows that $(\widetilde{\phi},\widetilde{\Psi})$ satisfies Condition (i) of Lemma 3.1. Since $\phi_n\to\widetilde{\phi}$ and $\|\phi_n\|^2\le C_\rho$ for any n, we have $\|\widetilde{\phi}\|^2\le C_\rho$. By the convergence of (ϕ_n,Ψ_n), $\Gamma_{xz_1,n}\to\widetilde{\Gamma}_{xz_1}$. Since the roots of a polynomial continuously depends on its coefficients, we have

$$\begin{aligned}&\rho_{\min}\left(\Gamma_{xz_1,n}\Gamma'_{xz_1,n}\right)\to\rho_{\min}\left(\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}'_{xz_1}\right),\qquad\rho_{\min}(\Psi_n)\to\rho_{\min}(\widetilde{\Psi})\quad\text{and}\\&\rho_{\max}(\Psi_n)\to\rho_{\max}(\widetilde{\Psi}),\end{aligned}$$

which together with the assumption that $\Gamma_{xz_1,n}$ and Ψ_n satisfy Condition (ii) of Lemma 3.1 implies that

$$c_\rho\le\rho_{\min}\left(\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}'_{xz_1}\right)\quad\text{and}\quad c_\rho\le\rho_{\min}(\widetilde{\Psi})\le\rho_{\max}(\widetilde{\Psi})\le C_\rho.$$

This shows that Condition (ii) of Lemma 3.1 holds for $(\widetilde{\phi},\widetilde{\Psi})$. For any $\delta\in B^c_{N,\rho}$, by the triangle inequality, the Cauchy–Schwarz inequality and $\|\delta\|\ge c_\rho^2C_\rho^{-2}C_\Delta^{-1}(1+C_\Delta)^{-1}2^{-1}$,

$$\begin{aligned}&\|\delta\|^{-1}\left\|\left(\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}^{-1}_{z_1z_1}\widetilde{\Gamma}_{z_1v}-\widetilde{\Gamma}_{xv}\right)\delta-\widetilde{\Gamma}_{xu}\right\|\\&\ge\|\delta\|^{-1}\left\|\left(\Gamma_{xz_{1,n}}\Gamma^{-1}_{z_1z_1,n}\Gamma_{z_1v,n}-\Gamma_{xv,n}\right)\delta-\Gamma_{xu,n}\right\|\\&\quad-\left\|\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}^{-1}_{z_1z_1}\widetilde{\Gamma}_{z_1v}-\Gamma_{xz_{1,n}}\Gamma^{-1}_{z_1z_1,n}\Gamma_{z_1v,n}\right\|\\&\quad-\|\widetilde{\Gamma}_{xv}-\Gamma_{xv,n}\|-2C_\rho^2C_\Delta(1+C_\Delta)c_\rho^{-2}\|\Gamma_{xu,n}-\widetilde{\Gamma}_{xu}\|,\end{aligned}$$

which together with the convergence of (ϕ_n,Ψ_n) and Conditions (ii)–(iii) of Lemma 3.1 implies that

$$\|\delta\|^{-1}\left\|\left(\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}^{-1}_{z_1z_1}\widetilde{\Gamma}_{z_1v}-\widetilde{\Gamma}_{xv}\right)\delta-\widetilde{\Gamma}_{xu}\right\|$$

$$\geq c_\Gamma - \|\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}_{z_1z_1}^{-1}\widetilde{\Gamma}_{z_1v} - \Gamma_{xz_{1,n}}\Gamma_{z_1z_{1,n}}^{-1}\Gamma_{z_1v,n}\| \\ - \|\widetilde{\Gamma}_{xv} - \Gamma_{xv,n}\| - 2C_\rho^2 C_\Delta(1+C_\Delta)c_\rho^{-2}\|\Gamma_{xu,n} - \widetilde{\Gamma}_{xu}\|$$

for any n. Let n go to infinity, we get

$$\|\delta\|^{-1}\|(\widetilde{\Gamma}_{xz_1}\widetilde{\Gamma}_{z_1z_1}^{-1}\widetilde{\Gamma}_{z_1v} - \widetilde{\Gamma}_{xv})\delta - \widetilde{\Gamma}_{xu}\| \geq c_\Gamma$$

for any $\delta \in B_\varepsilon^c$. This shows that Condition (iii) of Lemma 3.1 also holds for $(\widetilde{\phi}, \widetilde{\Psi})$. Hence the set of (ϕ, Ψ) which satisfies Conditions (i)–(iii) of Lemma 3.1 is closed. By Conditions (i)–(ii) of the lemma, we know that this set is compact because it is also bounded. Let F^* denote the normal distribution with mean ϕ and variance–covariance matrix Ψ. Then $\overline{v}_{*,F^*}$ is the image of (ϕ, Ψ) under a continuous mapping, which implies that $\{\overline{v}_{*,F^*} : F^* \in \mathcal{F}^*\}$ is compact. Therefore, the set $\{v_{*,F^*} : F^* \in \mathcal{F}^*\}$ is compact, and hence closed. This proves Assumption D.1(v).

Assumption D.1(vi) is used to show that $\theta_F \in \mathrm{int}(\Theta)$ and $\theta_F^* \in \mathrm{int}(\Theta)$ for any $F \in \mathcal{F}$. By $\theta_F = \theta_0$ and Condition (iv) of Lemma 3.1, we have $\theta_F \in \mathrm{int}(\Theta)$ and $\theta_F^* \in \mathrm{int}(\Theta)$.

Finally, Assumption D.1(vii) is the same as Condition (v) of Lemma 3.1. □

Appendix E: Proof of some auxiliary results in Sections 4 and 5 of CLS

Proof of Lemma B.2. (i) Let $g_{2,j}(w, \theta)$ denote the jth ($j = 1, \dots, r_2$) component of $g_2(w, \theta)$. By the mean value expansion,

$$g_{2,j}(w, \theta_1) - g_{2,j}(w, \theta_2) = g_{2,j,\theta}(w, \widetilde{\theta}_{1,2})(\theta_1 - \theta_2) \tag{E.1}$$

for any $j = 1, \dots, r_2$, where $\widetilde{\theta}_{1,2}$ is some vector between θ_1 and θ_2. By (E.1) and the Cauchy–Schwarz inequality

$$\left|\mathbb{E}_F[g_{2,j}(w, \theta_1) - g_{2,j}(w, \theta_2)]\right| \leq \mathbb{E}_F\Big[\sup_{\theta\in\Theta}\|g_{2,\theta}(W, \theta)\|\Big]\|\theta_1 - \theta_2\|, \tag{E.2}$$

for any $j = 1, \dots, r_2$. By (E.2), we deduce that

$$\begin{aligned}\|M_{2,F}(\theta_1) - M_{2,F}(\theta_2)\| &\leq \sqrt{r_2}\mathbb{E}_F\Big[\sup_{\theta\in\Theta}\|g_{2,\theta}(W, \theta)\|\Big]\|\theta_1 - \theta_2\| \\ &\leq C_{M,1}\sqrt{r_2}\|\theta_1 - \theta_2\| \end{aligned}\tag{E.3}$$

for any $F \in \mathcal{F}$, where $C_{M,1} \equiv \sup_{F\in\mathcal{F}} \mathbb{E}_F[\sup_{\theta\in\Theta}\|g_{2,\theta}(W, \theta)\|]$ and $C_{M,1} < \infty$ by Assumption 3.2(ii). This immediately proves the claim in (i). The claim in (ii) follows by similar argument and its proof is omitted.

(iii) By the mean value expansion,

$$\begin{aligned}&g_{2,j_1}(w, \theta_1)g_{2,j_2}(w, \theta_1) - g_{2,j_1}(w, \theta_2)g_{2,j_2}(w, \theta_2) \\ &\quad = \big[g_{2,j_1,\theta}(w, \widetilde{\theta}_{1,2})g_{2,j_2}(w, \widetilde{\theta}_{1,2}) + g_{2,j_1}(w, \widetilde{\theta}_{1,2})g_{2,j_2,\theta}(w, \widetilde{\theta}_{1,2})\big](\theta_1 - \theta_2)\end{aligned} \tag{E.4}$$

for any $j_1, j_2 = 1, \ldots, r_2$, where $\widetilde{\theta}_{1,2}$ is some vector between θ_1 and θ_2 and may take different values from the $\widetilde{\theta}_{1,2}$ in (E.1). By (E.4), the triangle inequality and the Cauchy–Schwarz inequality

$$\begin{aligned}
&\left|\mathbb{E}_F\left[g_{2,j_1}(w,\theta_1)g_{2,j_2}(w,\theta_1) - g_{2,j_1}(w,\theta_2)g_{2,j_2}(w,\theta_2)\right]\right| \\
&\quad \le 2\mathbb{E}_F\left[\sup_{\theta\in\Theta}\left\|g_2(W,\theta)\right\|\left\|g_{2,\theta}(W,\theta)\right\|\right]\|\theta_1-\theta_2\| \\
&\quad \le \mathbb{E}_F\left[\sup_{\theta\in\Theta}\left(\left\|g_2(W,\theta)\right\|^2 + \left\|g_{2,\theta}(W,\theta)\right\|^2\right)\right]\|\theta_1-\theta_2\|
\end{aligned} \tag{E.5}$$

for any $j_1, j_2 = 1, \ldots, r_2$, where the second inequality is by the simple inequality that $|ab| \le (a^2+b^2)/2$. By (E.5),

$$\begin{aligned}
&\left\|\mathbb{E}_F\left[g_2(W,\theta_1)g_2(W,\theta_1)' - g_2(W,\theta_2)g_2(W,\theta_2)'\right]\right\| \\
&\quad \le r_2\mathbb{E}_F\left[\sup_{\theta\in\Theta}\left(\left\|g_2(W,\theta)\right\|^2 + \left\|g_{2,\theta}(W,\theta)\right\|^2\right)\right]\|\theta_1-\theta_2\| \\
&\quad \le r_2 C_{M,2}\|\theta_1-\theta_2\|
\end{aligned} \tag{E.6}$$

for any $F \in \mathcal{F}$, where $C_{M,2} \equiv \sup_{F\in\mathcal{F}} \mathbb{E}_F[\sup_{\theta\in\Theta}(\|g_2(W,\theta)\|^2 + \|g_{2,\theta}(W,\theta)\|^2)]$ and $C_{M,2} < \infty$ by Assumption 3.2(ii). Using the triangle inequality, and the inequality in (E.2), we deduce that

$$\begin{aligned}
&\left|\mathbb{E}_F\left[g_{2,j_1}(w,\theta_1)\right]\mathbb{E}_F\left[g_{2,j_2}(w,\theta_1)\right] - \mathbb{E}_F\left[g_{2,j_1}(w,\theta_2)\right]\mathbb{E}_F\left[g_{2,j_2}(w,\theta_2)\right]\right| \\
&\quad \le \left|\mathbb{E}_F\left[g_{2,j_1}(w,\theta_1) - g_{2,j_1}(w,\theta_2)\right]\mathbb{E}_F\left[g_{2,j_2}(w,\theta_1)\right]\right| \\
&\qquad + \left|\mathbb{E}_F\left[g_{2,j_1}(w,\theta_2)\right]\mathbb{E}_F\left[g_{2,j_2}(w,\theta_2) - g_{2,j_2}(w,\theta_1)\right]\right| \\
&\quad \le 2\mathbb{E}_F\left[\sup_{\theta\in\Theta}\left\|g_2(W,\theta)\right\|\right]\mathbb{E}_F\left[\sup_{\theta\in\Theta}\left\|g_{2,\theta}(W,\theta)\right\|\right]\|\theta_1-\theta_2\|
\end{aligned} \tag{E.7}$$

for any $j_1, j_2 = 1, \ldots, r_2$. By (E.7),

$$\left\|\mathbb{E}_F\left[g_2(w,\theta_1)\right]\mathbb{E}_F\left[g_2(w,\theta_1)'\right] - \mathbb{E}_F\left[g_2(w,\theta_2)\right]\mathbb{E}_F\left[g_2(w,\theta_2)'\right]\right\| \le r_2 C_{M,3}\|\theta_1-\theta_2\| \tag{E.8}$$

for any $F \in \mathcal{F}$, where $C_{M,3} \equiv 2\sup_{F\in\mathcal{F}} \mathbb{E}_F[\sup_{\theta\in\Theta}\|g_2(W,\theta)\|]\mathbb{E}_F[\sup_{\theta\in\Theta}\|g_{2,\theta}(W,\theta)\|]$ and $C_{M,3} < \infty$ by Assumption 3.2(ii).

