Econometrica Supplementary Material

SUPPLEMENT TO "MACRO-FINANCE DECOUPLING: ROBUST EVALUATIONS OF MACRO ASSET PRICING MODELS"

(*Econometrica*, Vol. 90, No. 2, March 2022, 685–713)

XU CHENG
Department of Economics, University of Pennsylvania

WINSTON WEI DOU
Finance Department, The Wharton School, University of Pennsylvania

ZHIPENG LIAO
Department of Economics, University of California, Los Angeles

This supplemental appendix provides the following supporting materials. Sections SA–SC provide the proofs of Lemmas A1–A7 in the appendix to the main text Cheng, Dou, and Liao (2022). Section SA provides the proofs of several lemmas on the asymptotic convergence of the random components in the test statistic $\mathcal{T}$ and the conditional critical value $c_\alpha(\hat{d})$. Section SB verifies the bounded Lipschitz properties of the test statistic and the conditional critical value, which are used to show their weak convergence in large samples. Section SC includes some auxiliary lemmas. Section SD provides additional theoretical results on the power of the proposed conditional test. Section SE provides comparison with some power envelopes through simulations. Section SF collects details and additional results of the empirical application.

SA. PROOFS FOR ASYMPTOTIC CONVERGENCE

THIS SECTION provides the proofs of Lemmas A1, A2, A6, and A7 in the Appendix to Cheng, Dou, and Liao (2022) under their Assumptions 1, 2, 3, and 4. The proofs of Lemmas A3, A4, and A5 in the Appendix to Cheng, Dou, and Liao (2022) are in Sections SB and SC of this supplement. Throughout this supplement, we use $\lambda_{\min}(A)$ to denote the minimal eigenvalue of a real symmetric matrix A, and $\|\cdot\|$ denotes the matrix Frobenius norm.

PROOF OF LEMMA A1: (a) Define $C_{1,n} \equiv \sup_{\theta\in\Theta} |\bar{g}(\theta)'(\hat{\Omega}(\theta))^{-1}\bar{g}(\theta) - G(\theta)'(\Omega(\theta))^{-1} \times G(\theta)|$, where $G(\theta) \equiv \mathbb{E}[\bar{g}(\theta)]$. Then by Assumptions 1(i), (iii), 2(i), and 3(iii),

$$C_{1,n} = o_p(1), \tag{SA.1}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. Consider any $\varepsilon > 0$. By the definition of $\hat{\theta}$ and θ_0,

$$\begin{aligned}
&\sup_{\mathbb{P}\in\mathcal{P}_0} \mathbb{P}\big(\|\hat{\theta} - \theta_0\| \geq \varepsilon\big) \\
&\leq \sup_{\mathbb{P}\in\mathcal{P}_0} \mathbb{P}\Big(\min_{\theta\in B_\varepsilon^c(\theta_0)} \bar{g}(\theta)'\big(\hat{\Omega}(\theta)\big)^{-1}\bar{g}(\theta) \leq \bar{g}(\theta_0)'\big(\hat{\Omega}(\theta_0)\big)^{-1}\bar{g}(\theta_0)\Big) \\
&\leq \sup_{\mathbb{P}\in\mathcal{P}_0} \mathbb{P}\Big(\min_{\theta\in B_\varepsilon^c(\theta_0)} G(\theta)'\big(\Omega(\theta)\big)^{-1}G(\theta) \leq 2C_{1,n}\Big)
\end{aligned}$$

Xu Cheng: xucheng@econ.upenn.edu
Winston Wei Dou: wdou@wharton.upenn.edu
Zhipeng Liao: zhipeng.liao@econ.ucla.edu

 https://doi.org/10.3982/ECTA18506

$$\leq \sup_{\mathbb{P}\in\mathcal{P}_0} \mathbb{P}\big(c_\lambda^{-1}\delta_\varepsilon^2 \leq 2C_{1,n}\big), \tag{SA.2}$$

where the second inequality is by the definition of $C_{1,n}$ and the third inequality is by Assumption 3. Combining the results in (SA.1) and (SA.2), we deduce that

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in\mathcal{P}_0} \mathbb{P}\big(\|\hat{\theta}-\theta_0\| \geq \varepsilon\big) = 0 \quad \text{for any } \varepsilon > 0. \tag{SA.3}$$

Let $C_{2,n} \equiv \sup_{\theta\in\Theta} \|q(\theta) - Q(\theta)\|$. Then, by Assumption 1(ii),

$$C_{2,n} = o_p(1) \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}. \tag{SA.4}$$

Applying the first-order expansion, we get

$$g(\hat{\theta}) = g(\theta_0) + q(\tilde{\theta})n^{1/2}(\hat{\theta} - \theta_0), \tag{SA.5}$$

where $\tilde{\theta}$ is the mean value between θ_0 and $\hat{\theta}$ and it may vary across rows. By Assumption 1(ii), the consistency of $\hat{\theta}$, and (SA.4),

$$q(\tilde{\theta}) = Q(\tilde{\theta}) + o_p(1) = Q + o_p(1) = O_p(1) \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}_0. \tag{SA.6}$$

Similarly, we can show that

$$q(\hat{\theta}) = Q(\hat{\theta}) + o_p(1) = Q + o_p(1) = O_p(1) \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}_0. \tag{SA.7}$$

By Assumption 2(i), (ii) and the consistency of $\hat{\theta}$,

$$\hat{\Omega} \equiv \hat{\Omega}(\hat{\theta}) = \Omega + o_p(1) \tag{SA.8}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. Applying the chain rule, we get the first-order condition of $\hat{\theta}$:

$$0_{d_\theta\times 1} = 2q(\hat{\theta})'\hat{\Omega}^{-1}g(\hat{\theta}) - \left(n^{-1/2}g(\hat{\theta})'\hat{\Omega}^{-1}\frac{\partial\hat{\Omega}(\hat{\theta})}{\partial\theta_j}\hat{\Omega}^{-1}g(\hat{\theta})\right)_{j=1,\dots,d_\theta}, \tag{SA.9}$$

where $(a_j)_{j=1,\dots,d_\theta} \equiv (a_1, \dots, a_{d_\theta})'$ for any real numbers $a_1, \dots, a_{d_\theta}$. By Assumptions 1(i), (iii), 2(iv), and 3(iii), the consistency of $\hat{\theta}$, (SA.5), (SA.6), (SA.7), and (SA.8),

$$\begin{aligned} &q(\hat{\theta})'\hat{\Omega}^{-1}g(\hat{\theta}) \\ &\quad = q(\hat{\theta})'\hat{\Omega}^{-1}g(\theta_0) + q(\hat{\theta})'\hat{\Omega}^{-1}q(\tilde{\theta})n^{1/2}(\hat{\theta}-\theta_0) \\ &\quad = Q'\Omega^{-1}g(\theta_0) + \big(Q'\Omega^{-1}Q + o_p(1)\big)n^{1/2}(\hat{\theta}-\theta_0) + o_p(1) \end{aligned} \tag{SA.10}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. Similarly, we can show that for $j = 1, \dots, d_\theta$,

$$n^{-1/2}g(\hat{\theta})'\hat{\Omega}^{-1}\frac{\partial\hat{\Omega}(\hat{\theta})}{\partial\theta_j}\hat{\Omega}^{-1}g(\hat{\theta}) = n^{1/2}(\hat{\theta}-\theta_0)o_p(1) + o_p(1)$$
$$\text{uniformly over } \mathbb{P} \in \mathcal{P}_0. \tag{SA.11}$$

Combining the results in (SA.9), (SA.10), (SA.11), and applying Assumptions 1(iii) and 3(iii), we deduce that

$$n^{1/2}(\hat{\theta}-\theta_0) = -\left(Q'\Omega^{-1}Q\right)^{-1}Q'\Omega^{-1}g(\theta_0)+o_p(1) = O_p(1) \tag{SA.12}$$

uniformly over $\mathbb{P}\in\mathcal{P}_0$, which proves the first claim of the lemma.

(b) This claim follows by Assumptions 1 and 3(iii), (SA.5), (SA.6), and (SA.12).

(c) This claim has been proved in (SA.8).

(d) By Assumptions 2(iv) and 3(iii), and (SA.8),

$$\lim_{n\to\infty}\inf_{\mathbb{P}\in\mathcal{P}_0}\mathbb{P}\left(K^{-1}\leq\lambda_{\min}(\hat{\Omega})\leq\lambda_{\max}(\hat{\Omega})\leq K\right)=1. \tag{SA.13}$$

By Assumptions 1(iii) and 3(iii), (SA.7), and (SA.8), we have

$$q(\hat{\theta})'\hat{\Omega}^{-1}q(\hat{\theta}) = Q'\Omega^{-1}Q+o_p(1) \quad \text{uniformly over } \mathbb{P}\in\mathcal{P}_0, \tag{SA.14}$$

which together with Assumption 1(iii), 2(iv), and 3(iii) implies that

$$\lim_{n\to\infty}\inf_{\mathbb{P}\in\mathcal{P}_0}\mathbb{P}\left(K^{-1}\leq\lambda_{\min}\left(q(\hat{\theta})'(\hat{\Omega})^{-1}q(\hat{\theta})\right)\leq\lambda_{\max}\left(q(\hat{\theta})'(\hat{\Omega})^{-1}q(\hat{\theta})\right)\leq K\right)=1. \tag{SA.15}$$

Let $\|\cdot\|_S$ denote the matrix operator norm. By (7.2.13) in Horn and Johnson (1990),

$$\left\|\hat{\Omega}^{1/2}-\Omega^{1/2}\right\|_S \leq \|\hat{\Omega}-\Omega\|_S\left\|\Omega^{-1/2}\right\|_S, \tag{SA.16}$$

which together with Assumption 3(iii) and (SA.8) implies that

$$\left\|\hat{\Omega}^{1/2}-\Omega^{1/2}\right\|_S = o_p(1) \quad \text{uniformly over } \mathbb{P}\in\mathcal{P}_0. \tag{SA.17}$$

By (SA.17) and the relation between the operator norm and the Frobenius norm,

$$\left\|\hat{\Omega}^{1/2}-\Omega^{1/2}\right\| = o_p(1) \quad \text{uniformly over } \mathbb{P}\in\mathcal{P}_0. \tag{SA.18}$$

By Assumptions 1(iii) and 3(iii), (SA.7), and (SA.18),

$$\hat{\Omega}^{-1/2}q(\hat{\theta}) = \Omega^{-1/2}Q+o_p(1) = O_p(1) \tag{SA.19}$$

uniformly over $\mathbb{P}\in\mathcal{P}_0$. The claim in the lemma follows by (SA.14), (SA.15), and (SA.19).

(e) By Assumption 2(i), (ii) and the consistency of $\hat{\theta}$,

$$\sup_{\theta\in\Theta}\left\|\hat{\Omega}(\theta,\hat{\theta})-\Omega(\theta,\theta_0)\right\| = o_p(1) \quad \text{uniformly over } \mathbb{P}\in\mathcal{P}_0. \tag{SA.20}$$

By (SA.13) and (SA.20),

$$\begin{aligned}\sup_{\theta\in\Theta}\left\|\left(\hat{\Omega}(\theta,\hat{\theta})-\Omega(\theta,\theta_0)\right)\hat{\Omega}^{-1}\right\| &\leq \left(\lambda_{\min}(\hat{\Omega})\right)^{-1}\sup_{\theta\in\Theta}\left\|\hat{\Omega}(\theta,\hat{\theta})-\Omega(\theta,\theta_0)\right\| \\ &= o_p(1)\end{aligned} \tag{SA.21}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$, where the inequality is by the Cauchy–Schwarz inequality. Similarly,

$$\begin{aligned}
\sup_{\theta\in\Theta}\left\|\Omega(\theta,\theta_0)\left(\hat{\Omega}^{-1}-\Omega^{-1}\right)\right\| &= \sup_{\theta\in\Theta}\left\|\Omega(\theta,\theta_0)\hat{\Omega}^{-1}(\hat{\Omega}-\Omega)\Omega^{-1}\right\| \\
&\leq \left(\lambda_{\min}(\hat{\Omega})\lambda_{\min}(\Omega)\right)^{-1}\sup_{\theta,\tilde{\theta}\in\Theta}\left\|\Omega(\theta,\tilde{\theta})\right\|\|\hat{\Omega}-\Omega\| \\
&= o_p(1) \qquad \text{(SA.22)}
\end{aligned}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$, where the inequality is by the Cauchy–Schwarz inequality, and the last equality is by Assumptions 2(iv) and 3(iii), (SA.8), and (SA.13). Collecting the results in (SA.21) and (SA.22), we deduce that

$$\begin{aligned}
&\sup_{\theta\in\Theta}\left\|\hat{V}(\theta)-V(\theta)\right\| \\
&\quad\leq \sup_{\theta\in\Theta}\left\|\left(\hat{\Omega}(\theta,\hat{\theta})-\Omega(\theta,\theta_0)\right)\hat{\Omega}^{-1}\right\| + \sup_{\theta\in\Theta}\left\|\Omega(\theta,\theta_0)\left(\hat{\Omega}^{-1}-\Omega^{-1}\right)\right\| \\
&\quad= o_p(1) \qquad \text{(SA.23)}
\end{aligned}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. By Assumptions 2(iv) and 3(iii),

$$\sup_{\theta\in\Theta}\left\|V(\theta)\right\| \leq \sup_{\theta\in\Theta}\left\|S_0\Omega(\theta,\theta_0)\Omega^{-1}\right\| \leq \left(\lambda_{\min}(\Omega)\right)^{-1}\sup_{\theta\in\Theta}\left\|\Omega(\theta,\theta_0)\right\| \leq c_\lambda^{-1}C_\Omega, \qquad \text{(SA.24)}$$

which finishes the proof. *Q.E.D.*

PROOF OF LEMMA A2: To link $\hat{\xi}$ and ξ^*, we first define

$$\tilde{\xi} \equiv \left(\tilde{v}', \tilde{m}(\cdot)', \operatorname{vec}\left(V(\cdot)\right)', \operatorname{vech}(\Omega)', \operatorname{vech}\left(\Omega_0(\cdot)\right)', \operatorname{vech}(M)'\right)', \qquad \text{(SA.25)}$$

where $\tilde{v} \equiv \Omega^{1/2}M\Omega^{-1/2}g(\theta_0)$ and $\tilde{m}(\cdot) \equiv g_0(\cdot) - V(\cdot)\tilde{v}$. The difference between $\tilde{\xi}$ and ξ^* lies in the first two elements, where $\tilde{v}$ and $\tilde{m}(\cdot)$ in $\tilde{\xi}$ involve the empirical process $g(\cdot)$, but $\Omega^{1/2}Mv^*$ and $m^*(\cdot)$ in ξ^* involve the limiting Gaussian process $\psi(\cdot)$. Under Assumption 1(i), $g(\cdot) - \mathbb{E}[g(\cdot)]$ weakly converges to $\psi(\cdot)$, which is equivalent to

$$\lim_{n\to\infty}\sup_{\mathbb{P}\in\mathcal{P}_0}\sup_{f\in\mathrm{BL}_1}\left\|\mathbb{E}\left[f\left(g-\mathbb{E}[g]\right)\right]-E\left[f(\psi)\right]\right\| = 0. \qquad \text{(SA.26)}$$

Furthermore, Assumptions 1(iii), 2(iv), and 3(iii) imply that $\tilde{v}$ and $\tilde{m}(\cdot)$ are Lipschitz in $g(\cdot)$, and hence any bounded Lipschitz function of $\tilde{\xi}$ can be written as a bounded Lipschitz function of $\tilde{\xi}_1$, where $\tilde{\xi}_1$ replaces $g_0(\cdot)$ in $\tilde{\xi}$ with $g_0(\cdot) - E[g_0(\cdot)]$, which together with (SA.26) implies that

$$\lim_{n\to\infty}\sup_{\mathbb{P}\in\mathcal{P}_0}\sup_{f\in\mathrm{BL}_1}\left\|\mathbb{E}\left[f(\tilde{\xi})\right]-E\left[f(\xi^*)\right]\right\| = 0. \qquad \text{(SA.27)}$$

Next, note that the difference between $\hat{\xi}$ and $\tilde{\xi}$ is that $\hat{\Omega}$, $\hat{M}$, and $\hat{V}(\theta)$ in $\hat{\xi}$ are replaced by their probability limits in Lemma A1(c), (d), (e) of Cheng, Dou, and Liao (2022). Thus, we have $\hat{\xi} = \tilde{\xi} + o_p(1)$ uniformly over $\mathbb{P} \in \mathcal{P}_0$ following Lemma A1(c), (d), (e) in Cheng, Dou, and Liao (2022), which implies

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in\mathcal{P}_0} \sup_{f\in \mathrm{BL}_1} \left\| \mathbb{E}\left[f(\hat{\xi})\right] - \mathbb{E}\left[f(\tilde{\xi})\right]\right\| = 0. \tag{SA.28}$$

The desirable result follows from (SA.27), (SA.28), and the triangle inequality. *Q.E.D.*

PROOF OF LEMMA A6: Since $0 \leq L(v; d^*) \leq v'Mv$ for any $v \in \mathbb{R}^k$ and $ut_C(u) \leq C$ for any $u \geq 0$, we have

$$\left|L_C\left(v; d^*\right) - \overline{L}_C\left(v; d^*\right)\right| = L\left(v; d^*\right)t_C\left(v'Mv\right)I\left\{\|v\|^2 > C\right\} \leq CI\left\{\|v\|^2 > C\right\} \tag{SA.29}$$

for any $v \in \mathbb{R}^k$, which implies that

$$P\left(\left|L_C\left(v^*, d^*\right) - \overline{L}_C\left(v^*, d^*\right)\right| > \varepsilon\right) \leq P\left(I\left\{\left\|v^*\right\|^2 > C\right\} > \varepsilon/C\right) \leq P\left(\left\|v^*\right\|^2 > C\right). \tag{SA.30}$$

Since $\|v^*\|^2$ follows the chi-square distribution with degree of freedom k, there exists a constant C_δ such that $P(\|v^*\|^2 > C_\delta) \leq \delta/4$ which together with (SA.30) implies that, for any $C \geq C_\delta$,

$$P\left(\left|L_C\left(v^*, d^*\right) - \overline{L}_C\left(v^*, d^*\right)\right| > \varepsilon\right) \leq \delta/4. \tag{SA.31}$$

By the union bound of probability and (SA.31), we have, for any $C \geq C_\delta$,

$$\begin{aligned} &P\left(L_C\left(v^*, d^*\right) > c_{\alpha,C}\left(d^*\right) + \varepsilon\right) \\ &\quad \leq P\left(\overline{L}_C\left(v^*, d^*\right) + \left|L_C\left(v^*, d^*\right) - \overline{L}_C\left(v^*, d^*\right)\right| > c_{\alpha,C}\left(d^*\right) + \varepsilon\right) \\ &\quad \leq P\left(\overline{L}_C\left(v^*, d^*\right) > c_{\alpha,C}\left(d^*\right)\right) + \delta/4 \leq \alpha + \delta/4, \end{aligned} \tag{SA.32}$$

where the last inequality is by the definition of $c_{\alpha,C}(d^*)$. Since $R_C(\xi^*) = L_C(v^*, d^*)$ by definition, the claim of the lemma follows from (SA.32). *Q.E.D.*

PROOF OF LEMMA A7: (a) Since $v^* = O_p(1)$, by Assumptions 1(iii), 2(iv), and 3(iii), Lemma A1(b), (d) in Cheng, Dou, and Liao (2022), and (SA.18), we have, uniformly over $\mathbb{P} \in \mathcal{P}_0$,

$$\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta}) = \Omega^{1/2}M\Omega^{-1/2}\left(\Omega^{1/2}v^* - g(\theta_0)\right) + o_p(1) = O_p(1), \tag{SA.33}$$

which together with Lemma A1(e) in Cheng, Dou, and Liao (2022) implies that

$$\sup_{\theta\in\Theta}\left\|\hat{V}(\theta)\left(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta})\right)\right\| = O_p(1) \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}_0. \tag{SA.34}$$