By the definition of $\Omega_{2,F}(\theta)$, the triangle inequality and the results in (E.6) and (E.8)

$$\left\|\Omega_{2,F}(\theta_1) - \Omega_{2,F}(\theta_2)\right\| \le r_2(C_{M,2}+C_{M,3})\|\theta_1-\theta_2\|, \tag{E.9}$$

which immediately proves the claim in (iii). □

Proof of Lemma B.3. By Lemma B.1(i),

$$\overline{g}_2(\theta) = M_{2,F_n}(\theta) + \left[n^{-1}\sum_{i=1}^n g_2(W_i,\theta) - M_{2,F_n}(\theta)\right] = M_{2,F_n}(\theta) + o_p(1), \tag{E.10}$$

uniformly over $\theta \in \Theta$. As $g_1(W,\theta)$ is a subvector of $g_2(W,\theta)$, by (E.10) and Assumption 3.2(ii),

$$\overline{g}_1(\theta)'\overline{g}_1(\theta) = M_{1,F_n}(\theta)'M_{1,F_n}(\theta) + o_p(1) \tag{E.11}$$

uniformly over $\theta \in \Theta$. By Assumptions 3.1(i)–(ii) and $F_n \in \mathcal{F}$, $M_{1,F_n}(\theta)'M_{1,F_n}(\theta)$ is uniquely minimized at θ_{F_n}, which together with the uniform convergence in (E.11) implies that

$$\widetilde{\theta}_1 - \theta_{F_n} \to_p 0. \tag{E.12}$$

To show the consistency of $\overline{\Omega}_2$, note that

$$\begin{aligned}\overline{\Omega}_2 &= n^{-1}\sum_{i=1}^n g_2(W_i,\widetilde{\theta}_1)g_2(W_i,\widetilde{\theta}_1)' - \overline{g}_2(\widetilde{\theta}_1)\overline{g}_2(\widetilde{\theta}_1)' \\ &= \mathbb{E}_{F_n}\big[g_2(W,\widetilde{\theta}_1)g_2(W,\widetilde{\theta}_1)'\big] - M_{2,F_n}(\widetilde{\theta}_1)'M_{2,F_n}(\widetilde{\theta}_1) + o_p(1) \\ &= \Omega_{2,F_n}(\widetilde{\theta}_1) + o_p(1) = \Omega_{2,F_n} + o_p(1),\end{aligned} \tag{E.13}$$

where the first equality is by the definition of $\overline{\Omega}_2$, the second equality holds by (E.10), Lemma B.1(ii) and Assumption 3.2(ii), the third equality follows from the definition of $\Omega_{2,F_n}(\theta)$, and the last equality holds by Lemma B.2(iii) and (E.12). This shows the consistency of $\overline{\Omega}_2$. □

In the rest of the Supplemental Appendix, we use C denote a generic fixed positive finite constant whose value does not depend on F or n.

Proof of Lemma B.4. As $\overline{g}_1(\theta)$ is a subvector of $\overline{g}_2(\theta)$, and $\overline{\Omega}_{1,n}$ is a submatrix of $\overline{\Omega}_{2,n}$, using (E.10), (E.13), and Assumptions 3.2(ii)–(iii), we have

$$\overline{g}_1(\theta)'(\overline{\Omega}_1)^{-1}\overline{g}_1(\theta) = M_{1,F_n}(\theta)'\Omega_{1,F_n}^{-1}M_{1,F_n}(\theta) + o_p(1), \tag{E.14}$$

uniformly over Θ. By Assumptions 3.2(ii)–(iii),

$$C^{-1} \le \rho_{\min}\big(\Omega_{1,F_n}^{-1}\big) \le \rho_{\max}\big(\Omega_{1,F_n}^{-1}\big) \le C, \tag{E.15}$$

which together with Assumptions 3.1(i)–(ii) implies that $M_{1,F_n}(\theta)'\Omega_{1,F_n}^{-1}M_{1,F_n}(\theta)$ is uniquely minimized at θ_{F_n}. By the standard arguments for the consistency of an extremum estimator, we have

$$\widehat{\theta}_1 - \theta_{F_n} = o_p(1). \tag{E.16}$$

Using (E.16), Lemma B.1(iv) and Assumption 3.2(ii), we have

$$\begin{aligned}\overline{g}_1(\widehat{\theta}_1) &= \overline{g}_1(\theta_{F_n}) + \big[M_{1,F_n}(\widehat{\theta}_1) - M_{1,F_n}(\theta_{F_n})\big] + o_p\big(n^{-1/2}\big) \\ &= \overline{g}_1(\theta_{F_n}) + \big[G_{1,F_n}(\theta_{F_n}) + o_p(1)\big](\widehat{\theta}_1 - \theta_{F_n}) + o_p\big(n^{-1/2}\big).\end{aligned} \tag{E.17}$$

Similarly,

$$n^{-1}\sum_{i=1}^{n} g_{1,\theta}(W_i,\widehat{\theta}_1)=G_{1,F_n}(\widehat{\theta}_1)+o_p(1)=G_{1,F_n}+o_p(1), \tag{E.18}$$

where the first equality follows from Lemma B.1(iii) and the second equality follows by (E.16) and Lemma B.2(ii). From the first-order condition for the GMM estimator $\widehat{\theta}_1$, we deduce that

$$\begin{aligned}0&=\left[n^{-1}\sum_{i=1}^{n} g_{1,\theta}(W_i,\widehat{\theta}_1)\right]'(\overline{\Omega}_1)^{-1}\overline{g}_1(\widehat{\theta}_1)\\&=\big(G'_{1,F_n}\Omega^{-1}_{1,F_n}+o_p(1)\big)\big[\overline{g}_1(\theta_{F_n})+\big(G_{1,F_n}+o_p(1)\big)(\widehat{\theta}_1-\theta_{F_n})+o_p\big(n^{-1/2}\big)\big],\end{aligned} \tag{E.19}$$

where the second equality follows from Assumptions 3.2(ii)–(iii), (E.13), (E.17), and (E.18). By (E.19), $\mathbb{E}_{F_n}[g_1(W,\theta_{F_n})]=0$, and Assumption 3.2,

$$n^{1/2}(\widehat{\theta}_1-\theta_{F_n})=\big(\Gamma_{1,F_n}+o_p(1)\big)\mu_n\big(g_1(W,\theta_{F_n})\big)+o_p(1). \tag{E.20}$$

By Assumptions 3.2 and Lemma B.1(v), $\Gamma_{1,F_n}=O(1)$, and $\mu_n(g_1(W,\theta_{F_n}))=o_p(1)$, which together with (E.20) implies that

$$n^{1/2}(\widehat{\theta}_1-\theta_{F_n})=\Gamma_{1,F_n}\mu_n\big(g_1(W,\theta_{F_n})\big)+O_p(1),$$

where $\Gamma_{1,F_n}\mu_n(g_1(W,\theta_{F_n}))=O_p(1)$. This completes the proof. □

Proof of Lemma B.5. By (E.10), (E.13), and Assumptions 3.2(ii)–(iii), we have

$$\overline{g}_2(\theta)'(\overline{\Omega}_2)^{-1}\overline{g}_2(\theta)=M_{2,F_n}(\theta)'\Omega^{-1}_{2,F_n}M_{2,F_n}(\theta)+o_p(1)=Q_{F_n}(\theta)+o_p(1) \tag{E.21}$$

uniformly over Θ. By Assumption 3.1(iii), $Q_{F_n}(\theta)$ is uniquely minimized at $\theta^*_{F_n}$. The consistency result $\widehat{\theta}_2-\theta^*_{F_n}\to_p 0$ follows from standard arguments for the consistency of an extremum estimator. □

Proof of Lemma B.6. By the definition of $\widehat{\theta}_2$,

$$\overline{g}_2(\widehat{\theta}_2)'(\overline{\Omega}_2)^{-1}\overline{g}_2(\widehat{\theta}_2)\le\overline{g}_2(\theta_{F_n})'(\overline{\Omega}_2)^{-1}\overline{g}_2(\theta_{F_n}), \tag{E.22}$$

which implies that

$$\|\overline{g}_2(\widehat{\theta}_2)\|^2\le\rho_{\max}(\overline{\Omega}_2)\rho^{-1}_{\min}(\overline{\Omega}_2)\|\overline{g}_2(\theta_{F_n})\|^2. \tag{E.23}$$

By (E.13) and Assumptions 3.2(ii)–(iii),

$$C^{-1}\le\rho_{\min}(\overline{\Omega}_2)\le\rho_{\max}(\overline{\Omega}_2)\le C \tag{E.24}$$

with probability approaching 1. By Lemma B.1(i), $M_{1,F_n}(\theta_{F_n})=0_{r_1\times 1}$ and $\delta_{F_n}=o(1)$,

$$\|\overline{g}_2(\theta_{F_n})\|^2=o_p(1), \tag{E.25}$$

which combined with (E.23) and (E.24) implies that

$$\|\overline{g}_2(\widehat{\theta}_2)\| = o_p(1). \tag{E.26}$$

Moreover, by (E.26), Lemma B.1(i) and the triangle inequality,

$$\|M_{2,F_n}(\widehat{\theta}_2)\| \leq \|\overline{g}_2(\widehat{\theta}_2) - M_{2,F_n}(\widehat{\theta}_2)\| + \|\overline{g}_2(\widehat{\theta}_2)\| = o_p(1), \tag{E.27}$$

which immediately implies that

$$\|M_{1,F_n}(\widehat{\theta}_2)\| = o_p(1). \tag{E.28}$$

The first result in Lemma B.6 follows by (E.28) and the unique identification of θ_{F_n} maintained by Assumptions 3.1(i)–(ii).

Using $\widehat{\theta}_2 - \theta_{F_n} = o_p(1)$, Lemma B.1(iv) and Assumption 3.2(ii), we have

$$\begin{aligned}\overline{g}_2(\widehat{\theta}_2) &= \overline{g}_2(\theta_{F_n}) + \big[M_{2,F_n}(\widehat{\theta}_2) - M_{2,F_n}(\theta_{F_n})\big] + o_p\big(n^{-1/2}\big) \\ &= \overline{g}_2(\theta_{F_n}) + \big[G_{2,F_n}(\theta_{F_n}) + o_p(1)\big](\widehat{\theta}_2 - \theta_{F_n}) + o_p\big(n^{-1/2}\big). \end{aligned} \tag{E.29}$$

Similarly,

$$n^{-1}\sum_{i=1}^{n} g_{2,\theta}(W_i, \widehat{\theta}_2) = G_{2,F_n}(\widehat{\theta}_2) + o_p(1) = G_{2,F_n}(\theta_{F_n}) + o_p(1), \tag{E.30}$$

where the first equality follows from Lemma B.1(iii) and the second equality follows by $\widehat{\theta}_2 - \theta_{F_n} = o_p(1)$ and Lemma B.2(ii). From the first-order condition for the GMM estimator $\widehat{\theta}_2$, we deduce that

$$\begin{aligned} 0 &= \left[n^{-1}\sum_{i=1}^{n} g_{2,\theta}(W_i, \widehat{\theta}_2)\right]' (\overline{\Omega}_2)^{-1}\overline{g}_2(\widehat{\theta}_2) \\ &= \big(G'_{2,F_n}\Omega_{2,F_n}^{-1} + o_p(1)\big)\big[\overline{g}_2(\theta_{F_n}) + \big(G_{2,F_n} + o_p(1)\big)(\widehat{\theta}_2 - \theta_{F_n}) + o_p\big(n^{-1/2}\big)\big], \end{aligned} \tag{E.31}$$

where the second equality follows from Assumptions 3.2(ii)–(iii), (E.13), (E.29), and (E.30). By (E.31) and Assumption 3.2,

$$n^{1/2}(\widehat{\theta}_2 - \theta_{F_n}) = \big(\Gamma_{2,F_n} + o_p(1)\big)\big\{\mu_n\big(g_2(W, \theta_{F_n})\big) + n^{1/2}\mathbb{E}_{F_n}\big[g_2(W, \theta_{F_n})\big]\big\} + o_p(1), \tag{E.32}$$

where $\Gamma_{2,F_n} = -(G'_{2,F_n}\Omega_{2,F_n}^{-1}G_{2,F_n})^{-1}G'_{2,F_n}\Omega_{2,F_n}^{-1}$. □

Proof for the claim in equation (4.3). Consider the case $n^{1/2}\delta_{F_n} \to d \in \mathbb{R}^{r^*}$. By Lemma 4.1,

$$\begin{aligned} n^{1/2}\big[\widehat{\theta}(\omega) - \theta_{F_n}\big] &= n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) + \omega\big[n^{1/2}(\widehat{\theta}_2 - \theta_{F_n}) - n^{1/2}(\widehat{\theta}_1 - \theta_{F_n})\big] \\ &\to_D \Gamma^*_{1,F}\mathcal{Z}_{d,2,F} + \omega\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)\mathcal{Z}_{d,2,F}, \end{aligned} \tag{E.33}$$

where $\mathcal{Z}_{d,2,F}$ has the same distribution as $\mathcal{Z}_{2,F} + d_0$. This implies that

$$\ell\big(\widehat{\theta}(\omega)\big) = n\big[\widehat{\theta}_n(\omega) - \theta_{F_n}\big]' \Upsilon \big[\widehat{\theta}_n(\omega) - \theta_{F_n}\big] \to_D \lambda_F(\omega), \tag{E.34}$$

where

$$\lambda_F(\omega) = \mathcal{Z}'_{d,2,F}\Gamma^{*\prime}_{1,F}\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F} + 2\omega\mathcal{Z}'_{d,2,F}\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F} \\ + \omega^2\mathcal{Z}'_{d,2,F}\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\Upsilon\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)\mathcal{Z}_{d,2,F}.$$