Therefore, by the triangle inequality, Assumption 1, and (SA.34),

$$\begin{aligned} &\sup_{\theta\in\Theta}\left\|n^{-1/2}\left(\hat{m}(\theta) + \hat{V}(\theta)\hat{\Omega}^{1/2}\hat{M}v^*\right) - \mathbb{E}\left[g_0(\theta)\right]\right\| \\ &\quad \leq \sup_{\theta\in\Theta}\left\|n^{-1/2}g_0(\theta) - \mathbb{E}\left[g_0(\theta)\right]\right\| + \sup_{\theta\in\Theta}\left\|n^{-1/2}\hat{V}(\theta)\left(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta})\right)\right\| \\ &\quad = o_p(1) \end{aligned} \tag{SA.35}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. By Assumptions 1(iii), 2(i), 3(iii), and 4(i), and (SA.35), we can apply similar arguments in the proof of (SA.3) to deduce that

$$\hat{\theta}^* = \theta_0 + o_p(1) \tag{SA.36}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$. Applying the chain rule, we get the first-order condition of $\hat{\theta}^*$:

$$\begin{aligned} 0_{d_\theta \times 1} = {} & 2\Big(\frac{\partial}{\partial \theta'}\big(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*\big)\Big)' \big(\hat{\Omega}_0^*\big)^{-1}\big(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*\big) \\ & - \Bigg(\big(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*\big)'\big(\hat{\Omega}_0^*\big)^{-1}\frac{\partial \hat{\Omega}_0(\hat{\theta}^*)}{\partial \theta_j}\big(\hat{\Omega}_0^*\big)^{-1} \\ & \times \big(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*\big)\Bigg)_{j=1,\ldots,d_\theta}, \end{aligned} \tag{SA.37}$$

where $\hat{\Omega}_0^* \equiv \hat{\Omega}_0(\hat{\theta}^*)$ and $\hat{V}_S(\hat{\theta}^*) \equiv \hat{V}(\hat{\theta}^*)\hat{\Omega}^{1/2}\hat{M}$. Using (SA.36) and similar arguments for showing (SA.5) and (SA.6), we obtain

$$g_0(\hat{\theta}^*) = g_0(\theta_0) + n^{1/2}(\hat{\theta}^* - \theta_0)\big(Q_0 + o_p(1)\big) = O_p(1) + n^{1/2}(\hat{\theta}^* - \theta_0)O_p(1) \tag{SA.38}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. By (SA.34) and (SA.38),

$$\begin{aligned} \hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^* &= g_0(\hat{\theta}^*) + \hat{V}(\hat{\theta}^*)\big(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta})\big) \\ &= O_p(1) + n^{1/2}(\hat{\theta}^* - \theta_0)O_p(1), \end{aligned} \tag{SA.39}$$

which together with Assumptions 2 and 3(iii), (SA.36), and the Cauchy–Schwarz inequality implies that, for any $j = 1, \ldots, d_\theta$,

$$\begin{aligned} & n^{-1/2}\big(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*\big)'\big(\hat{\Omega}_0^*\big)^{-1}\frac{\partial \hat{\Omega}_0(\hat{\theta}^*)}{\partial \theta_j}\big(\hat{\Omega}_0^*\big)^{-1}\big(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*\big) \\ & \quad = n^{1/2}(\hat{\theta}^* - \theta_0)o_p(1) + o_p(1), \\ & \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}. \end{aligned} \tag{SA.40}$$

By Assumptions 2(iii, iv) and 3(iii),

$$\max_{1 \le j \le d_\theta} \sup_{\theta \in \Theta} \big\| \hat{V}(\theta)/\partial \theta_j \big\| = O_p(1) \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}_0, \tag{SA.41}$$

which combined with (SA.33) implies that

$$\frac{\partial}{\partial \theta'}\hat{V}(\hat{\theta}^*)\big(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta})\big) = O_p(1) \tag{SA.42}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$. By Assumption 1(ii) and (SA.36),

$$n^{-1/2}\frac{\partial g_0(\hat{\theta}^*)}{\partial \theta'} = Q_0 + o_p(1) \quad \text{uniformly over } \mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}. \tag{SA.43}$$

Collecting the results in (SA.42) and (SA.43), we have

$$\begin{aligned}
&n^{-1/2}\frac{\partial}{\partial\theta'}\left(\hat{m}(\hat{\theta}^*)+\hat{V}_S(\hat{\theta}^*)\boldsymbol{v}^*\right)\\
&\quad=n^{-1/2}\frac{\partial}{\partial\theta'}g_0(\hat{\theta}^*)+n^{-1/2}\frac{\partial}{\partial\theta'}\hat{V}(\hat{\theta}^*)\left(\hat{\Omega}^{1/2}\hat{M}\boldsymbol{v}^*-g(\hat{\theta})\right)\\
&\quad=Q_0+o_p(1) \qquad\text{(SA.44)}
\end{aligned}$$

uniformly over $\mathbb{P}\in\mathcal{P}_0\cap\mathcal{P}_{00}$. Apply the first-order expansion to get

$$\begin{aligned}
&\hat{m}(\hat{\theta}^*)+\hat{V}_S(\hat{\theta}^*)\boldsymbol{v}^*\\
&\quad=g_0(\hat{\theta}^*)+\hat{V}(\hat{\theta}^*)\left(\hat{\Omega}^{1/2}\hat{M}\boldsymbol{v}^*-g(\hat{\theta})\right)\\
&\quad=g_0(\theta_0)+\hat{V}(\theta_0)\left(\hat{\Omega}^{1/2}\hat{M}\boldsymbol{v}^*-g(\hat{\theta})\right)\\
&\qquad+\frac{\partial g_0(\tilde{\theta}^*)}{\partial\theta'}(\hat{\theta}^*-\theta_0)+\frac{\partial\hat{V}(\tilde{\theta}^*)}{\partial\theta'}(\hat{\theta}^*-\theta_0)\left(\hat{\Omega}^{1/2}\hat{M}\boldsymbol{v}^*-g(\hat{\theta})\right)\\
&\quad=g_0(\theta_0)+\hat{V}(\theta_0)\left(\hat{\Omega}^{1/2}\hat{M}\boldsymbol{v}^*-g(\hat{\theta})\right)\\
&\qquad+\left(Q_0+o_p(1)\right)n^{1/2}(\hat{\theta}^*-\theta_0)+o_p(1) \qquad\text{(SA.45)}
\end{aligned}$$

uniformly over $\mathbb{P}\in\mathcal{P}_0\cap\mathcal{P}_{00}$, where the third equality is by Assumption 1(ii), (SA.33), (SA.36), and (SA.41). By $V(\theta_0)=S_0$, Lemma A1(b), (e) in Cheng, Dou, and Liao (2022), and (SA.33),

$$\hat{V}(\theta_0)\left(\hat{\Omega}^{1/2}\hat{M}\boldsymbol{v}^*-g(\hat{\theta})\right)=S_0\Omega^{1/2}M\Omega^{-1/2}\left(\Omega^{1/2}\boldsymbol{v}^*-g(\theta_0)\right)+o_p(1) \qquad\text{(SA.46)}$$

uniformly over $\mathbb{P}\in\mathcal{P}_0$. By (SA.45) and (SA.46), we have, uniformly over $\mathbb{P}\in\mathcal{P}_0\cap\mathcal{P}_{00}$,

$$\begin{aligned}
&\hat{m}(\hat{\theta}^*)+\hat{V}_S(\hat{\theta}^*)\boldsymbol{v}^*\\
&\quad=\left(Q_0+o_p(1)\right)n^{1/2}(\hat{\theta}^*-\theta_0)\\
&\qquad+Q_0\left(Q'\Omega^{-1}Q\right)^{-1}Q'\Omega^{-1}g(\theta_0)+S_0\Omega^{1/2}M\boldsymbol{v}^*+o_p(1). \qquad\text{(SA.47)}
\end{aligned}$$

By Assumptions 3(iii) and 4(ii), $Q_0'\Omega_0^{-1}Q_0$ is positive definite. Therefore, collecting the results in (SA.37), (SA.40), (SA.44), (SA.47), and applying Assumption 2(i), (ii) and (SA.36) to $\hat{\Omega}_0^*$, we obtain

$$\begin{aligned}
n^{1/2}(\hat{\theta}^*-\theta_0)&=-\left(Q'\Omega^{-1}Q\right)^{-1}Q'\Omega^{-1}g(\theta_0)\\
&\qquad-\left(Q_0'\Omega_0^{-1}Q_0\right)^{-1}Q_0'\Omega_0^{-1}S_0\Omega^{1/2}M\boldsymbol{v}^*+o_p(1) \qquad\text{(SA.48)}
\end{aligned}$$

uniformly over $\mathbb{P}\in\mathcal{P}_0\cap\mathcal{P}_{00}$, which proves part (a) of the lemma.

(b) By (SA.48), Assumptions 1(iii), 2(iv), 3(iii), and 4(ii),

$$n^{1/2}(\hat{\theta}^*-\theta_0)=O_p(1)\quad\text{uniformly over }\mathbb{P}\in\mathcal{P}_0\cap\mathcal{P}_{00}. \qquad\text{(SA.49)}$$

By Assumptions 2(iv), 3(iii) and 4(ii), (SA.47), and (SA.48), we have, uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$,

$$\begin{aligned}\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^* &= -Q_0(Q_0'\Omega_0^{-1}Q_0)^{-1}Q_0'\Omega_0^{-1}S_0\Omega^{1/2}Mv^* + S_0\Omega^{1/2}Mv^* + o_p(1) \\ &= \Omega_0^{1/2}M_0\Omega_0^{-1/2}S_0\Omega^{1/2}Mv^* + o_p(1) \\ &= \Omega_0^{1/2}M_0\Omega_0^{-1/2}S_0\Omega^{1/2}v^* + o_p(1) = O_p(1), \end{aligned} \tag{SA.50}$$

where $M_0 \equiv I_{k_0} - \Omega_0^{-1/2}Q_0(Q_0'\Omega_0^{-1}Q_0)^{-1}Q_0'\Omega_0^{-1/2}$ and the third equality is by $M_0\Omega_0^{-1/2}Q_0 = 0_{k_0\times 1}$. By Assumptions 2(i), (ii) and 3(iii), (SA.36), and (SA.50), we deduce that, uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$,

$$(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*)'(\hat{\Omega}_0^*)^{-1}(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*) = v^{*\prime}\tilde{M}_0v^* + o_p(1), \tag{SA.51}$$

where $\tilde{M}_0 \equiv (\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2})$. Since $v^* = O_p(1)$, Lemma A1(d) implies that

$$v^{*\prime}\hat{M}v^* = v^{*\prime}Mv^* + o_p(1) \tag{SA.52}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$, which together with (SA.51) finishes the proof.

(c) Since $M \equiv I_k - \Omega^{-1/2}Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1/2}$, $M^2 = M$. Moreover,

$$\begin{aligned}\tilde{M}_0^2 &= (\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2})(\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2}) \\ &= (\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2}) = \tilde{M}_0 \end{aligned} \tag{SA.53}$$

and

$$\begin{aligned}\tilde{M}_0M &= \tilde{M}_0 - (\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2})\Omega^{-1/2}Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1/2} \\ &= \tilde{M}_0, \end{aligned} \tag{SA.54}$$

where the second equality is by $M_0\Omega_0^{-1/2}Q_0 = 0_{k_0\times 1}$. Similarly, $M\tilde{M}_0 = \tilde{M}_0$. Therefore, $(M - \tilde{M}_0)^2 = M^2 - M\tilde{M}_0 - \tilde{M}_0M + \tilde{M}_0^2 = M - \tilde{M}_0$, which implies that $M - \tilde{M}_0$ is an idempotent matrix. The rank of $M - \tilde{M}_0$ equals the trace of $M - \tilde{M}_0$ since $M - \tilde{M}_0$ is idempotent. By the definition of M and $\tilde{M}_0$,

$$\operatorname{tr}(M) = k - \operatorname{tr}(\Omega^{-1/2}Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1/2}) = k - d_\theta \tag{SA.55}$$

and

$$\begin{aligned}\operatorname{tr}(\tilde{M}_0) &= \operatorname{tr}((\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2})) \\ &= \operatorname{tr}(M_0(\Omega_0^{-1/2}S_0\Omega S_0'\Omega_0^{-1/2})) = \operatorname{tr}(M_0) = k_0 - d_\theta, \end{aligned} \tag{SA.56}$$

which implies that $\operatorname{tr}(M - \tilde{M}_0) = k - k_0 = k_1$. Therefore, $M - \tilde{M}_0$ is an idempotent matrix with rank k_1 which together with $v^* \sim N(0, I_k)$ proves the claim (c). *Q.E.D.*

SB. PROOFS FOR BOUNDED LIPSCHITZ CONDITIONS

This section contains the proofs of Lemmas A4 and A5 in Cheng, Dou, and Liao (2022), and the auxiliary results used to show them.

LEMMA SB1: *Consider $\xi \equiv (x', d')'$, where d satisfies Assumptions 2 and 3. Then we have, for any given $C > 0$:*
(i) *$R(\xi)$ is bounded and Lipschitz in ξ on the set $\{\xi : R(\xi) \geq 0$ and $x'\Omega_d^{-1}x \leq C\}$;*
(ii) *$L(v; d)$ is bounded and Lipschitz in d on the set $\{(v, d) : L(v; d) \geq 0$ and $\|v\| \leq C\}$.*

PROOF OF LEMMA SB1: (i) Let $S_\xi \equiv \{\xi : R(\xi) \geq 0 \text{ and } x'\Omega_d^{-1}x \leq C\}$. By the definition of $R(\xi)$, $R(\xi) \leq x'\Omega_d^{-1}x \leq C$ for any $\xi \in S_\xi$, which shows that $R(\xi)$ is bounded on S_ξ. Next, we want to show that for any $\xi_1, \xi_2 \in S_\xi$,

$$\left|R(\xi_1) - R(\xi_2)\right| \leq C_R \|\xi_1 - \xi_2\|_s \tag{SB.1}$$

for some constant C_R. By the triangle inequality,

$$\left|R(\xi_1) - R(\xi_2)\right| \leq R(\xi_1) + R(\xi_2) \leq x_2'\Omega_{d,2}^{-1}x_2 + x_1'\Omega_{d,1}^{-1}x_1 = 2C, \tag{SB.2}$$

which implies that the claimed result holds with a Lipschitz constant $C_R = 2$ if $\|\xi_1 - \xi_2\|_s > C$. Thus, it is only necessary to consider the case that $\|\xi_1 - \xi_2\|_s \leq C$.

Define $A_j(\theta) \equiv m_{d,j}(\theta) + V_{d,j}(\theta)x_j$ for $j = 1, 2$. Consider any $\xi_1, \xi_2 \in S_\xi$, by the triangle inequality,

$$\begin{aligned}
&\left|R(\xi_1) - R(\xi_2)\right| \\
&\quad \leq \left|x_1'\Omega_{d,1}^{-1}x_1 - x_2'\Omega_{d,2}^{-1}x_2\right| \\
&\qquad + \left|\min_{\theta\in\Theta} A_1(\theta)'\big(\Omega_{0,d,1}(\theta)\big)^{-1}A_1(\theta) - \min_{\theta\in\Theta} A_2(\theta)'\big(\Omega_{0,d,2}(\theta)\big)^{-1}A_2(\theta)\right|.
\end{aligned} \tag{SB.3}$$

By the triangle inequality, the Cauchy–Schwarz inequality, $x_1'\Omega_{d,1}^{-1}x_1 \leq C$, and $x_2'\Omega_{d,2}^{-1}x_2 \leq C$,

$$\begin{aligned}
&\left|x_1'\Omega_{d,1}^{-1}x_1 - x_2'\Omega_{d,2}^{-1}x_2\right| \\
&\quad \leq \left|(x_1 - x_2)'\Omega_{d,1}^{-1}x_1\right| + \left|x_2'\Omega_{d,2}^{-1}(\Omega_{d,1} - \Omega_{d,2})\Omega_{d,1}^{-1}x_1\right| + \left|x_2'\Omega_{d,2}^{-1}(x_1 - x_2)\right| \\
&\quad \leq \left[\frac{\big(x_1'\Omega_{d,1}^{-1}x_1\big)^{1/2}}{\big(\lambda_{\min}(\Omega_{d,1})\big)^{1/2}} + \frac{\big(x_2'\Omega_{d,2}^{-1}x_2\big)^{1/2}}{\big(\lambda_{\min}(\Omega_{d,2})\big)^{1/2}}\right]\|x_1 - x_2\| \\
&\qquad + \frac{\big(x_1'\Omega_{d,1}^{-1}x_1\big)^{1/2}\big(x_2'\Omega_{d,2}^{-1}x_2\big)^{1/2}}{\big(\lambda_{\min}(\Omega_{d,1})\lambda_{\min}(\Omega_{d,2})\big)^{1/2}}\|\Omega_{d,1} - \Omega_{d,2}\| \\
&\quad \leq 2c_\lambda^{-1/2}C^{1/2}\|x_1 - x_2\| + c_\lambda^{-1}C\|\Omega_{d,1} - \Omega_{d,2}\| \leq C_1\|\xi_1 - \xi_2\|_s
\end{aligned} \tag{SB.4}$$

for some constant C_1.

Let θ_j denote the minimizer of $A_j(\theta)'(\Omega_{0,d,j}(\theta))^{-1}A_j(\theta)$ for $j = 1, 2$. By the triangle inequality, we have

$$\left|\min_{\theta\in\Theta} A_1(\theta)'\big(\Omega_{0,d,1}(\theta)\big)^{-1}A_1(\theta) - \min_{\theta\in\Theta} A_2(\theta)'\big(\Omega_{0,d,2}(\theta)\big)^{-1}A_2(\theta)\right|$$

$$
\begin{aligned}
&\le \max_{\{\theta_1,\theta_2\}} \left| A_1(\theta)'\left(\Omega_{0,d,1}(\theta)\right)^{-1} A_1(\theta) - A_2(\theta)'\left(\Omega_{0,d,2}(\theta)\right)^{-1} A_2(\theta) \right| \\
&\le \max_{\{\theta_1,\theta_2\}} \left| \left(A_1(\theta) - A_2(\theta)\right)'\left(\Omega_{0,d,1}(\theta)\right)^{-1} A_1(\theta) \right| \\
&\quad + \max_{\{\theta_1,\theta_2\}} \left| A_2(\theta)'\left(\Omega_{0,d,1}(\theta)\right)^{-1}\left(A_1(\theta) - A_2(\theta)\right) \right| \\
&\quad + \max_{\{\theta_1,\theta_2\}} \left| A_2(\theta)'\left(\left(\Omega_{0,d,1}(\theta)\right)^{-1} - \left(\Omega_{0,d,2}(\theta)\right)^{-1}\right) A_2(\theta) \right|. \qquad \text{(SB.5)}
\end{aligned}
$$

We next investigate the three terms after the second inequality of (SB.5) one by one. By the triangle inequality and the Cauchy–Schwarz inequality,

$$
\begin{aligned}
&\max_{\{\theta_1,\theta_2\}} \left\| A_1(\theta) - A_2(\theta) \right\| \\
&\quad \le \sup_{\theta\in\Theta} \left\| m_{d,1}(\theta) - m_{d,2}(\theta) \right\| + \|x_1\| \sup_{\theta\in\Theta} \left\| V_{d,1}(\theta) - V_{d,2}(\theta) \right\| + \sup_{\theta\in\Theta} \left\| V_{d,2}(\theta) \right\| \|x_1 - x_2\| \\
&\quad \le \|\xi_1 - \xi_2\|_s + \lambda_{\max}(\Omega_{d,1})\left(x_1'\Omega_{d,1}^{-1}x_1\right)^{1/2} \|\xi_1 - \xi_2\|_s + C_V \|\xi_1 - \xi_2\|_s, \qquad \text{(SB.6)}
\end{aligned}
$$

where $C_V \equiv c_\lambda^{-1} C_\Omega$ and the second inequality is by the definition of $\|\xi_1 - \xi_2\|_s$ and $\sup_{\theta\in\Theta} \|V_{2,d}(\theta)\| \le c_\lambda^{-1} C_\Omega$ (which is proved in Lemma A1(e) of Cheng, Dou, and Liao (2022)). Therefore, by Assumption 2(iv) and (SB.6),

$$
\max_{\{\theta_1,\theta_2\}} \left\| A_1(\theta) - A_2(\theta) \right\| \le \left(1 + (CC_\Omega)^{1/2} + C_V\right) \|\xi_1 - \xi_2\|_s, \qquad \text{(SB.7)}
$$

which together with the triangle inequality, Assumption 2(iv), the restrictions on ξ_1 and ξ_2, and $\|\xi_1 - \xi_2\|_s \le C$ implies that

$$
\begin{aligned}
&\max_{\{\theta_1,\theta_2\}} \left\| A_1(\theta) \right\| \\
&\quad \le \left\| A_1(\theta_1) \right\| + \left\| A_1(\theta_2) \right\| \\
&\quad \le \left\| A_1(\theta_1) \right\| + \left\| A_2(\theta_2) \right\| + \left\| A_1(\theta_2) - A_2(\theta_2) \right\| \\
&\quad \le C_\Omega^{1/2}\left(\left(A_1(\theta_1)'\left(\Omega_{d,1,0}(\theta_1)\right)^{-1} A_1(\theta_1)\right)^{1/2} + \left(A_2(\theta_2)'\left(\Omega_{d,2,0}(\theta_2)\right)^{-1} A_2(\theta_2)\right)^{1/2}\right) \\
&\qquad + \left(1 + (CC_\Omega)^{1/2} + C_V\right) \|\xi_1 - \xi_2\|_s \\
&\quad \le 2(C_\Omega C)^{1/2} + \left(1 + (CC_\Omega)^{1/2} + C_V\right) C. \qquad \text{(SB.8)}
\end{aligned}
$$