Now we consider $\mathbb{E}[\lambda_F(\omega)]$ using the equalities in Lemma B.9 below. First,

$$\mathbb{E}\big[\mathcal{Z}'_{d,2,F}\Gamma^{*\prime}_{1,F}\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F}\big] = \mathrm{tr}(\Upsilon\Sigma_{1,F}) \tag{E.35}$$

because $\Gamma^*_1\mathcal{Z}_{d,2,F} = \Gamma_{1,F}\mathcal{Z}_{1,F}$ and $\Gamma_{1,F}\mathbb{E}[\mathcal{Z}_{1,F}\mathcal{Z}'_{1,F}]\Gamma'_{1,F} = \Sigma_{1,F}$ by definition. Second,

$$\begin{aligned} &\mathbb{E}\big[\mathcal{Z}'_{d,2,F}\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F}\big] \\ &= \mathrm{tr}\big(\Upsilon\Gamma^*_{1,F}\mathbb{E}\big[\mathcal{Z}_{d,2,F}\mathcal{Z}'_{d,2,F}\big]\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\big) \\ &= \mathrm{tr}\big(\Upsilon\Gamma^*_{1,F}\big[d_0d_0' + \Omega_{2,F}\big]\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\big) \\ &= \mathrm{tr}\big(\Upsilon(\Sigma_{2,F} - \Sigma_{1,F})\big), \end{aligned} \tag{E.36}$$

where the last equality holds by Lemma B.9. Third,

$$\begin{aligned} &\mathbb{E}\big[\mathcal{Z}'_{d,2,F}\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\Upsilon\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)\mathcal{Z}_{d,2,F}\big] \\ &= \mathrm{tr}\big(\Upsilon\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)\big[d_0d_0' + \Omega_{2,F}\big]\big(\Gamma_{2,F} - \Gamma^*_{1,F}\big)'\big) \\ &= d_0'\Gamma'_{2,F}\Upsilon\Gamma_{2,F}d_0 - \mathrm{tr}\big(\Upsilon(\Sigma_{2,F} - \Sigma_{1,F})\big) \end{aligned} \tag{E.37}$$

by Lemma B.9. Combining the results in (E.35)–(E.37), we obtain

$$\mathbb{E}\big[\lambda_F(\omega)\big] = \mathrm{tr}(\Upsilon\Sigma_{1,F}) - 2\omega\,\mathrm{tr}\big(\Upsilon(\Sigma_{1,F} - \Sigma_{2,F})\big) \\ + \omega^2\big[d_0'\Gamma'_{2,F}\Upsilon\Gamma_{2,F}d_0 + \mathrm{tr}\big(\Upsilon(\Sigma_{1,F} - \Sigma_{2,F})\big)\big]. \tag{E.38}$$

Note that $d_0'\Gamma'_{2,F}\Upsilon\Gamma_{2,F}d_0 = d_0'(\Gamma_{2,F} - \Gamma^*_{1,F})'\Upsilon(\Gamma_{2,F} - \Gamma^*_{1,F})d_0$ because $\Gamma^*_{1,F}d_0 = \mathbf{0}_{d_\theta}$. It is clear that the optimal weight ω^*_F in (4.3) minimizes the quadratic function of ω in (E.38). □

Proof of Lemma B.9. By construction, $\Gamma^*_{1,F}d_0 = \mathbf{0}_{d_\theta \times 1}$. For ease of notation, we write $\Omega_{2,F}$ and $G_{2,F}$ as

$$\Omega_{2,F} = \begin{pmatrix} \Omega_{1,F} & \Omega_{1r^*} \\ \Omega_{r^*1,F} & \Omega_{r^*,F} \end{pmatrix} \quad \text{and} \quad G_{2,F} = \begin{pmatrix} G_{1,F} \\ G_{r^*,F} \end{pmatrix}. \tag{E.39}$$

To prove part (b), we have

$$\Gamma^*_{1,F}\Omega_{2,F}\Gamma^*_{1,F} = [\Gamma_{1,F}, \mathbf{0}_{d_\theta \times r^*}]\begin{pmatrix} \Omega_{1,F} & \Omega_{1r^*} \\ \Omega_{r^*1,F} & \Omega_{r^*,F} \end{pmatrix}[\Gamma_{1,F}, \mathbf{0}_{d_\theta \times r^*}]$$

$$= \Gamma_{1,F}\Omega_{1,F}\Gamma'_{1,F} = \left(G'_{1,F}\Omega^{-1}_{1,F}G_{1,F}\right)^{-1} = \Sigma_{1,F}. \tag{E.40}$$

To show part (c), note that

$$\begin{aligned}\Gamma^*_{1,F}\Omega_{2,F}\Gamma'_{2,F} &= -[\Gamma_{1,F}, \mathbf{0}_{d_\theta\times r^*}]\Omega_{2,F}\Omega^{-1}_{2,F}G_{2,F}\left(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F}\right)^{-1} \\ &= -\Gamma_{1,F}G_{1,F}\left(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F}\right)^{-1} = \left(G'_{2,F}\Omega^{-1}_{2,F}G_{2,F}\right)^{-1} = \Sigma_{2,F}\end{aligned} \tag{E.41}$$

because $-\Gamma_{1,F}G_{1,F} = I_{d_\theta\times d_\theta}$. Part (d) follows from the definition of $\Gamma_{2,F}$. □

Proof of Lemma 4.2. We first prove the consistency of $\widehat{\Omega}_k$, $\widehat{G}_k$, and $\widehat{\Sigma}_k$ for $k = 1, 2$. By Lemma 4.1, we have $\widehat{\theta}_1 = \theta_{F_n} + o_p(1)$. Using the same arguments in showing (E.13), we can show that

$$\widehat{\Omega}_2 = \Omega_{2,F_n} + o_p(1) = \Omega_{2,F} + o_p(1), \tag{E.42}$$

where the second equality is by the assumption of the lemma that $v_{F_n} \to v_F$ for some $F \in \mathcal{F}$. As $\widehat{\Omega}_1$ is a submatrix of $\widehat{\Omega}_2$, by (E.42) we have

$$\widehat{\Omega}_1 = \Omega_{1,F_n} + o_p(1) = \Omega_{1,F} + o_p(1). \tag{E.43}$$

By the consistency of $\widehat{\theta}_1$ and the same arguments used to show (E.30), we have

$$n^{-1}\sum_{i=1}^{n} g_{2,\theta}(W_i, \widehat{\theta}_1) = G_{2,F_n}(\theta_{F_n}) + o_p(1) = G_{2,F} + o_p(1), \tag{E.44}$$

where the second equality is by (B.10) which is assumed in the lemma. As $n^{-1} \times \sum_{i=1}^{n} g_{1,\theta}(W_i, \widehat{\theta}_1)$ is a submatrix of $n^{-1}\sum_{i=1}^{n} g_{2,\theta}(W_i, \widehat{\theta}_1)$, by (E.44) we have

$$n^{-1}\sum_{i=1}^{n} g_{1,\theta}(W_i, \widehat{\theta}_1) = G_{1,F_n}(\theta_{F_n}) + o_p(1) = G_{1,F} + o_p(1). \tag{E.45}$$

From Assumption 3.2, (E.42), (E.43), (E.44), and (E.45), we see that $\widehat{\Omega}_k$ and $\widehat{G}_k$ are consistent estimators of $\Omega_{k,F}$ and $G_{k,F}$, respectively, for $k = 1, 2$. By the Slutsky theorem and Assumption 3.2, we know that $\widehat{\Sigma}_k$ is a consistent estimator of $\Sigma_{k,F}$ for $k = 1, 2$.

In the case where $n^{1/2}\delta_{F_n} \to d \in \mathbb{R}^{r^*}$, the desired result follows from Lemma 4.1, the consistency of $\widehat{\Sigma}_{1,F}$ and $\widehat{\Sigma}_{2,F}$, and the CMT. In the case where $\|n^{1/2}\delta_{F_n}\| \to \infty$, $\widetilde{\omega}_{\text{eo}} \to_p 0$ because $n^{1/2}\|\widehat{\theta}_2 - \widehat{\theta}_1\| \to_p \infty$ and

$$\begin{aligned}n^{1/2}(\widehat{\theta}_{\text{eo}} - \theta_{F_n}) &= n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) + \widetilde{\omega}_{\text{eo}} n^{1/2}(\widehat{\theta}_2 - \widehat{\theta}_1) \\ &= n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) + \frac{n^{1/2}(\widehat{\theta}_2 - \widehat{\theta}_1)\operatorname{tr}\left[\Upsilon(\widehat{\Sigma}_1 - \widehat{\Sigma}_2)\right]}{n(\widehat{\theta}_2 - \widehat{\theta}_1)'\Upsilon(\widehat{\theta}_2 - \widehat{\theta}_1) + \operatorname{tr}\left[\Upsilon(\widehat{\Sigma}_1 - \widehat{\Sigma}_2)\right]} \\ &\to_D \xi_{1,F}\end{aligned} \tag{E.46}$$

by Lemma 4.1. □

PROOF OF LEMMA B.15. By definition,

$$\xi_{1,F}'\Upsilon\xi_{1,F} = \mathcal{Z}_{1,F}'\Gamma_{1,F}'\Upsilon\Gamma_{1,F}\mathcal{Z}_{1,F} = \mathcal{Z}_1'\Omega_{1,F}^{1/2}\Gamma_{1,F}'\Upsilon\Gamma_{1,F}\Omega_{1,F}^{1/2}\mathcal{Z}_1, \tag{E.47}$$

where $\mathcal{Z}_1 \sim N(\mathbf{0}_{r_1}, I_{r_1\times r_1})$. By Assumptions 3.2(ii) and 3.2(iv), and the fact that Υ is a fixed matrix,

$$\sup_{F\in\mathcal{F}} \rho_{\max}\big(\Omega_{1,F}^{1/2}\Gamma_{1,F}'\Upsilon\Gamma_{1,F}\Omega_{1,F}^{1/2}\big) \le C. \tag{E.48}$$

By (E.48),

$$\sup_{h\in H}\mathbb{E}\big[\big(\xi_{1,F}'\Upsilon\xi_{1,F}\big)^2\big] \le \sup_{h\in H}\rho_{\max}^2\big(\Omega_{1,F}^{1/2}\Gamma_{1,F}'\Upsilon\Gamma_{1,F}\Omega_{1,F}^{1/2}\big)\mathbb{E}\big[\big(\mathcal{Z}_1'\mathcal{Z}_1\big)^2\big] \le 3r_1C, \tag{E.49}$$

where the second inequality is by $\mathbb{E}[(\mathcal{Z}_1'\mathcal{Z}_1)^2] \le 3r_1 + r_1(r_1-1) = r_1^2 + 2r_1$ which is implied by the assumption that $\mathcal{Z}_1$ is a r_1-dimensional standard normal random vector. The first inequality of this lemma follows as the upper bound in (E.49) does not depend on F.

For any $F \in \mathcal{F}$, define

$$B_F \equiv \big(\Gamma_{2,F} - \Gamma_{1,F}^*\big)'\Upsilon\big(\Gamma_{2,F} - \Gamma_{1,F}^*\big).$$

By the Cauchy–Schwarz inequality and the simple inequality $|ab| \le (a^2+b^2)/2$ (for any real numbers a and b),

$$\begin{aligned}\overline{\xi}_F'\Upsilon\overline{\xi}_F &\le 2\big(\mathcal{Z}_{d,2,F}'\Gamma_{1,F}^{*\prime}\Upsilon\Gamma_{1,F}^*\mathcal{Z}_{d,2,F} + \overline{\omega}_F^2\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}\big)\\ &= 2\big(\mathcal{Z}_{1,F}'\Gamma_{1,F}'\Upsilon\Gamma_{1,F}\mathcal{Z}_{1,F} + \overline{\omega}_F^2\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}\big),\end{aligned} \tag{E.50}$$

where the equality is by $\Gamma_{1,F}^*d_0 = \mathbf{0}_{d_\theta\times 1}$ (which is proved in Lemma B.9). By (E.50) and the simple inequality $(a+b)^2 \le 2(a^2+b^2)$ (for any real numbers a and b),

$$\big(\overline{\xi}_F'\Upsilon\overline{\xi}_F\big)^2 \le 8\big(\mathcal{Z}_{1,F}'\Gamma_{1,F}'\Upsilon\Gamma_{1,F}\mathcal{Z}_{1,F}\big)^2 + 8\big(\overline{\omega}_F^2\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}\big)^2. \tag{E.51}$$

By the first inequality of this lemma, we have $\sup_{h\in H}\mathbb{E}[(\xi_{1,F}'\Upsilon\xi_{1,F})^2] \le C$. Hence by (E.51), to show the second inequality of this lemma, it is sufficient to prove that

$$\sup_{h\in H}\mathbb{E}\big[\big(\overline{\omega}_F^2\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}\big)^2\big] \le C. \tag{E.52}$$