By the same arguments, the inequality in (SB.8) applies to $\max_{\{\theta_1,\theta_2\}} \|A_2(\theta)\|$. By the Cauchy–Schwarz inequality, Assumption 3(iii), (SB.7), and (SB.8),

$$
\begin{aligned}
&\max_{\{\theta_1,\theta_2\}} \left| \left(A_1(\theta) - A_2(\theta)\right)'\left(\Omega_{d,1,0}(\theta)\right)^{-1} A_1(\theta) \right| \\
&\quad \le \max_{\{\theta_1,\theta_2\}} \left(\lambda_{\min}\left(\Omega_{0,d,1}(\theta)\right)\right)^{-1} \left\| A_1(\theta) - A_2(\theta) \right\| \left\| A_1(\theta) \right\| \le C_2 \|\xi_1 - \xi_2\|_s \qquad \text{(SB.9)}
\end{aligned}
$$

for some constant C_2. Similarly, we can show that

$$
\max_{\{\theta_1,\theta_2\}} \left| A_2(\theta)'\left(\Omega_{0,d,1}(\theta)\right)^{-1}\left(A_1(\theta) - A_2(\theta)\right) \right| \le C_2 \|\xi_1 - \xi_2\|_s. \qquad \text{(SB.10)}
$$

By the Cauchy–Schwarz inequality, Assumption 3(iii), and (SB.8),

$$\begin{aligned}
&\max_{\{\theta_1,\theta_2\}}\left|A_2(\theta)'\left(\left(\Omega_{0,d,1}(\theta)\right)^{-1}-\left(\Omega_{0,d,2}(\theta)\right)^{-1}\right)A_2(\theta)\right| \\
&\quad\leq \frac{\max_{\{\theta_1,\theta_2\}}\left\|A_2(\theta)\right\|^2\left\|\Omega_{0,d,1}(\theta)-\Omega_{0,d,2}(\theta)\right\|}{\left(\inf_{\theta\in\Theta}\lambda_{\min}\left(\Omega_{0,d,1}(\theta)\right)\inf_{\theta\in\Theta}\lambda_{\min}\left(\Omega_{0,d,2}(\theta)\right)\right)^{-1}} \\
&\quad\leq C_3\sup_{\theta\in\Theta}\left\|\Omega_{0,d,1}(\theta)-\Omega_{0,d,2}(\theta)\right\| \qquad \text{(SB.11)}
\end{aligned}$$

for some constant C_3. Collecting the results in (SB.5), (SB.9), (SB.10), and (SB.11), we get

$$\begin{aligned}
&\left|\min_{\theta\in\Theta}A_1(\theta)'\left(\Omega_{0,d,1}(\theta)\right)^{-1}A_1(\theta)-\min_{\theta\in\Theta}A_2(\theta)'\left(\Omega_{0,d,2}(\theta)\right)^{-1}A_2(\theta)\right| \\
&\quad\leq (2C_2+C_3)\|\xi_1-\xi_2\|_s, \qquad \text{(SB.12)}
\end{aligned}$$

which together with (SB.3) and (SB.4) implies that the Lipschitz constant is $C_R = C_1 + 2C_2 + C_3$.

(ii) Note that under the condition $L(v;d)\geq 0$ and $\|v\|\leq C$, we have $0\leq L(v;d)\leq v'M_d v\leq C^2$. To show $L(v;d)$ is Lipschitz in d, we write

$$L(v;d)\equiv v'M_d v-\min_{\theta\in\Theta}\left(m_d(\theta)+V_{d,S}(\theta)v\right)'\left(\Omega_{d,0}(\theta)\right)^{-1}\left(m_d(\theta)+V_{d,S}(\theta)v\right), \qquad \text{(SB.13)}$$

where $V_{d,S}(\cdot)\equiv V_d(\cdot)\Omega_d^{1/2}M_d$. This functional form is analogous to $R(\xi)$, with Ω_d and $V_d(\cdot)$ in $R(\xi)$ replaced by M_d and $V_{d,S}(\cdot)$, respectively. Given that $V_{d,S}(\cdot)$ is Lipschitz in d (established in Lemma SC4 below) and $\sup_{\theta\in\Theta}\|V_{d,S}(\theta)\|\leq C_\Omega^{1/2}C_V$ (by Assumptions 2(iv) and 3(iii) and Lemma A1(e) of Cheng, Dou, and Liao (2022)), showing $L(v;d)$ is Lipschitz in d is analogous to showing $R(\xi)$ is Lipschitz in ξ. The only difference is that M_d is not a full rank matrix, unlike Ω_d, which is the reason that we have to bound $\|v\|$ directly instead of bounding $v'M_d v$. Because (SB.4) in the proof of part (i) uses the full rank condition of Ω_d, we replace (SB.4) with the following argument to show $v'M_d v$ is Lipschitz in M_d. Given $\|v\|\leq C$, we have

$$\begin{aligned}
\left|v'M_{d,1}v-v'M_{d,2}v\right| &= \left|v'(M_{d,1}-M_{d,2})v\right| \\
&\leq \|v\|^2\|M_{d,1}-M_{d,2}\|\leq C^2\|M_{d,1}-M_{d,2}\| \qquad \text{(SB.14)}
\end{aligned}$$

by the Cauchy–Schwarz inequality. The rest of the proof is analogous to that in the proof of Lemma SB1(i) and hence is omitted. *Q.E.D.*

PROOF OF LEMMA A4: The truncation function $t_C(u)$ satisfies the following properties: (i) for any $u\in\mathbb{R}^+$, $0\leq t_C(u)\leq 1$ and $ut_C(u)\leq u$; (ii) for any u_1, $u_2\in\mathbb{R}$, $|t_C(u_1)-t_C(u_2)|\leq C^{-1}|u_1-u_2|$, which implies that $t_C(u)$ is Lipschitz in u. Therefore,

$$0\leq R_C(\xi)\leq\left(x'\Omega_d^{-1}x\right)t_C\left(x'\Omega_d^{-1}x\right)\leq C, \qquad \text{(SB.15)}$$

which means that $R_C(\xi)$ is bounded.

Next, we show that $R_C(\xi)$ is Lipschitz in ξ. That is, for any ξ_j with $R(\xi_j) \geq 0$ $(j = 1, 2)$,

$$\left|R_C(\xi_1) - R_C(\xi_2)\right| \leq C_R \|\xi_1 - \xi_2\|_s, \tag{SB.16}$$

where C_R is a constant. Without loss of generality, we assume that $x_2'\Omega_{d,2}^{-1}x_2 \leq x_1'\Omega_{d,1}^{-1}x_1$. By the triangle inequality,

$$\begin{aligned}\left|R_C(\xi_1) - R_C(\xi_2)\right| \leq {} & \left|\left(R(\xi_1) - R(\xi_2)\right)t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right)\right| \\ & + \left|\left(t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right) - t_C\left(x_2'\Omega_{d,2}^{-1}x_2\right)\right)R(\xi_2)\right|. \end{aligned}\tag{SB.17}$$

We have (SB.16) holds with $C_R = C_{R_1} + C_{R_2}$ if we can show that

$$\left|\left(R(\xi_1) - R(\xi_2)\right)t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right)\right| \leq C_{R_1}\|\xi_1 - \xi_2\|_s \tag{SB.18}$$

and

$$\left|\left(t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right) - t_C\left(x_2'\Omega_{d,2}^{-1}x_2\right)\right)R(\xi_2)\right| \leq C_{R_2}\|\xi_1 - \xi_2\|_s \tag{SB.19}$$

for some constants C_{R_1} and C_{R_2}.

We first consider (SB.18). First, note that it holds trivially if $x_1'\Omega_{d,1}^{-1}x_1 > 2C$ because, in this case, $t_C(x_1'\Omega_{d,1}^{-1}x_1) = 0$. Second, given $x_2'\Omega_{d,2}^{-1}x_2 \leq x_1'\Omega_{d,1}^{-1}x_1 \leq 2C$, we deduce that

$$\left|\left(R(\xi_1) - R(\xi_2)\right)t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right)\right| \leq \left|R(\xi_1) - R(\xi_2)\right| \leq C_{R_1}\|\xi_1 - \xi_2\|_s, \tag{SB.20}$$

where the first inequality is by property (i) of $t_C(u)$, and the second inequality is by Lemma SB1(i).

Next, we show (SB.19). First, note that it holds trivially if $2C < x_2'\Omega_{d,2}^{-1}x_2$. In this case,

$$t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right) = t_C\left(x_2'\Omega_{d,2}^{-1}x_2\right) = 0 \tag{SB.21}$$

following the definition of $t_C(u)$. Second, given $x_2'\Omega_{d,2}^{-1}x_2 \leq 2C$, we have

$$\left|\left(t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right) - t_C\left(x_2'\Omega_{d,2}^{-1}x_2\right)\right)R(\xi_2)\right| \leq \left|R(\xi_2)\right| \leq x_2'\Omega_{d,2}^{-1}x_2 \leq 2C. \tag{SB.22}$$

Thus, (SB.19) holds with $C_{R_2} = 1$ if $\|\xi_1 - \xi_2\|_s > 2C$. Third, it remains to consider the case where $x_2'\Omega_{d,2}^{-1}x_2 \leq 2C$ and $\|\xi_1 - \xi_2\|_s \leq 2C$. In this case, Lemma SC3 in Section SC implies that $x_1'\Omega_{d,1}^{-1}x_1 \leq C^*$ for some constant C^*. In this case,

$$\begin{aligned}&\left|\left(t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right) - t_C\left(x_2'\Omega_{d,2}^{-1}x_2\right)\right)R(\xi_2)\right| \\ &\quad \leq \left|t_C\left(x_1'\Omega_{d,1}^{-1}x_1\right) - t_C\left(x_2'\Omega_{d,2}^{-1}x_2\right)\right|x_2'\Omega_{d,2}^{-1}x_2 \leq 2C\left|x_1'\Omega_{d,1}^{-1}x_1 - x_2'\Omega_{d,2}^{-1}x_2\right|\end{aligned} \tag{SB.23}$$

using $0 \leq R(\xi_2) \leq x_2'\Omega_{d,2}^{-1}x_2 \leq 2C$ and $|t_C(u_1) - t_C(u_2)| \leq C^{-1}|u_1 - u_2|$ which follows from property (ii) of $t_C(u)$. Then we can show $|x_1'\Omega_{d,1}^{-1}x_1 - x_2'\Omega_{d,2}^{-1}x_2| \leq C_{R_2}\|\xi_1 - \xi_2\|_S$ for some constant C_{R_2} by the same arguments that show (SB.4) but with $x_1'\Omega_{d,1}^{-1}x_1 \leq 2C$ replaced by $x_1'\Omega_{d,1}^{-1}x_1 \leq C^*$. *Q.E.D.*

LEMMA SB2: *Given $L(v; d) \geq 0$, $\overline{L}_C(v; d)$ is bounded and Lipschitz in d.*

PROOF OF LEMMA SB2: The proof is analogous to that of Lemma A4 with the truncation function $t_C(x'\Omega_d^{-1}x)$ replaced by $t_C(v'M_d v)I\{\|v\|^2 \le C\}$ and Lemma SB1(i) replaced by Lemma SB1(ii), and hence is omitted. *Q.E.D.*

PROOF OF LEMMA A5: Since $0 \le L(v; d) \le v'M_d v$ for any $v \in \mathbb{R}^k$ and $ut_C(u) \le C$ for any $u \ge 0$, we have

$$\overline{L}_C(v; d) = L(v; d)t_C\big(v'M_d v\big)I\big\{\|v\|^2 \le C\big\} \le L(v; d)t_C\big(v'M_d v\big) \le C, \tag{SB.24}$$

which implies that $c_{\alpha,C}(d)$ is bounded. For any v, any d_1 and d_2, by Lemma SB2 there exists a constant C_L such that

$$\big|\overline{L}_C(v; d_1) - \overline{L}_C(v; d_2)\big| \le C_L\|d_1 - d_2\|_s. \tag{SB.25}$$

Since $\overline{L}_C(v; d_1) \ge \overline{L}_C(v; d_2) - C_L\|d_1 - d_2\|_s$ for any v and $P^*(\overline{L}_C(v^*; d_1) > c_{\alpha,C}(d_1)) \le \alpha$, we have $P(\overline{L}_C(v^*; d_2) > c_{\alpha,C}(d_1) + C_L\|d_1 - d_2\|_s) \le \alpha$, which implies that

$$c_{\alpha,C}(d_2) \le c_{\alpha,C}(d_1) + C_L\|d_1 - d_2\|_s. \tag{SB.26}$$

Similarly, we also have

$$c_{\alpha,C}(d_1) \le c_{\alpha,C}(d_2) + C_L\|d_1 - d_2\|_s. \tag{SB.27}$$

Combining (SB.26) and (SB.27), we get

$$\big|c_{\alpha,C}(d_1) - c_{\alpha,C}(d_2)\big| \le C_L\|d_1 - d_2\|_s, \tag{SB.28}$$

which shows the claim of the lemma. *Q.E.D.*

SC. ADDITIONAL AUXILIARY LEMMAS

This section contains the proof of Lemma A3 in Cheng, Dou, and Liao (2022) and some other auxiliary results.

PROOF OF LEMMA A3: For any square matrices A_1 and A_2, let $\mathrm{diag}(A_1, A_2)$ denote the block-diagonal matrix created by aligning the input matrices A_1 and A_2 along the diagonal. Since $\hat{\theta} \in \Theta$ is the minimizer of $g(\theta)'(\hat{\Omega}(\theta))^{-1}g(\theta)$,

$$\begin{aligned} R(\hat{\xi}) &\ge g(\hat{\theta})'\hat{\Omega}^{-1}g(\hat{\theta}) - g_0(\hat{\theta})'\hat{\Omega}_0^{-1}g_0(\hat{\theta}) \\ &= g(\hat{\theta})'\big(\hat{\Omega}^{-1} - \mathrm{diag}\big(\hat{\Omega}_0^{-1}, 0_{k_1\times k_1}\big)\big)g(\hat{\theta}), \end{aligned} \tag{SC.1}$$

where $\hat{\Omega} \equiv \hat{\Omega}(\hat{\theta}, \hat{\theta})$ and $\hat{\Omega}_0$ is the leading $k_0 \times k_0$ submatrix of $\hat{\Omega}$. Let $\hat{\Omega}_{0,1}$ denote the upper-right $k_0 \times k_1$ submatrix of $\hat{\Omega}$ and $\hat{\Omega}_{1,0} \equiv \hat{\Omega}_{0,1}'$. Since $\hat{\Omega}$ is positive definite,

$$\hat{\Omega}\big(\hat{\Omega}^{-1} - \mathrm{diag}\big(\hat{\Omega}_0^{-1}, 0_{k_1\times k_1}\big)\big)\hat{\Omega} = \mathrm{diag}\big(0_{k_0\times k_0}, \hat{\Omega}_1 - \hat{\Omega}_{1,0}\hat{\Omega}_0^{-1}\hat{\Omega}_{0,1}\big) \tag{SC.2}$$

is a positive semi-definite matrix, which together with (SC.1) proves the first claim of the lemma.

To prove the second claim, we first notice that $\hat{m}(\hat{\theta}) = g_0(\hat{\theta}) - \hat{V}(\hat{\theta})g(\hat{\theta}) = 0_{k_0\times 1}$, which implies that, for any $v \in \mathbb{R}^k$,

$$L(v; \hat{d}) \geq v'\hat{M}v - v'\hat{V}_S(\hat{\theta})'\hat{\Omega}_0^{-1}\hat{V}_S(\hat{\theta})v, \tag{SC.3}$$

where $\hat{V}_S(\theta) \equiv S_0\hat{\Omega}^{1/2}\hat{M}$. Therefore,

$$L(x; \hat{d}) \geq v'\hat{M}\hat{\Omega}^{1/2}\big(\hat{\Omega}^{-1} - \mathrm{diag}\big(\hat{\Omega}_0^{-1}, 0_{k_1\times k_1}\big)\big)\hat{\Omega}^{1/2}\hat{M}v \geq 0, \tag{SC.4}$$

where the second inequality holds since the matrix in (SC.2) is positive semi-definite. *Q.E.D.*

LEMMA SC3: *For any ξ_1 and ξ_2 with $x_2'\Omega_{d,2}^{-1}x_2 \leq C$ and $\|\xi_1 - \xi_2\|_s \leq C$, where C is a constant, we have $x_1'\Omega_{d,1}^{-1}x_1 \leq C^*$ for some constant C^*, which depends the constant C of the lemma, and C_Ω and c_λ in Assumptions 2(iv) and 3(iii), respectively.*

PROOF OF LEMMA SC3: Since $\Omega_{d,1}^{-1}$ is symmetric and positive definite under Assumption 3(iii),

$$\begin{aligned} x_1'\Omega_{d,1}^{-1}x_1 &\leq 2x_2'\Omega_{d,1}^{-1}x_2 + 2(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2) \\ &= 2x_2'\Omega_{d,2}^{-1}x_2 + 2x_2'\big(\Omega_{d,1}^{-1} - \Omega_{d,2}^{-1}\big)x_2 + 2(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2) \\ &\leq 2C + 2x_2'\big(\Omega_{d,1}^{-1} - \Omega_{d,2}^{-1}\big)x_2 + 2(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2), \end{aligned} \tag{SC.5}$$

where the second inequality is by $x_2'\Omega_{d,2}^{-1}x_2 \leq C$ as assumed in the lemma. By Assumption 3(iii) and $\|\xi_1 - \xi_2\|_s \leq C$,

$$(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2) \leq \big(\lambda_{\min}(\Omega_{d,1})\big)^{-1}\|x_1 - x_2\|^2 \leq C^2c_\lambda^{-1}. \tag{SC.6}$$

Similarly, by Assumption 3(iii) and $\|\xi_1 - \xi_2\|_s \leq C$,

$$\begin{aligned} \big|x_2'\big(\Omega_{d,1}^{-1} - \Omega_{d,2}^{-1}\big)x_2\big|^2 &= \big|x_2'\Omega_{d,1}^{-1}(\Omega_{d,1} - \Omega_{d,2})\Omega_{d,2}^{-1}x_2\big|^2 \\ &\leq \big(x_2'\Omega_{d,1}^{-2}x_2\big)\big(x_2'\Omega_{d,2}^{-2}x_2\big)\|\Omega_{d,1} - \Omega_{d,2}\|^2 \\ &\leq \frac{\lambda_{\max}(\Omega_{d,2})\big(x_2'\Omega_{d,2}^{-1}x_2\big)^2}{\lambda_{\min}(\Omega_{d,2})\big(\lambda_{\min}(\Omega_{d,1})\big)^2}\|\Omega_{d,1} - \Omega_{d,2}\|^2 \\ &\leq C^4c_\lambda^{-3}C_\Omega, \end{aligned} \tag{SC.7}$$

where the last inequality is by $x_2'\Omega_{d,2}^{-1}x_2 \leq C$, Assumptions 2(iv) and 3(iii), and $\|\xi_1 - \xi_2\|_s \leq C$. The claim of the lemma follows from (SC.5)–(SC.7). *Q.E.D.*

LEMMA SC4: *For any ξ_1 and ξ_2, define $V_{d,S,j}(\theta) \equiv V_{d,j}(\theta)\Omega_{d,j}^{1/2}M_{d,j}$ for $j = 1, 2$. Then, we have*

$$\sup_{\theta\in\Theta}\big\|V_{d,S,1}(\theta) - V_{d,S,2}(\theta)\big\| \leq C_\Omega^{1/2}\big(1 + c_\lambda^{-1}C_\Omega^{1/2} + c_\lambda^{-1}C_\Omega\big)\|\xi_1 - \xi_2\|_s, \tag{SC.8}$$

where C_Ω and c_λ are in Assumptions 2(iv) and 3(iii), respectively.

PROOF OF LEMMA SC4: By definition,

$$\begin{aligned} & V_{d,S,1}(\theta) - V_{d,S,2}(\theta) \\ & \quad = \left[V_{d,1}(\theta) - V_{d,2}(\theta)\right]\Omega_{d,1}^{1/2} M_{d,1} \\ & \qquad + V_{d,2}(\theta)\left(\Omega_{d,1}^{1/2} - \Omega_{d,2}^{1/2}\right) M_{d,1} + V_{d,2}(\theta)\Omega_{d,2}^{1/2}(M_{d,1} - M_{d,2}). \end{aligned} \tag{SC.9}$$

By the properties of $\Omega_{d,j}$ and $M_{d,j}$,

$$\sup_{\theta\in\Theta}\left\|\left[V_{d,1}(\theta) - V_{d,2}(\theta)\right]\Omega_{d,1}^{1/2} M_{d,1}\right\| \leq C_{\Omega}^{1/2} \sup_{\theta\in\Theta}\left\|V_{d,1}(\theta) - V_{d,2}(\theta)\right\|. \tag{SC.10}$$

By the properties of $V_{d,j}(\theta)$ and $M_{d,j}$, and (7.2.13) in Horn and Johnson (1990),

$$\begin{aligned} \sup_{\theta\in\Theta}\left\|V_{d,2}(\theta)\left(\Omega_{d,1}^{1/2} - \Omega_{d,2}^{1/2}\right) M_{d,1}\right\| &\leq \sup_{\theta\in\Theta}\left\|V_{d,2}(\theta)\right\| \left\|\Omega_{d,1}^{1/2} - \Omega_{d,2}^{1/2}\right\|_S \\ &\leq c_{\lambda}^{-1} C_{\Omega} \left\|\Omega_{d,2}^{-1/2}\right\|_S \|\Omega_{d,1} - \Omega_{d,2}\| \\ &\leq c_{\lambda}^{-2} C_{\Omega} \|\Omega_{d,1} - \Omega_{d,2}\|. \end{aligned} \tag{SC.11}$$

Similarly,

$$\sup_{\theta\in\Theta}\left\|V_{d,2}(\theta)\Omega_{d,2}^{1/2}(M_{d,1} - M_{d,2})\right\| \leq c_{\lambda}^{-1} C_{\Omega}^{3/2} \|M_{d,1} - M_{d,2}\|. \tag{SC.12}$$

The desirable result follows by (SC.9)–(SC.12) and the triangle inequality. *Q.E.D.*

SD. THEORETICAL POWER PROPERTIES OF THE NEW TEST

In this section, we investigate the power properties of the conditional specification test in two cases. First, when the asset pricing moments are globally misspecified, we show that the conditional specification test rejects these moments wpa1, and thus is consistent regardless of the identification strength in the baseline moments. Second, when baseline moments provide strong identification and the asset pricing moments are locally misspecified, we show that the conditional test has the same asymptotic local power as the C test. Thus, it shares the power optimality of the C test in standard scenarios.