Recall that we have defined $A_F = \Upsilon(\Sigma_{1,F} - \Sigma_{2,F})$ in Theorem 5.2. By the definition,

$$\begin{aligned}&\overline{\omega}_F^2\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}\\ &= \frac{\big(\mathrm{tr}(A_F)\big)^2\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}}{\big(\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F} + \mathrm{tr}(A_F)\big)^2}\\ &= \mathrm{tr}(A_F)\frac{\mathrm{tr}(A_F)}{\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F} + \mathrm{tr}(A_F)}\frac{\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F}}{\mathcal{Z}_{d,2,F}'B_F\mathcal{Z}_{d,2,F} + \mathrm{tr}(A_F)}.\end{aligned} \tag{E.53}$$

By Lemma 2.1 in Cheng and Liao (2015), $\mathrm{tr}(A_F) \geq 0$ for any $F \in \mathcal{F}$. This together with $\mathcal{Z}'_{d,2,F} B_F \mathcal{Z}_{d,2,F} \geq 0$ implies that

$$\frac{\mathrm{tr}(A_F)}{\mathcal{Z}'_{d,2,F} B_F \mathcal{Z}_{d,2,F} + \mathrm{tr}(A_F)} \leq 1 \quad \text{and} \quad \frac{\mathcal{Z}'_{d,2,F} B_F \mathcal{Z}_{d,2,F}}{\mathcal{Z}'_{d,2,F} B_F \mathcal{Z}_{d,2,F} + \mathrm{tr}(A_F)} \leq 1. \tag{E.54}$$

By (E.54) and $\mathrm{tr}(A_F) \geq 0$,

$$\overline{\omega}_F^2 \mathcal{Z}'_{d,2,F} B_F \mathcal{Z}_{d,2,F} \leq \mathrm{tr}(A_F) = \mathrm{tr}(\Upsilon \Sigma_{1,F}) - \mathrm{tr}(\Upsilon \Sigma_{2,F}), \tag{E.55}$$

where the equality is by $A_F = \Upsilon(\Sigma_{1,F} - \Sigma_{2,F})$. By (E.55) and the simple inequality $(a+b)^2 \leq 2(a^2+b^2)$,

$$\mathbb{E}\left[\left(\overline{\omega}_F^2 \mathcal{Z}'_{d,2,F} B_F \mathcal{Z}_{d,2,F}\right)^2\right] \leq 2\left(\mathrm{tr}(\Upsilon \Sigma_{1,F})\right)^2 + 2\left(\mathrm{tr}(\Upsilon \Sigma_{2,F})\right)^2. \tag{E.56}$$

By Assumptions 3.2(ii) and 3.2(iv),

$$\begin{aligned} \rho_{\min}\left(G'_{k,F} \Omega_{k,F}^{-1} G_{k,F}\right) &\geq \rho_{\min}\left(\Omega_{k,F}^{-1}\right) \rho_{\min}\left(G'_{k,F} G_{k,F}\right) \\ &= \rho_{\min}\left(G'_{k,F} G_{k,F}\right) / \rho_{\max}(\Omega_{k,F}) \geq C^{-1} \end{aligned} \tag{E.57}$$

for any $F \in \mathcal{F}$ and for $k = 1, 2$. By (E.57) and the definition of $\Sigma_{k,F}$ $(k = 1, 2)$,

$$\rho_{\max}(\Sigma_{k,F}) = \rho_{\min}^{-1}\left(G'_{k,F} \Omega_{k,F}^{-1} G_{k,F}\right) \leq C \tag{E.58}$$

for any $F \in \mathcal{F}$. As Υ and $\Sigma_{k,F}$ are positive definite symmetric matrix, by the standard trace inequality ($\mathrm{tr}(AB) \leq \mathrm{tr}(A)\rho_{\max}(B)$ for Hermitian matrices $A \geq 0$ and $B \geq 0$),

$$\mathrm{tr}(\Upsilon \Sigma_{k,F}) \leq \mathrm{tr}(\Upsilon)\rho_{\max}(\Sigma_{k,F}) \leq C \quad \text{for } k = 1, 2, \tag{E.59}$$

for any $F \in \mathcal{F}$. Collecting the results in (E.56) and (E.59), we immediately get (E.52). This completes the proof. □

Proof of Lemma B.16. First, note that

$$\min\{x, \zeta\} - x = (\zeta - x) I\{x > \zeta\}. \tag{E.60}$$

Hence we have

$$\begin{aligned} &\sup_{h \in H} \left| \mathbb{E}\left[\min\{\overline{\xi}'_F \Upsilon \overline{\xi}_F, \zeta\} - \overline{\xi}'_F \Upsilon \overline{\xi}_F\right] \right| \\ &\leq \sup_{h \in H} \mathbb{E}\left[\left|\zeta - \overline{\xi}'_F \Upsilon \overline{\xi}_F\right| I\{\overline{\xi}'_F \Upsilon \overline{\xi}_F > \zeta\}\right] \\ &\leq \zeta \sup_{h \in H} \mathbb{E}\left[I\{\overline{\xi}'_F \Upsilon \overline{\xi}_F > \zeta\}\right] + \sup_{h \in H} \mathbb{E}\left[\overline{\xi}'_F \Upsilon \overline{\xi}_F I\{\zeta^{-1} > \left(\overline{\xi}'_F \Upsilon \overline{\xi}_F\right)^{-1}\}\right] \\ &\leq 2\zeta^{-1} \sup_{h \in H} \mathbb{E}\left[\left(\overline{\xi}'_F \Upsilon \overline{\xi}_F\right)^2\right] \leq 2C\zeta^{-1}, \end{aligned} \tag{E.61}$$

where the first inequality is by the Jensen's inequality, the second inequality is by the Markov inequality, the third inequality is by the monotonicity of expectation and the last inequality is by Lemma B.15. Using the same arguments, we can show that

$$\sup_{h\in H}\left|\mathbb{E}\left[\min\{\xi_{1,F}'Y\xi_{1,F},\zeta\}-\xi_{1,F}'Y\xi_{1,F}\right]\right|\leq 2C\zeta^{-1}. \tag{E.62}$$

Collecting the results in (E.61) and (E.62), and applying the triangle inequality, we deduce that

$$\sup_{h\in H}\left[\left|g_{\zeta}(h)-g(h)\right|\right]\leq 4C\zeta^{-1}. \tag{E.63}$$

The claimed result of this lemma follows by (E.63) as C is a fixed constant.

By the triangle inequality, the Jensen's inequality and Lemma B.15,

$$\begin{aligned}\sup_{h\in H}\left|g(h)\right| &= \sup_{h\in H}\left|\mathbb{E}\left[\overline{\xi}_F'Y\overline{\xi}_F-\xi_{1,F}'Y\xi_{1,F}\right]\right| \\ &\leq \sup_{h\in H}\mathbb{E}\left[\overline{\xi}_F'Y\overline{\xi}_F\right]+\sup_{h\in H}\mathbb{E}\left[\xi_{1,F}'Y\xi_{1,F}\right]\leq C,\end{aligned}$$

which completes the proof of the lemma. □

Proof of Lemma A.1. By definition,

$$\begin{aligned}&\mathbb{E}\left[\|\widehat{\theta}_{\mathrm{eo}}-\theta\|^2\right]-\mathbb{E}\left[\|\widehat{\theta}_1-\theta\|^2\right] \\ &= \mathbb{E}\left[\frac{k^2\sigma^4(Y-X)'(Y-X)}{\left(2k\sigma^2+(Y-X)'(Y-X)\right)^2}\right] \\ &\quad+\mathbb{E}\left[\frac{2k\sigma^2(X-\theta)'(Y-X)}{2k\sigma^2+(Y-X)'(Y-X)}\right]\end{aligned} \tag{E.64}$$

Let

$$\begin{aligned}J_1&\equiv\mathbb{E}\left[\frac{(X-\theta)'(Y-X)}{2k\sigma^2+(Y-X)'(Y-X)}\right] \quad\text{and} \\ J_2&\equiv\mathbb{E}\left[\frac{(Y-X)'(Y-X)}{\left(2k\sigma^2+(Y-X)'(Y-X)\right)^2}\right].\end{aligned} \tag{E.65}$$

Let $X^*=\sigma^{-1}(X-\theta)$, $Y^*=\sigma^{-1}(Y-\theta)$ and $Z^*=(X^{*\prime},Y^{*\prime})'$. Then we can write

$$\begin{aligned}J_1&=\mathbb{E}\left[\frac{(X-\theta)'(Y-X)}{2k\sigma^2+(Y-X)'(Y-X)}\right] \\ &=\mathbb{E}\left[\frac{X^{*\prime}(Y^*-X^*)}{2k+\left(Y^*-X^*\right)'\left(Y^*-X^*\right)}\right]=\mathbb{E}\left[\frac{Z^{*\prime}D_1Z^*}{2k+Z^{*\prime}D_2Z^*}\right],\end{aligned} \tag{E.66}$$

where

$$D_1=\begin{pmatrix}-I_k & 0_k\\ I_k & 0_k\end{pmatrix}\quad\text{and}\quad D_2=\begin{pmatrix}I_k & -I_k\\ -I_k & I_k\end{pmatrix}. \tag{E.67}$$

Note that

$$\mathbb{E}\left[D_1 Z^* Z^{*\prime} D_1'\right] = D_2 \tag{E.68}$$

by definition and the Gaussian assumption. Let $\eta(x) = x/(x'D_2x + 2k)$. Its derivative is

$$\frac{\partial \eta(x)'}{\partial x} = \frac{1}{x'D_2x + 2k} I_k - \frac{2}{\left(x'D_2x + 2k\right)^2} D_2 x x'. \tag{E.69}$$

By Lemma 1 of Hansen (2016), which is a matrix version of the Stein's lemma (Stein (1981)),

$$\begin{aligned}
J_1 = \mathbb{E}\left(\eta\left(Z^*\right)' D_1 Z^*\right) &= \mathbb{E}\left[\operatorname{tr}\left(\frac{\partial \eta\left(Z^*\right)'}{\partial x} D_1\right)\right] \\
&= \mathbb{E}\left[\frac{\operatorname{tr}(D_1)}{2k + Z^{*\prime} D_2 Z^*}\right] - 2\mathbb{E}\left[\frac{\operatorname{tr}\left(D_2 Z^* Z^{*\prime} D_1\right)}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right] \\
&= \mathbb{E}\left[\frac{-k}{2k + Z^{*\prime} D_2 Z^*}\right] - 2\mathbb{E}\left[\frac{Z^{*\prime} D_1 D_2 Z^*}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right] \\
&= \mathbb{E}\left[\frac{-k}{2k + Z^{*\prime} D_2 Z^*}\right] + 2\mathbb{E}\left[\frac{Z^{*\prime} D_2 Z^*}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right] \\
&= \mathbb{E}\left[\frac{2-k}{2k + Z^{*\prime} D_2 Z^*}\right] + \mathbb{E}\left[\frac{-4k}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right],
\end{aligned} \tag{E.70}$$

where the fourth equality follows from

$$D_1 D_2 = \begin{pmatrix} -I_k & I_k \\ I_k & -I_k \end{pmatrix} = -D_2. \tag{E.71}$$

Moreover,

$$\begin{aligned}
k^2\sigma^4 J_2 &= \mathbb{E}\left[\frac{k^2\sigma^4 (Y-X)'(Y-X)}{\left(2k\sigma^2 + (Y-X)'(Y-X)\right)^2}\right] \\
&= \mathbb{E}\left[\frac{k^2\sigma^2}{2k + Z^{*\prime} D_2 Z^*}\right] - \mathbb{E}\left[\frac{2k^3\sigma^2}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right],
\end{aligned} \tag{E.72}$$

which together with (E.70) implies that

$$\begin{aligned}
&\mathbb{E}\left[\|\widehat{\theta}_{eo} - \theta\|^2\right] - \mathbb{E}\left[\|\widehat{\theta}_1 - \theta\|^2\right] \\
&\quad = \sigma^2 \mathbb{E}\left[\frac{2k(2-k) + k^2}{2k + Z^{*\prime} D_2 Z^*}\right] - \sigma^2 \mathbb{E}\left[\frac{2k^3 + 8k^2}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right] \\
&\quad = \sigma^2 \mathbb{E}\left[\frac{k(4-k)}{2k + Z^{*\prime} D_2 Z^*}\right] - \sigma^2 \mathbb{E}\left[\frac{2k^2(k+4)}{\left(2k + Z^{*\prime} D_2 Z^*\right)^2}\right].
\end{aligned} \tag{E.73}$$

The asserted result follows from the fact that D_2 is positive semidefinite and the second term on the right-hand side of the second equality of (E.73) is always negative. □

Appendix F: Asymptotic risk of the pre-test GMM estimator

In this section, we establish similar results in Theorem 5.1 for the pre-test GMM estimator based on the J-test statistic. The pre-test estimator is defined as

$$\widehat{\theta}_{\text{pre}} = 1\{J_n > c_\alpha\}\widehat{\theta}_1 + 1\{J_n \leq c_\alpha\}\widehat{\theta}_2, \tag{F.1}$$

where $J_n = n\overline{g}_2(\widehat{\theta}_2)'(\widehat{\Omega}_2)^{-1}\overline{g}_2(\widehat{\theta}_2)$ and c_α is the $100(1-\alpha)$th quantile of the chi-squared distribution with degree of freedom $r_2 - d_\theta$.