ASSUMPTION SD1: *The following conditions hold for any* $\mathbb{P} \in \mathcal{P}_{1,\infty} \subset \mathcal{P}$: (i) $\inf_{\theta\in\Theta} \|\mathbb{E}[\bar{g}_1(\theta)]\| > c_{g_1}$ *for some* $c_{g_1} > 0$; (ii) $\lambda_{\min}(\Omega_0(\theta_0)) \geq c_\lambda$, $\lambda_{\min}(\hat{\Omega}) \geq c_\lambda$, *and* $\lambda_{\min}(\hat{Q}'\hat{Q}) \geq c_\lambda$ *wpa*1.

Assumption SD1(i) implies that there are globally misspecified moments in $\mathbb{E}[\bar{g}_1(\theta_0)] = 0_{k_1\times 1}$. Assumption SD1(ii) requires that the eigenvalues of $\hat{\Omega}$ and $\hat{Q}'\hat{Q}$ are bounded away from zero wpa1. In view of Assumptions 1(ii) and 2(i), this condition holds if the eigenvalues of $\Omega(\theta_1)$ and $Q(\theta_1)Q(\theta_1)'$ are bounded away from zero, where θ_1 denotes the pseudo true value under misspecification. Therefore, Assumption SD1(ii) is the counterpart of Assumption 3(iii) under the alternative.

THEOREM SD1: *Suppose Assumptions* 1, 2, *and* SD1 *hold*. *For any* $\mathbb{P} \in \mathcal{P}_{1,\infty}$, $\mathbb{P}(\mathcal{T} > c_\alpha(\hat{d})) \to 1$ *as* $n \to \infty$.

PROOF OF THEOREM SD1: We first show that the test statistic $\mathcal{T}$, written as $R(\hat{\xi})$, diverges at rate n under global misspecification. By Assumptions 2(i), (iv) and SD1(ii),

$$R(\hat{\xi}) = \min_{\theta\in\Theta} g(\theta)'\big(\hat{\Omega}(\theta)\big)^{-1} g(\theta) - \min_{\theta\in\Theta} g_0(\theta)'\big(\hat{\Omega}_0(\theta)\big)^{-1} g_0(\theta)$$
$$\geq (C_\Omega + 1)^{-1} \min_{\theta\in\Theta} \big\| g(\theta)\big\|^2 - c_\lambda^{-1} \big\| g_0(\theta_0)\big\|^2 \quad \text{wpa1}, \tag{SD.1}$$

where

$$\big\| g(\theta)\big\|^2 \geq \frac{1}{2} \big\| \mathbb{E}\big[g(\theta)\big]\big\|^2 - \big\| g(\theta) - \mathbb{E}\big[g(\theta)\big]\big\|^2. \tag{SD.2}$$

By Assumption SD1(i), there exists a constant $c_{g_1} > 0$ such that $\min_{\theta\in\Theta} \|\mathbb{E}[\bar{g}(\theta)]\|^2 \geq c_{g_1}$, which combined with (SD.1), (SD.2), and Assumptions 1(i) and 2(iv) implies that

$$R(\hat{\xi}) \geq n\Big(K^{-1} \min_{\theta\in\Theta} \big\|\mathbb{E}\big[\bar{g}(\theta)\big]\big\|^2 - o_p(1)\Big) \geq n c_{g_1} K^{-1} \quad \text{wpa1}. \tag{SD.3}$$

The critical value satisfies $c_\alpha(\hat{d}) \leq q_{1-\alpha}(\chi_k^2)$ wpa1, because $L(v^*; \hat{d}) \leq v^{*\prime}\hat{M}v^* \leq \|v^*\|^2$ wpa1 given that $\hat{M}$ is an idempotent matrix wpa1 under Assumption SD1(ii) and $q_{1-\alpha}(\chi_k^2)$ is the $1-\alpha$ quantile of $\|v^*\|^2$. Therefore, by (SD.3) and $c_\alpha(\hat{d}) \leq q_{1-\alpha}(\chi_k^2)$ wpa1, we have

$$\mathbb{P}\big(R(\hat{\xi}) > c_\alpha(\hat{d})\big) \geq \mathbb{P}\big(n c_{g_1} K^{-1} > q_{1-\alpha}\big(\chi_k^2\big)\big) - o(1), \tag{SD.4}$$

where the right-hand side of the above inequality goes to 1 as $n \to \infty$. *Q.E.D.*

The consistency of the conditional specification test holds no matter whether the parameter θ_0 (or its subvector) is strongly, weakly, or not identified by the baseline moments. We next study the local power of the conditional specification test when the baseline moments provide strong identification.

ASSUMPTION SD2: *The following conditions hold for any* $\mathbb{P} \in \mathcal{P}_{1,A} \subset \mathcal{P}$:
(i) $\mathbb{E}[\bar{g}_1(\theta_0)] = an^{-1/2}$ *for some* $a \in \mathbb{R}^{k_1}$ *with* $\|a\| < \infty$;
(ii) *Assumptions* 3(*ii*) *and* 3(*iii*) *hold for any* $\mathbb{P} \in \mathcal{P}_{1,A}$.

THEOREM SD2: *Suppose Assumptions* 1, 2, 4, *and* SD2 *hold*. *For any* $\mathbb{P} \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}$, *we have*

$$\mathbb{P}\big(\mathcal{T} > c_\alpha(\hat{d})\big) \to P\big(\chi_{k_1}^2\big(a_\Omega' M a_\Omega\big) > q_{1-\alpha}\big(\chi_{k_1}^2\big)\big), \quad \text{as } n \to \infty,$$

where $a_\Omega \equiv \Omega^{-1/2}a$ *and* $\chi_{k_1}^2(a_\Omega' M a_\Omega)$ *denotes a non-central chi-square random variable with degree of freedom* k_1 *and non-central parameter* $a_\Omega' M a_\Omega$.

PROOF OF THEOREM SD2: Under Assumptions 1 and 2, the strong identification in baseline moments in Assumption 4, and the local misspecification in Assumption SD2, $\hat{\theta}$ and $\hat{\theta}_0$ are consistent by the standard arguments and results in (A.19) and (A.20) of Cheng, Dou, and Liao (2022) remain valid. Therefore,

$$R(\hat{\xi}) \to_d \big(\Omega^{-1/2}v + a_\Omega\big)' M \big(\Omega^{-1/2}v + a_\Omega\big) - v_0'\Omega_0^{-1/2} M_0 \Omega_0^{-1/2} v_0, \tag{SD.5}$$

where $a_\Omega \equiv \Omega^{-1/2}a$, v is a multivariate normal random variable with mean zero and variance Ω, and v_0 denotes the leading k_0 subvector of v. By the standard arguments in the GMM literature (e.g., Hall (2005, Section 5)), we have

$$\big(\Omega^{-1/2}v + a_\Omega\big)' M \big(\Omega^{-1/2}v + a_\Omega\big) - v_0'\Omega_0^{-1/2} M_0 \Omega_0^{-1/2} v_0 \sim \chi^2_{k_1}\big(a_\Omega' M a_\Omega\big). \tag{SD.6}$$

We next study $c_\alpha(\hat{d})$ under the local misspecification. Since $\hat{\theta}$ is $n^{1/2}$ consistent under the local misspecification, Lemma A7 of Cheng, Dou, and Liao (2022) remains valid for any $\mathbb{P} \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}$. Therefore, for any $\mathbb{P} \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}$,

$$L\big(v^*, \hat{d}\big) = v^{*\prime}(M - \tilde{M}_0)v^* + o_p(1) \sim \chi^2_{k_1}. \tag{SD.7}$$

By (SD.7) and arguments analogous to those used to show Theorem 2(ii), we have $c_\alpha(\hat{d}) \to_p q_{1-\alpha}(\chi^2_{k_1})$, which together with (SD.5) and (SD.6) proves the claim of the theorem. *Q.E.D.*

As long as $a_\Omega' M a_\Omega > 0$, we have $P(\chi^2_{k_1}(a_\Omega' M a_\Omega) > q_{1-\alpha}(\chi^2_{k_1})) > \alpha$. Moreover, this probability is strictly increasing in the non-central parameter $a_\Omega' M a_\Omega$. If the baseline moments $\mathbb{E}[\bar{g}_0(\theta_0)] = 0_{k_0\times 1}$ only depend on a subvector $\theta_{c,0}$ of θ_0 with dimension d_c and strongly identify $\theta_{c,0}$, arguments analogous to those used to show Theorem SD2 give

$$\mathbb{P}\big(\mathcal{T} > c_\alpha(\hat{d})\big) \to P\big(\chi^2_{k_1+d_c-d_\theta}\big(a_\Omega' M a_\Omega\big) > q_{1-\alpha}\big(\chi^2_{k_1+d_c-d_\theta}\big)\big) \quad \text{as } n \to \infty. \tag{SD.8}$$

When the baseline moments provide strong identification, the conditional specification test is asymptotically equivalent to the C test following Theorems 2, SD1, and SD2. [1] In particular, it shares the same (asymptotic) local power function with the C test and thus achieves optimality under local misspecification (Newey (1985)). Nevertheless, the conditional specification test compares favorably to the C test for its correct asymptotic size even with weak identification in the baseline moments, an important property for its applications to many macro-finance asset pricing models.

SE. COMPARISON TO SOME POWER ENVELOPES

In this section, we derive some power envelopes in a Gaussian experiment as in Section 4.1 of Cheng, Dou, and Liao (2022). These power envelopes are akin to those in Section 3.4 of Andrews and Mikusheva (2016). We compare the power of the proposed conditional specification test to these power envelopes through simulation studies.