Theorem F.1. *Suppose that Assumptions* 3.1–3.3 *hold. The bounds of the asymptotic risk difference satisfy*

$$\text{Asy}\underline{RD}(\widehat{\theta}_{\text{pre}}, \widehat{\theta}_1) = \min\Big\{\inf_{h\in H}[g_p(h)], 0\Big\},$$

$$\text{Asy}\overline{RD}(\widehat{\theta}_{\text{pre}}, \widehat{\theta}_1) = \max\Big\{\sup_{h\in H}[g_p(h)], 0\Big\},$$

where $g_p(h) \equiv \mathbb{E}[\overline{\xi}'_{p,F}\Upsilon\overline{\xi}_{p,F} - \xi'_{1,F}\Upsilon\xi_{1,F}]$ *and* $\overline{\xi}_{p,F}$ *is defined in* (F.3) *below.*

Proof of Theorem F.1. The two equalities and inequalities in the theorem follow by the same arguments in the proof of Theorem 5.1 with Lemma 4.2 for $\widehat{\theta}_{\text{eo}}$ replaced by Lemma F.1 for $\widehat{\theta}_{\text{pre}}$, Lemma B.15 replaced by Lemma F.2, and Lemma B.16 replaced by Lemma F.3. Its proof is hence omitted. □

By definition,

$$\begin{aligned}\mathbb{E}[\overline{\xi}'_{p,F}\Upsilon\overline{\xi}_{p,F}] &= \mathbb{E}[\mathcal{Z}'_{d,2,F}\Gamma^{*\prime}_{1,F}\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F}] + 2\mathbb{E}[\overline{\omega}_{p,F}\mathcal{Z}'_{d,2,F}(\Gamma_{2,F} - \Gamma^*_{1,F})'\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F}] \\ &\quad + \mathbb{E}[\overline{\omega}^2_{p,F}\mathcal{Z}'_{d,2,F}(\Gamma_{2,F} - \Gamma^*_{1,F})'\Upsilon(\Gamma_{2,F} - \Gamma^*_{1,F})\mathcal{Z}_{d,2,F}] \\ &= \text{tr}(\Upsilon\Sigma_{1,F}) + 2\mathbb{E}[\overline{\omega}_{p,F}\mathcal{Z}'_{d,2,F}(\Gamma_{2,F} - \Gamma^*_{1,F})'\Upsilon\Gamma^*_{1,F}\mathcal{Z}_{d,2,F}] \\ &\quad + \mathbb{E}[\overline{\omega}^2_{p,F}\mathcal{Z}'_{d,2,F}(\Gamma_{2,F} - \Gamma^*_{1,F})'\Upsilon(\Gamma_{2,F} - \Gamma^*_{1,F})\mathcal{Z}_{d,2,F}].\end{aligned} \tag{F.2}$$

The asymptotic risk of the pre-test estimator $\widehat{\theta}_p$ in Figure 2 is simulated based on the formula in (F.2).

The following lemma provides the asymptotic distribution of the pre-test GMM estimator under various sequence of DGPs, which is used to show Theorem F.1.

Lemma F.1. *Suppose that Assumptions* 3.1–3.3 *hold. Consider* $\{F_n\}$ *such that* $v_{F_n} \to v_F$ *for some* $F \in \mathcal{F}$.

(a) *If* $n^{1/2}\delta_{F_n} \to d$ *for some* $d \in \mathbb{R}^{r^*}$, *then*

$$J_n \to_D J_\infty(h_{d,F}) \equiv (\mathcal{Z}_{2,F} + d_0)'L_F(\mathcal{Z}_{2,F} + d_0),$$

where $L_F \equiv \Omega_{2,F}^{-1} - \Omega_{2,F}^{-1} G_{2,F}(G_{2,F}' \Omega_{2,F}^{-1} G_{2,F})^{-1} G_{2,F}' \Omega_{2,F}^{-1}$ *and* $d_0 = (0_{1\times r_1}, d')'$, *and*

$$n^{1/2}(\widehat{\theta}_{\text{pre}} - \theta_{F_n}) \to_D \overline{\xi}_{p,F} \equiv (1 - \overline{\omega}_{p,F})\xi_{1,F} + \overline{\omega}_{p,F}\xi_{2,F}, \tag{F.3}$$

where $\overline{\omega}_{p,F} = 1\{J_\infty(h_{d,F}) \le c_\alpha\}$.

(b) *If* $\|n^{1/2}\delta_{F_n}\| \to \infty$, *then* $\overline{\omega}_{p,F} \to_p 0$ *and* $n^{1/2}(\widehat{\theta}_{\text{pre}} - \theta_{F_n}) \to_D \xi_{1,F}$.

Proof of Lemma F.1. (a) By Assumption 3.2(ii), (E.29), and (E.32),

$$\begin{aligned} \overline{g}_2(\widehat{\theta}_2) &= \overline{g}_2(\theta_{F_n}) + \left[G_{2,F_n}(\theta_{F_n}) + o_p(1)\right](\widehat{\theta}_2 - \theta_{F_n}) + o_p\left(n^{-1/2}\right) \\ &= \overline{g}_2(\theta_{F_n}) + G_{2,F_n}\Gamma_{2,F_n}\overline{g}_2(\theta_{F_n}) + o_p\left(n^{-1/2}\right) \\ &= (I_{r_2} + G_{2,F_n}\Gamma_{2,F_n})\overline{g}_2(\theta_{F_n}) + o_p\left(n^{-1/2}\right), \end{aligned} \tag{F.4}$$

which implies that

$$J_n = n\overline{g}_2(\theta_{F_n})' L_{F_n} \overline{g}_2(\theta_{F_n}) + o_p(1), \tag{F.5}$$

where $L_{F_n} \equiv \Omega_{2,F_n}^{-1} - \Omega_{2,F_n}^{-1} G_{2,F_n}(G_{2,F_n}' \Omega_{2,F_n}^{-1} G_{2,F_n})^{-1} G_{2,F_n}' \Omega_{2,F_n}^{-1}$.

By $n^{1/2}\delta_{F_n} \to d$ and Lemma B.1(v),

$$n^{1/2}\Omega_{2,F_n}^{-1/2}\overline{g}_2(\theta_{F_n}) = \Omega_{2,F_n}^{-1/2}\mu_n\left(g_2(W,\theta_{F_n})\right) + \Omega_{2,F_n}^{-1/2} n^{1/2}\delta_{F_n} \to_D \mathcal{Z} + \Omega_{2,F}^{-1/2} d_0, \tag{F.6}$$

where $d_0' = (0_{1\times r_1}, d')$ and $\mathcal{Z}$ is a $r_2 \times 1$ standard normal random vector. By $v_{F_n} \to v_F$, (F.5), (F.6), and the CMT,

$$J_n \to_D (\mathcal{Z}_{2,F} + d_0)' L_F (\mathcal{Z}_{2,F} + d_0). \tag{F.7}$$

Recall that Lemma 4.1(a) implies that

$$n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) \to_D \xi_{1,F} \quad \text{and} \quad n^{1/2}(\widehat{\theta}_2 - \theta_{F_n}) \to_D \xi_{2,F}, \tag{F.8}$$

which together with (F.7) and the CMT implies that

$$\begin{aligned} n^{1/2}(\widehat{\theta}_{\text{pre}} - \theta_{F_n}) &= 1\{J_n > c_\alpha\} n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) + 1\{J_n \le c_\alpha\} n^{1/2}(\widehat{\theta}_2 - \theta_{F_n}) \\ &\to_D (1 - \overline{\omega}_{p,F})\xi_{1,F} + \overline{\omega}_{p,F}\xi_{2,F}, \end{aligned} \tag{F.9}$$

which completes the proof of the claim in (a).

(b) There are two cases to consider: (i) $\|\delta_{F_n}\| > C^{-1}$; and (ii) $\|\delta_{F_n}\| \to 0$. We first consider case (i). As $\overline{g}_1(\widehat{\theta}_2)$ is a subvector of $\overline{g}_2(\widehat{\theta}_2)$,

$$\begin{aligned} J_n &= n\overline{g}_2(\widehat{\theta}_2)'(\widehat{\Omega}_2)^{-1}\overline{g}_2(\widehat{\theta}_2) \\ &\ge n\rho_{\max}^{-1}(\widehat{\Omega}_2)\overline{g}_2(\widehat{\theta}_2)'\overline{g}_2(\widehat{\theta}_2) \\ &\ge n\rho_{\max}^{-1}(\widehat{\Omega}_2)\overline{g}_1(\widehat{\theta}_2)'\overline{g}_1(\widehat{\theta}_2). \end{aligned} \tag{F.10}$$

By (B.22) and (B.23) in the Appendix of CLS,

$$\|\widehat{\theta}_2 - \theta_{F_n}\| \ge C^{-1} \quad \text{with probability approaching 1,} \tag{F.11}$$

which together with Assumption 3.1(ii) and Lemma B.1(i) implies that

$$\overline{g}_1(\widehat{\theta}_2) = M_{1,F}(\widehat{\theta}_2) + o_p(1) \geq C \tag{F.12}$$

with probability approaching 1. By (E.42) and Assumption 3.2(ii), we have

$$\rho_{\max}(\widehat{\Omega}_2) \leq C \quad \text{with probability approaching 1.} \tag{F.13}$$

Combining the results in (F.10), (F.12), and (F.13), we deduce that

$$J_n \geq nC^{-1} \quad \text{with probability approaching 1,} \tag{F.14}$$

which immediately implies that

$$\overline{\omega}_{p,F} = 1\{J_n \leq c_\alpha\} = 0 \tag{F.15}$$

with probability approaching 1, as c_α is a fixed constant. By Lemma 4.1(b), (F.15), and the assumption that Θ is bounded, we have

$$\begin{aligned} n^{1/2}(\widehat{\theta}_{\text{pre}} - \theta_{F_n}) &= 1\{J_n > c_\alpha\} n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) + 1\{J_n \leq c_\alpha\} n^{1/2}(\widehat{\theta}_2 - \theta_{F_n}) \\ &= 1\{J_n > c_\alpha\} n^{1/2}(\widehat{\theta}_1 - \theta_{F_n}) + o_p(1) \to_D \xi_{1,F}, \end{aligned} \tag{F.16}$$

where the convergence in distribution is by the CMT.

We next consider the case that $\|\delta_{F_n}\| \to 0$ and $\|n^{1/2}\delta_{F_n}\| \to \infty$. In the proof of Lemma 4.1, we have shown that $\widehat{\theta}_2 - \theta_{F_n} = o_p(1)$, and that (F.4) and (F.5) hold in this case. It is clear that

$$n^{1/2}\overline{g}_2(\theta_{F_n}) = \mu_n(g_2(W, \theta_{F_n}) + \begin{pmatrix} 0_{r_1 \times 1} \\ n^{1/2}\delta_{F_n} \end{pmatrix}, \tag{F.17}$$

which implies that

$$\begin{aligned} n\overline{g}_2(\theta_{F_n})' L_{F_n} \overline{g}_2(\theta_{F_n}) = {} & [\mu_n(g_2(W, \theta_{F_n})]' L_{F_n} [\mu_n(g_2(W, \theta_{F_n})] \\ & + 2\begin{pmatrix} 0_{1\times r_1} & n^{1/2}\delta'_{F_n} \end{pmatrix} L_{F_n} [\mu_n(g_2(W, \theta_{F_n})] \\ & + \begin{pmatrix} 0_{1\times r_1} & n^{1/2}\delta'_{F_n} \end{pmatrix} L_{F_n} \begin{pmatrix} 0_{1\times r_1} & n^{1/2}\delta'_{F_n} \end{pmatrix}'. \end{aligned} \tag{F.18}$$

By Lemma B.1(v) and Assumptions 3.2(ii)–(iii),

$$[\mu_n(g_2(W, \theta_{F_n})]' L_{F_n} [\mu_n(g_2(W, \theta_{F_n})] = O_p(1). \tag{F.19}$$

In order to bound the third term in (F.18) from below, we shall show that for any $d_0 = (0_{1\times r_1}, d')'$ for $d \in \mathbb{R}^{r_*}$ with $\|d\| = 1$,

$$d_0' L_{F_n} d_0 \geq C^{-1} \tag{F.20}$$

By definition, L_{F_n} has d_θ many zero eigenvalues and $r_2 - d_\theta$ many of eigenvalues of ones. The matrix G_{2,F_n} contains the d_θ many eigenvectors of the zero eigenvalues of L_{F_n}, because

$$L_{F_n} G_{2,F_n} = 0_{r_2 \times d_\theta} \quad \text{and} \quad \rho_{\min}\big(G'_{2,F_n} G_{2,F_n}\big) \geq C^{-1}. \tag{F.21}$$

Let $G_{\perp,F_n}$ denote the orthogonal complement of G_{2,F_n} with $G'_{\perp,F_n} G_{\perp,F_n} = I_{r_2 - d_\theta}$. Then we have

$$\begin{pmatrix} G_{1,F_n} \\ G_{r^*,F_n} \end{pmatrix} a_1 + \begin{pmatrix} G_{1,\perp,F_n} \\ G_{r^*,\perp,F_n} \end{pmatrix} a_2 = \begin{pmatrix} 0_{r_1 \times 1} \\ d \end{pmatrix} \tag{F.22}$$

for some constant vectors $a_1 \in \mathbb{R}^{d_\theta}$ and $a_2 \in \mathbb{R}^{r_2 - d_\theta}$. As $\rho_{\min}(G'_{1,F_n} G_{1,F_n}) \geq C^{-1}$ by Assumption 3.2, we have

$$a_1 = -\big(G'_{1,F_n} G_{1,F_n}\big)^{-1} G'_{1,F_n} G_{1,\perp,F_n} a_2 \tag{F.23}$$

and

$$\big(G_{r^*,\perp,F_n} - G_{r^*,F_n}\big(G'_{1,F_n} G_{1,F_n}\big)^{-1} G'_{1,F_n} G_{1,\perp,F_n}\big) a_2 = d. \tag{F.24}$$

Let $H_{F_n} = G_{r^*,\perp,F_n} - G_{r^*,F_n}(G'_{1,F_n} G_{1,F_n})^{-1} G'_{1,F_n} G_{1,\perp,F_n}$. By $\rho_{\min}(G'_{1,F_n} G_{1,F_n}) \geq C^{-1}$, Assumptions 3.2(ii), (F.24), and the Cauchy–Schwarz inequality,

$$\|d\|^2 = a_2' H_{F_n} H'_{F_n} a_2 \leq C \|a_2\|^2, \tag{F.25}$$

which together with $\|d\| = 1$ implies that

$$\|a_2\|^2 \geq C^{-1}. \tag{F.26}$$

Using (F.21), (F.22), and (F.26), we deduce that

$$\begin{aligned} d_0' L_{F_n} d_0 &= (G_{2,F_n} a_1 + G_{\perp,F_n} a_2)' L_{F_n} (G_{2,F_n} a_1 + G_{\perp,F_n} a_2) \\ &= a_2' G'_{\perp,F_n} L_{F_n} G_{\perp,F_n} a_2 = \|a_2\|^2 \geq C^{-1}, \end{aligned} \tag{F.27}$$

which proves (F.20). By (F.20),

$$\begin{pmatrix} 0_{1 \times r_1} & n^{1/2} \delta'_{F_n} \end{pmatrix} L_{F_n} \begin{pmatrix} 0_{1 \times r_1} & n^{1/2} \delta'_{F_n} \end{pmatrix}' \geq C^{-1} n \|\delta_{F_n}\|^2, \tag{F.28}$$

which together with $n\|\delta_{F_n}\|^2 \to \infty$ implies that

$$\begin{pmatrix} 0_{1 \times r_1} & n^{1/2} \delta'_{F_n} \end{pmatrix} L_{F_n} \begin{pmatrix} 0_{1 \times r_1} & n^{1/2} \delta'_{F_n} \end{pmatrix}' \to \infty. \tag{F.29}$$

Collecting the results in (F.18), (F.19), and (F.29), and by the Cauchy–Schwarz inequality, we deduce that $n\overline{g}_2(\theta_{F_n})' L_{F_n} \overline{g}_2(\theta_{F_n}) \to_p \infty$, which together with (F.5) implies that

$$J_n \to_p \infty. \tag{F.30}$$

Using the same arguments in showing (F.16), we deduce that

$$n^{1/2}(\widehat{\theta}_{\text{pre}} - \theta_{F_n}) \to_D \xi_{1,F}. \tag{F.31}$$

This completes the proof. □

LEMMA F.2. *Under Assumptions* 3.2, *we have*

$$\sup_{h\in H}\mathbb{E}\big[\big(\overline{\xi}'_{p,F}\Upsilon\overline{\xi}_{p,F}\big)^2\big]\leq C. \tag{F.32}$$

PROOF OF LEMMA F.2. By the same arguments in showing (E.51), we have

$$\big(\overline{\xi}'_{p,F}\Upsilon\overline{\xi}_{p,F}\big)^2\leq 8\big(\mathcal{Z}'_{1,F}\Gamma'_{1,F}\Upsilon\Gamma_{1,F}\mathcal{Z}_{1,F}\big)^2+8\big(\overline{\omega}^2_{p,F}\mathcal{Z}'_{d,2,F}B_F\mathcal{Z}_{d,2,F}\big)^2. \tag{F.33}$$

By the first inequality in (B.58) in the Appendix of CLS, we have $\sup_{h\in H}\mathbb{E}[(\xi'_{1,F}\Upsilon\xi_{1,F})^2]\leq C$. Hence by (F.33), to show the inequality in (F.32), it is sufficient to prove that

$$\sup_{h\in H}\mathbb{E}\big[\big(\overline{\omega}^2_{p,F}\mathcal{Z}'_{d,2,F}B_F\mathcal{Z}_{d,2,F}\big)^2\big]\leq C. \tag{F.34}$$

By definition,

$$\overline{\omega}_{p,F}=I\big\{J_\infty(h_{d,F})\leq c_\alpha\big\}=I\big\{\mathcal{Z}'_{d,2,F}L_F\mathcal{Z}_{d,2,F}\leq c_\alpha\big\}. \tag{F.35}$$

By the simple inequality $(a+b)^2\geq a^2/2-2b^2$,

$$(z+d_0)'L_F(z+d_0)\geq d'_0L_Fd_0/2-2z'L_Fz \tag{F.36}$$

for any $z\in\mathbb{R}$, which together with Assumption 3.2 and (F.20) implies that

$$(z+d_0)'L_F(z+d_0)\geq\|d\|^2/C-2z'L_Fz\geq\|d\|^2/C-C\|z\|^2. \tag{F.37}$$

Under Assumption 3.2, $\|B_F\|\leq C$ for any $F\in\mathcal{F}$ which together with the simple inequality $(a+b)^2\leq 2(a^2+b^2)$ implies that

$$(z+d_0)'B_F(z+d_0)\leq 2C\big(\|d\|^2+\|z\|^2\big) \tag{F.38}$$

for any $z\in\mathbb{R}$. Collecting the results in (F.36) and (F.38), we get

$$\begin{aligned}&I\big\{(z+d_0)'L_F(z+d_0)\leq c_\alpha\big\}z'B_Fz\\&\quad\leq 2\mathrm{CI}\big\{\|d\|^2\leq c_\alpha C+C^2\|z\|^2\big\}\big(\|d\|^2+\|z\|^2\big)\\&\quad\leq 2C\big(c_\alpha C+\big(C^2+1\big)\|z\|^2\big),\end{aligned} \tag{F.39}$$

which implies that

$$\begin{aligned}&\sup_{h\in H}\mathbb{E}\big[\big(\overline{\omega}^2_{p,F}\mathcal{Z}'_{d,2,F}B_F\mathcal{Z}_{d,2,F}\big)^2\big]\\&\quad\leq 4C^2\mathbb{E}\big[\big(c_\alpha C+\big(C^2+1\big)\mathcal{Z}'_{2,F}\mathcal{Z}_{2,F}\big)^2\big]\\&\quad\leq C\big(c_\alpha+\mathbb{E}\big[\big(\mathcal{Z}'_{2,F}\mathcal{Z}_{2,F}\big)^2\big]\big)=C\big(c_\alpha+3\rho^2_{\max}(\Omega_2)r_2\big).\end{aligned} \tag{F.40}$$

This completes the proof. □

LEMMA F.3. *Let* $g_{p,\zeta}(h) \equiv \mathbb{E}[\min\{\overline{\xi}'_{p,F} \Upsilon \overline{\xi}_{p,F}, \zeta\} - \min\{\xi'_{1,F} \Upsilon \xi_{1,F}, \zeta\}]$. *Under Assumptions* 3.2, *we have*

$$\lim_{\zeta\to\infty} \sup_{h\in H} \left[\left| g_{p,\zeta}(h) - g_p(h) \right| \right] = 0, \tag{F.41}$$

where $\sup_{h\in H}[|g_p(h)|] \leq C$.

PROOF OF LEMMA F.3. The proof follows the same arguments of the proof of Lemma B.16 with the second inequality in (B.58) in the Appendix of CLS replaced by (F.32). □

APPENDIX G: SIMULATION RESULTS ON TRUNCATED RISK FOR SECTION 6

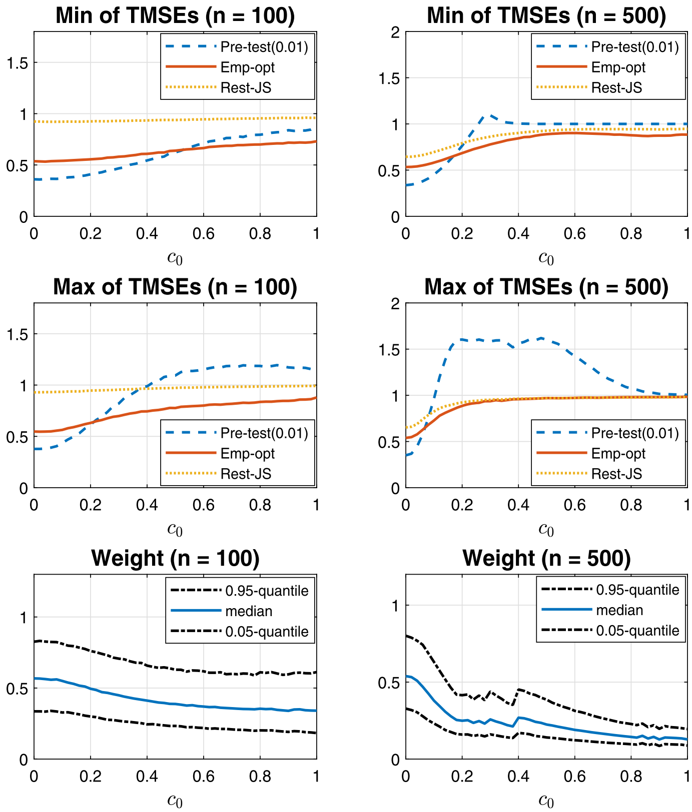

FIGURE G.1. Finite sample truncated MSEs of the pre-test and averaging GMM estimators in S1. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

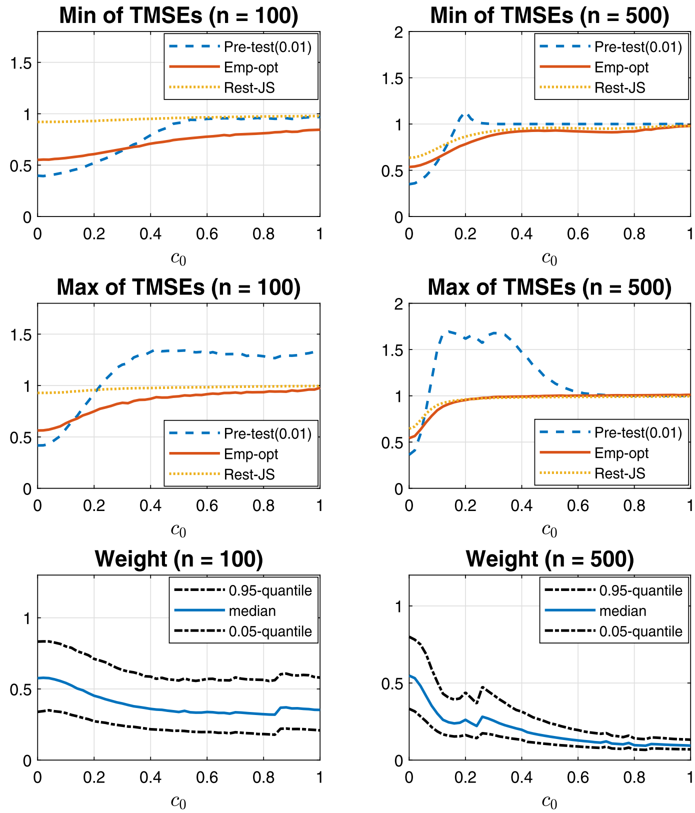

FIGURE G.2. Finite sample truncated MSEs of the pre-test and averaging GMM estimators in S2. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

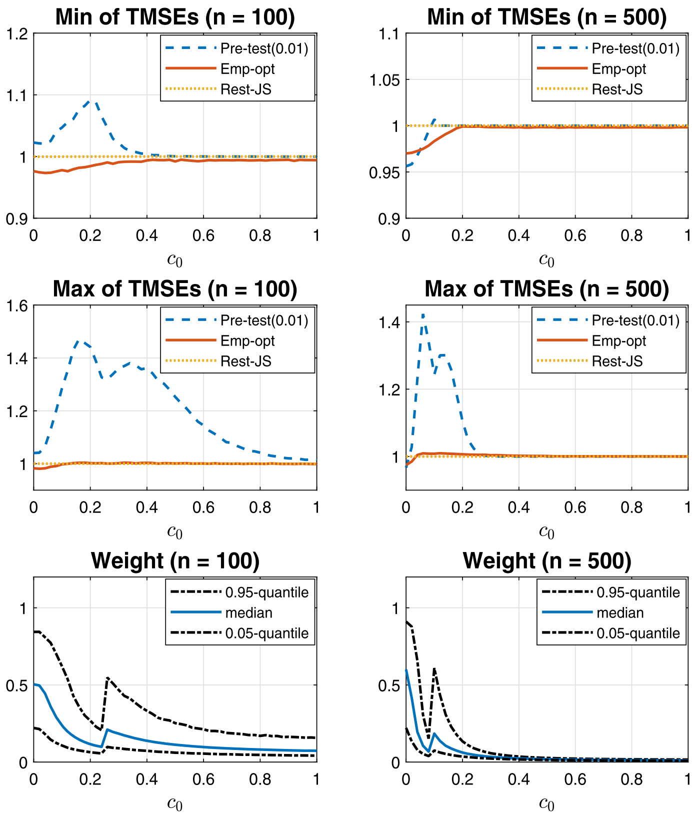

FIGURE G.3. Finite sample truncated MSEs of the pre-test and averaging GMM estimators in S3. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

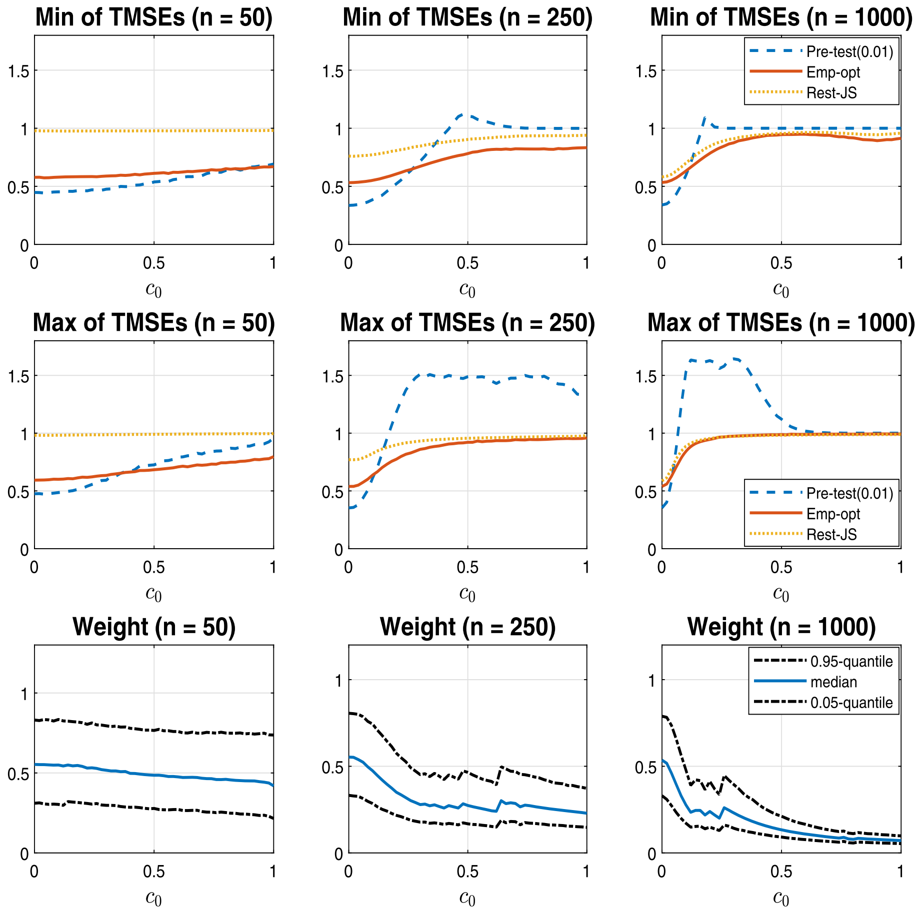

FIGURE G.4. Finite sample truncated MSEs of the pre-test and averaging GMM estimators in S1. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

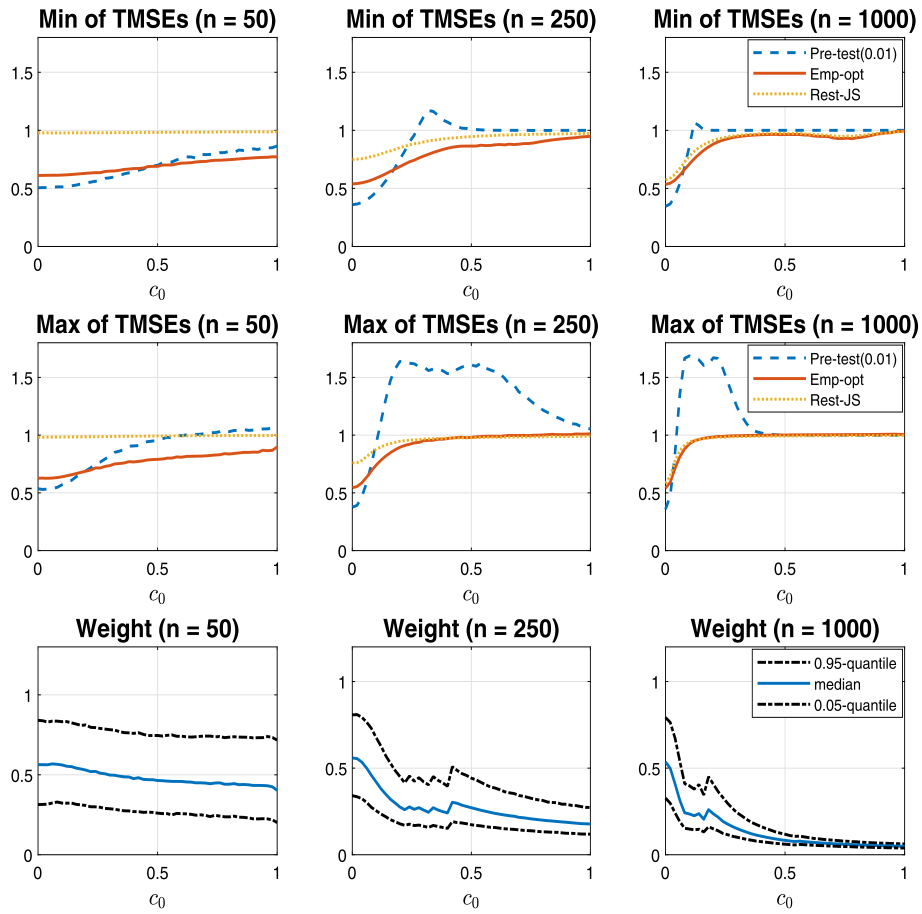

FIGURE G.5. Finite sample truncated MSEs of the pre-test and averaging GMM estimators in S2. Note: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

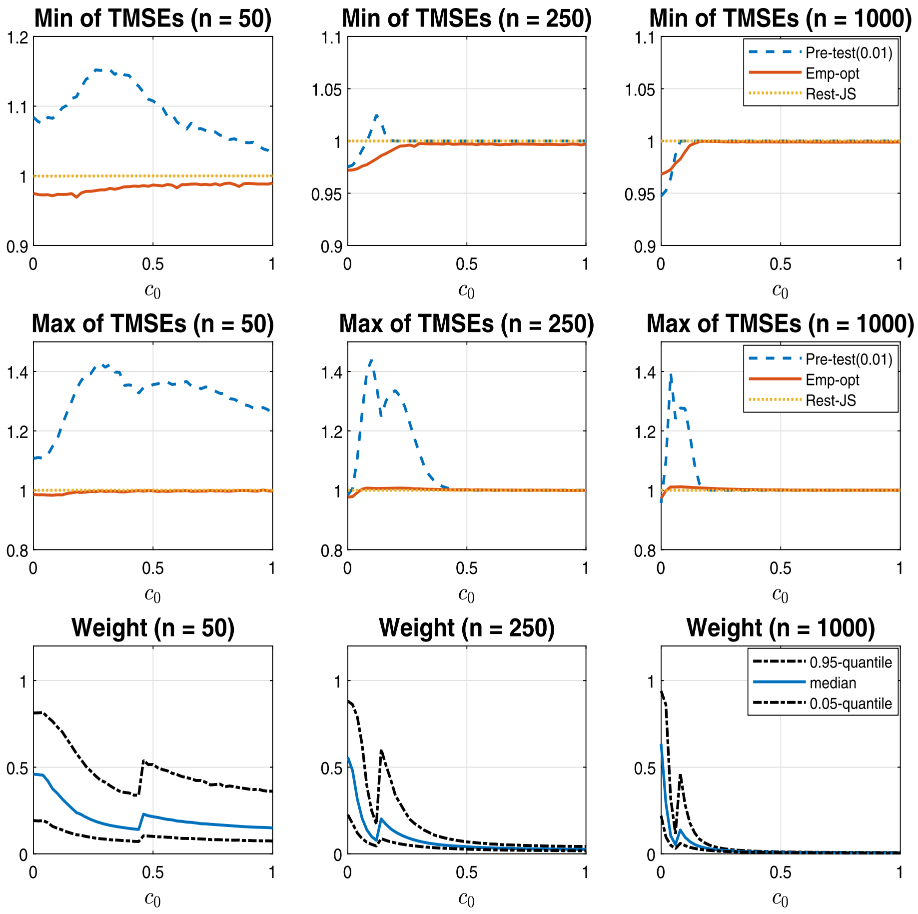

FIGURE G.6. Finite sample truncated MSEs of the pre-test and averaging GMM estimators in S3. Note: “Pre-test(0.01)” refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; “Emp-opt” refers to the averaging GMM estimator based on the empirical optimal weight; “Rest-JS” refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

TABLE G.1. The lower and upper bounds of the finite sample relative truncated MSEs.

		Design S1		Design S2		Design S3	
		Lower	Upper	Lower	Upper	Lower	Upper
$n = 50$	$\widehat{\theta}_{\text{oe}}$	0.5732	0.7968	0.6113	0.8980	0.9694	1.0012
	$\widehat{\theta}_{\text{JS}}$	0.9755	0.9959	0.9776	0.9978	0.9995	1.0003
	$\widehat{\theta}_{\text{pret}}$	0.4424	0.9574	0.5057	1.0973	1.0324	1.4283
$n = 100$	$\widehat{\theta}_{\text{oe}}$	0.5325	0.8789	0.5513	0.9781	0.9733	1.0040
	$\widehat{\theta}_{\text{JS}}$	0.9208	0.9911	0.9202	0.9956	0.9996	1.0002
	$\widehat{\theta}_{\text{pret}}$	0.3586	1.1940	0.3937	1.3539	0.9990	1.4709
$n = 250$	$\widehat{\theta}_{\text{oe}}$	0.5316	0.9587	0.5384	1.0118	0.9720	1.0079
	$\widehat{\theta}_{\text{JS}}$	0.7591	0.9787	0.7506	0.9923	0.9999	1.0000
	$\widehat{\theta}_{\text{pret}}$	0.3360	1.5106	0.3598	1.6392	0.9753	1.4394
$n = 500$	$\widehat{\theta}_{\text{oe}}$	0.5331	0.9846	0.5355	1.0112	0.9700	1.0096
	$\widehat{\theta}_{\text{JS}}$	0.6443	0.9823	0.6359	0.9953	1.0000	1.0000
	$\widehat{\theta}_{\text{pret}}$	0.3368	1.6196	0.3495	1.6937	0.9562	1.4236
$n = 1000$	$\widehat{\theta}_{\text{oe}}$	0.5335	0.9934	0.5341	1.0082	0.9681	1.0119
	$\widehat{\theta}_{\text{JS}}$	0.5803	0.9890	0.5737	0.9978	1.0000	1.0000
	$\widehat{\theta}_{\text{pret}}$	0.3395	1.6433	0.3451	1.6864	0.9473	1.3953

Note: 1. $\hat{\theta}_{\text{JS}}$ and $\hat{\theta}_{\text{pret}}$ denote the GMM averaging estimator based on the weight in (6.1) and the pre-testing GMM estimator based on J-test with nominal size 0.01, respectively; 2. the "Upper" and "Lower" refer to the upper bound and the lower bound of the finite sample relative MSEs among all DGPs considered in the simulation design given the sample size.

Appendix H: Simulation under the Student-t distribution

In this subsection, we report the simulation results on the finite sample properties of the pre-test and averaging GMM estimators, when the residual term u in the structure equation (6.4) of CLS is a Student-t random variable with degree of freedom 2. The simulation design is the same as the one in Section 6.1, except that we generate the structural error u in the following way:

$$u = \frac{u^*}{((\eta_1^2 + \eta_2^2)/2)^{1/2}},$$

where η_1 and η_2 are independent standard normal random variables which are independent with respect to $(Z_1, \ldots, Z_{18}, \varepsilon_1, \ldots, \varepsilon_6, u^*)$.[1] We call this simulation design as S4.

The finite sample untruncated and truncated MSEs are reported in Figures H.1–H.5. It is interesting to see that in this simulation design, both the pre-test GMM estimator and the averaging GMM estimator have smaller finite sample MSEs than the conservative GMM estimator. The main reason for this phenomenon is that the residual term u in the structural equation is Student-t with degree of freedom 2, which implies that u has infinite variance, and hence the conservative GMM estimator has large variance in finite samples. When the extra IVs Z_j^* $(j = 1, \ldots, 6)$ are used in the GMM estimation, the finite sample variances of the GMM estimator is greatly reduced. Therefore, the finite samples biases of the pre-testing GMM estimator and the averaging GMM estimator introduced by the extra IVs Z_j^* $(j = 1, \ldots, 6)$ are more than offset by the reduced finite sample variances, which enables both estimator have smaller finite sample MSEs.

[1]In this design, the structural error u does not enter the possibly invalid IVs (6.4). Therefore, the IVs and the regressors are normally distributed. We thank an anonymous who suggested this simulation design.

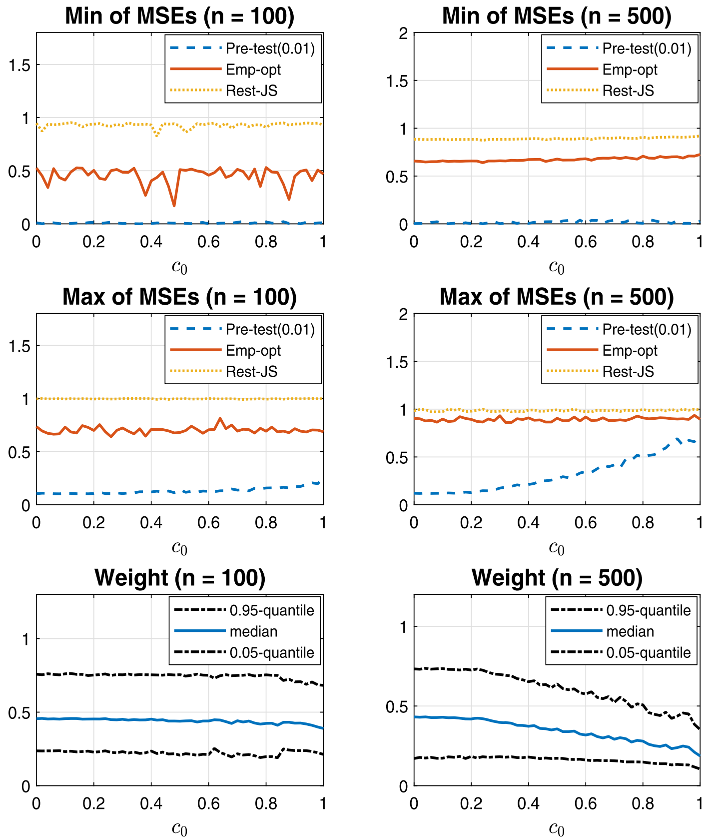

FIGURE H.1. Finite sample MSEs of the pre-test and averaging GMM estimators in S4. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively.

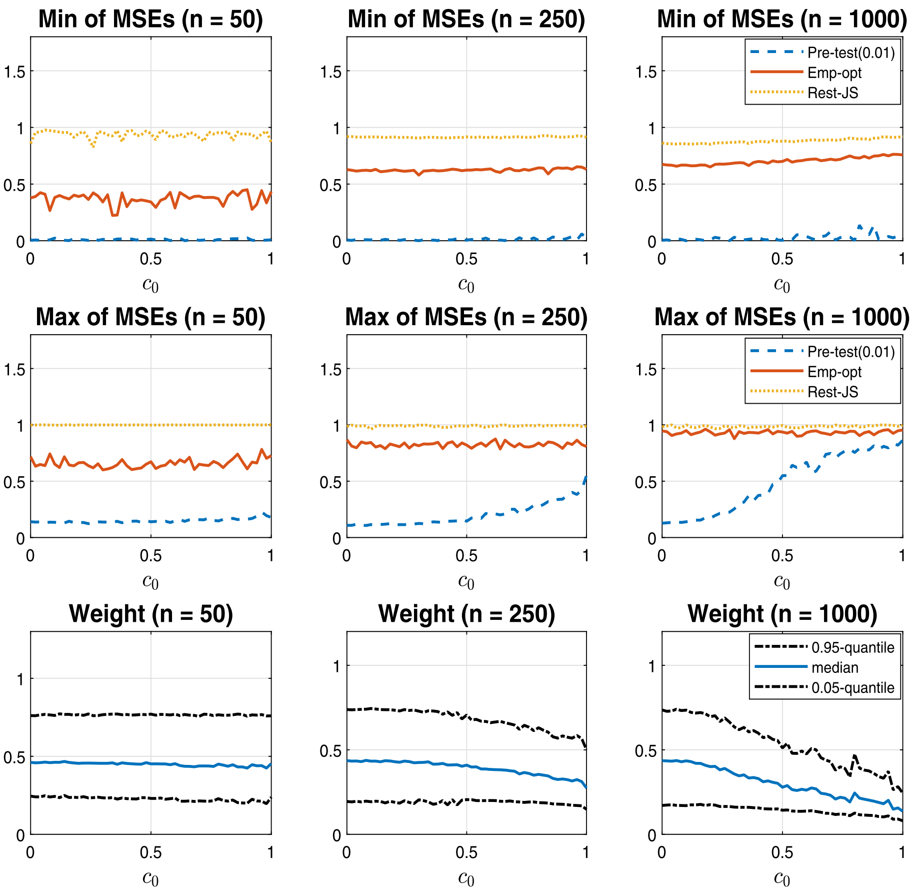

FIGURE H.2. Finite sample MSEs of the pre-test and averaging GMM estimators in S4. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively.

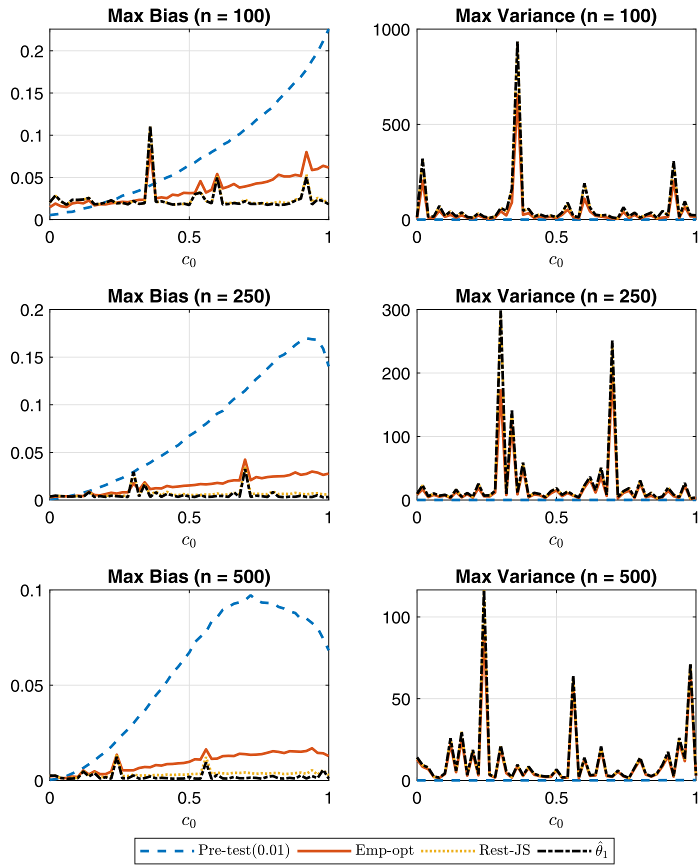

FIGURE H.3. Finite sample biases and variances in S4. *Note*: “Pre-test(0.01)” refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; “Emp-opt” refers to the averaging GMM estimator based on the empirical optimal weight; “Rest-JS” refers to the averaging estimators based on the restricted James–Stein weight, respectively.The truncation parameter for the truncated MSE is $\zeta = 1000$.

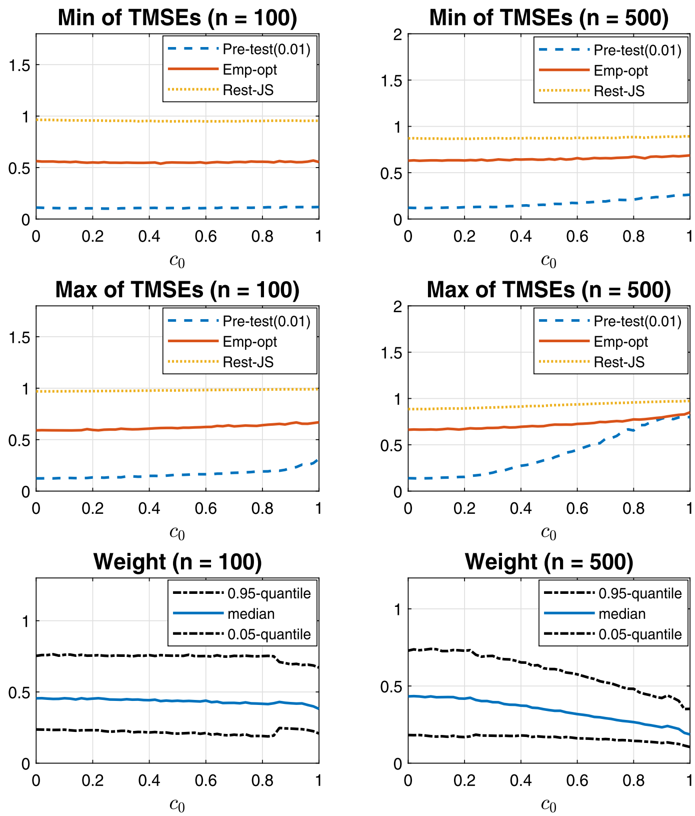

FIGURE H.4. Finite sample TMSEs of the pre-test and averaging GMM estimators in S4. *Note*: "Pre-test(0.01)" refers to the pre-test GMM estimator based on the J-test with nominal size 0.01; "Emp-opt" refers to the averaging GMM estimator based on the empirical optimal weight; "Rest-JS" refers to the averaging estimators based on the restricted James–Stein weight, respectively. The truncation parameter for the truncated MSE is $\zeta = 1000$.

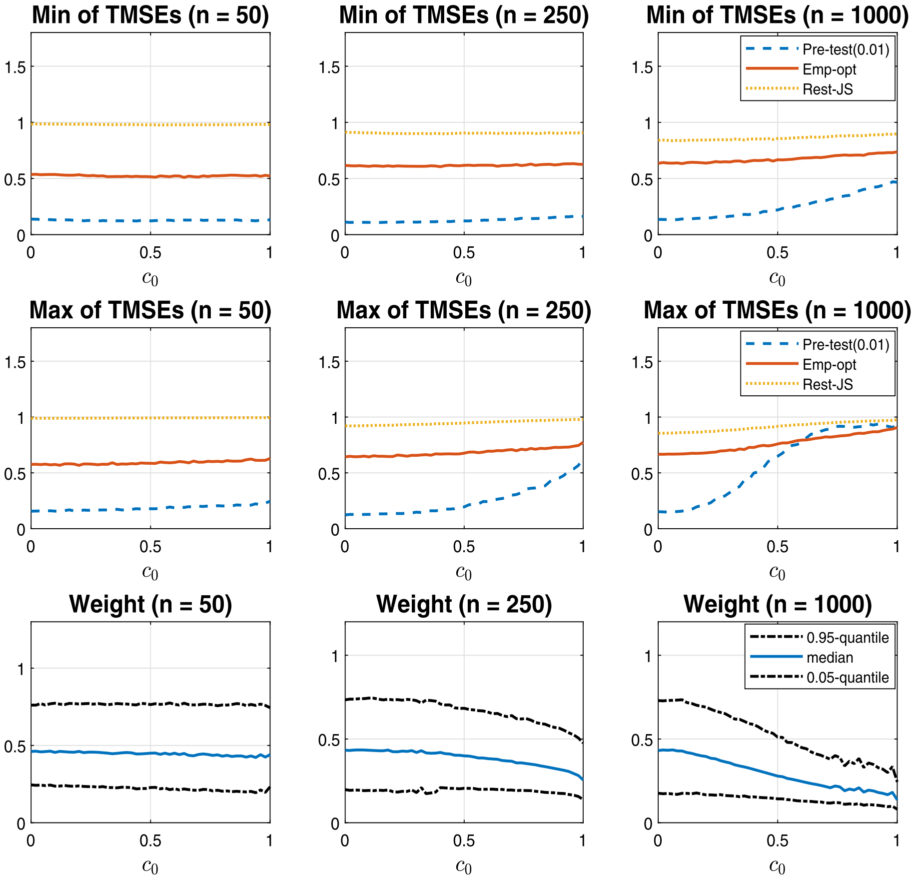

FIGURE H.5. Finite sample TMSEs of the pre-test and averaging GMM estimators in S4.

Reference

Cheng, X., Z. Liao, and R. Shi (2019), “On uniform asymptotic risk of averaging GMM estimators.” *Quantitative Economics*, 10, 931–979. https://doi.org/10.3982/QE711. [1]

Co-editor Karl Schmedders handled this manuscript.

Manuscript received 3 May, 2016; final version accepted 27 November, 2018; available online 9 January, 2019.