Setup. We observe (i) a Gaussian process $g_{0,\infty}(\cdot)$ with covariance matrix $\Omega_0(\cdot,\cdot)$, and (ii) a Gaussian random vector $g_\infty(\hat{\theta})$ which satisfies

$$g_\infty(\hat{\theta}) \equiv \big(I_k - Q\big(Q'\Omega^{-1}Q\big)^{-1}Q'\Omega^{-1}\big)g_\infty(\theta_0) = \Omega^{1/2}M\Omega^{-1/2}g_\infty(\theta_0), \tag{SE.1}$$

where $g_\infty(\theta_0) \equiv (g_{0,\infty}(\theta_0)', g_{1,\infty}(\theta_0)')'$ is normal with covariance matrix Ω, $g_{0,\infty}(\theta_0)$ and $g_{1,\infty}(\theta_0)$ are $k_0 \times 1$ and $k_1 \times 1$, respectively, $Q \equiv (Q_0', Q_1')'$, Q_0 and Q_1 are $k_0 \times d_\theta$ and $k_1 \times d_\theta$ ($k_1 \geq d_\theta$) matrices, respectively. We assume that Q_1 has full rank, and $\Omega_0(\cdot,\cdot)$, Ω, Q, and the covariance between $g_{0,\infty}(\cdot)$ and $g_\infty(\theta_0)$ are known.

[1]See, for example, Hall (2005) for detailed derivations for the C test.

We are interested in testing

$$H_0 : \eta = 0_{k_1\times 1} \quad \text{where } \eta \equiv \mathbb{E}\big[g_{1,\infty}(\theta_0)\big], \tag{SE.2}$$

while maintaining $\mathbb{E}[g_{0,\infty}(\theta_0)] = 0_{k_0\times 1}$ under both the null and the alternative hypotheses. The alternative hypothesis is written as

$$H_1 : \eta \neq 0. \tag{SE.3}$$

The true value of θ_0 is unknown under both the null and the alternative.

Power Envelopes. Let $Q^{\perp}$ denote the orthogonal complement of Q. It is clear that $Q'\Omega^{-1}$ and $Q^{\perp\prime}$ are the left eigenvectors of $I_k - Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}$ with respect to the (left) eigenvalues 0 and 1, respectively. Let $D = (Q^{\perp}, \Omega^{-1}Q)'$; then D is non-singular. Moreover,

$$Dg_\infty(\hat{\theta}) = D\big(I_k - Q\big(Q'\Omega^{-1}Q\big)^{-1}Q'\Omega^{-1}\big)g_\infty(\theta_0) = \begin{pmatrix} Q^{\perp\prime}g_\infty(\theta_0) \\ 0_{d_\theta\times k}\end{pmatrix}. \tag{SE.4}$$

Based on (SE.4), observing $g_\infty(\hat{\theta})$ is equivalent to observing

$$Y \equiv Q^{\perp\prime}g_\infty(\theta_0) \sim N\big(A(\eta), Q^{\perp\prime}\Omega Q^{\perp}\big), \quad \text{where } A(\eta) \equiv Q^{\perp\prime}\begin{pmatrix} 0_{k_0\times 1} \\ \eta \end{pmatrix}. \tag{SE.5}$$

We next consider inference of η based only on Y.

Since the uniformly most powerful (UMP) test does not exist for (SE.3), we follow Andrews and Mikusheva (2016) and derive several power envelopes by reducing the alternative hypothesis (SE.3) and/or imposing restrictions on the class of tests. If the alternative hypothesis (SE.3) is reduced to a single value η^* with $A(\eta^*) \neq 0$, then the Neyman–Pearson lemma implies that the UMP test rejects H_0 if

$$\frac{A(\eta^*)'\big(Q^{\perp\prime}\Omega Q^{\perp}\big)^{-1}Y}{\big(A(\eta^*)'\big(Q^{\perp\prime}\Omega Q^{\perp}\big)^{-1}A(\eta^*)\big)^{1/2}} > z_{1-\alpha}, \tag{SE.6}$$

where $z_{1-\alpha}$ denotes the $1-\alpha$ quantile of the standard normal distribution. It is clear that the optimality of the test in (SE.6) depends on η^* by construction. Its power may be low if the true value η under the alternative is different from η^*. In the simulation study below, we let the test in (SE.6) depend on the true value under the alternative η (i.e., we replace η^* by η) and call its power (as a function of η) as PE-1. Next, we consider a subset of alternative hypothesis (SE.3) which is proportional to a known vector η^* with $A(\eta^*) \neq 0$, that is, $H_1 : \eta = a\eta^*$. Since η^* is known, the subset of alternative hypothesis becomes $H_1 : a \neq 0$. As noticed in Andrews and Mikusheva (2016), the UMP unbiased test for this reduced problem rejects H_0 if

$$\left|\frac{A(\eta^*)'\big(Q^{\perp\prime}\Omega Q^{\perp}\big)^{-1}Y}{\big(A(\eta^*)'\big(Q^{\perp\prime}\Omega Q^{\perp}\big)^{-1}A(\eta^*)\big)^{1/2}}\right| > z_{1-\alpha/2}. \tag{SE.7}$$

In the simulation study below, we let the test in (SE.7) also depend on the true value η under the alternative and call its power (as a function of η) as PE-2. Both PE-1 and

PE-2 are infeasible because they require the knowledge of the true value η under the alternative. Finally, the following feasible test:

$$Y'\big(Q^{\perp\prime}\Omega Q^{\perp}\big)^{-1}Y > q_{1-\alpha}\big(\chi^2_{k-d_\theta}\big) \tag{SE.8}$$

is the UMP invariant test, whose power function is called PE-3. This is equivalent to the J test.

Conditional Specification Test. In this setup, the test statistic $\mathcal{T}$ in the paper is the QLR statistic written as

$$\mathcal{T} \equiv g_\infty(\hat{\theta})'\Omega^{-1}g_\infty(\hat{\theta}) - \min_{\theta\in\Theta} g_{0,\infty}(\theta)'\big(\Omega_0(\theta)\big)^{-1}g_{0,\infty}(\theta), \tag{SE.9}$$

where $\Omega_0(\theta) \equiv \Omega_0(\theta, \theta)$. We apply the conditional inference based on this test statistic. Define

$$m_{0,\infty}(\theta) \equiv g_{0,\infty}(\theta) - V(\theta)g_\infty(\hat{\theta}), \tag{SE.10}$$

where $V(\theta) \equiv \mathrm{Cov}(g_{0,\infty}(\theta), g_\infty(\theta_0))\Omega^{-1}$ is a known function of θ. Then, under the null hypothesis, $\mathrm{Cov}(m_{0,\infty}(\theta), g_\infty(\hat{\theta})) = 0$, which implies that $m_{0,\infty}(\theta)$ and $g_\infty(\hat{\theta})$ are independent by their joint normal distribution. The conditional inference is conducted using the critical value of

$$\begin{aligned}\mathcal{T} &= g_\infty(\hat{\theta})'\Omega^{-1}g_\infty(\hat{\theta}) \\ &\quad - \min_{\theta\in\Theta}\big(m_{0,\infty}(\theta) + V(\theta)g_\infty(\hat{\theta})\big)'\big(\Omega_0(\theta)\big)^{-1}\big(m_{0,\infty}(\theta) + V(\theta)g_\infty(\hat{\theta})\big)\end{aligned} \tag{SE.11}$$

conditioning on $m_{0,\infty}(\theta)$.

Simulation. Next, we compare the power of the proposed test with the three power envelopes through simulation studies. To this end, we consider a specific example where $d_\theta = 1$, $k_0 = qk_1$, $k_1 = 2$, and

$$g_{0,\infty}(\theta) \equiv g_{0,\infty}(\theta_0) + (\theta - \theta_0)Q_{0,\infty}, \tag{SE.12}$$

where $Q_{0,\infty}$ is a $k_0 \times 1$ random vector. The distribution of the random vector $(g_\infty(\theta_0)', Q_{0,\infty}')'$ is specified as follows:

$$\begin{pmatrix} g_{0,\infty}(\theta_0) \\ g_{1,\infty}(\theta_0) \\ Q_{0,\infty} \end{pmatrix} \sim N\left(\begin{pmatrix} 0_{k_0\times 1} \\ \eta \\ \mu_g \end{pmatrix}, \Sigma\right), \quad \text{where } \Sigma \equiv \begin{pmatrix} \Omega_{00} & \Omega_{01} & \Omega_{0g} \\ \Omega_{10} & \Omega_{11} & \Omega_{1g} \\ \Omega_{g0} & \Omega_{g1} & \Omega_{gg} \end{pmatrix}, \tag{SE.13}$$

μ_g is a $k_0 \times 1$ real vector. We shall consider two cases for $Q_{0,\infty}$. In the first case, $Q_{0,\infty}$ is a non-random vector as in the simple disaster risk model in Section 2 of Cheng, Dou, and Liao (2022). In this case, Ω_{gg}, Ω_{g1}, Ω_{1g}, Ω_{0g}, and Ω_{g0} are zero matrices, and $Q_{0,\infty} = \mu_g$. In the second case, $Q_{0,\infty}$ is a non-degenerate normal random vector.

By the definition of $g_{0,\infty}(\theta)$ and the joint distribution of $(g_\infty(\theta_0)', Q_{0,\infty}')'$, $\Omega_0(\theta)$ and $V(\theta)$, both of which show up in the conditional specification test, take the following form:

$$\Omega_0(\theta) = \Omega_{00} + (\theta - \theta_0)(\Omega_{0g} + \Omega_{g0}) + (\theta - \theta_0)^2\Omega_{gg},$$

$$V(\theta) = \begin{pmatrix}\Omega_{00} & \Omega_{01}\end{pmatrix}\Omega^{-1} + (\theta - \theta_0)\begin{pmatrix}\Omega_{g0} & \Omega_{g1}\end{pmatrix}\Omega^{-1}, \quad \text{where } \Omega = \begin{pmatrix}\Omega_{00} & \Omega_{01} \\ \Omega_{10} & \Omega_{11}\end{pmatrix}. \tag{SE.14}$$

We generate the covariance matrix Ω as follows:

$$\Omega = \begin{pmatrix}(1+\rho^2)^{-1}(I_{k_0} + \rho^2 1_{q\times q} \otimes \Omega_u) & \rho 1_{q\times 1} \otimes \Omega_u \\ \rho 1_{1\times q} \otimes \Omega_u & \Omega_u\end{pmatrix}, \quad \text{where } \Omega_u = \begin{pmatrix}1 & \lambda \\ \lambda & 1\end{pmatrix}, \tag{SE.15}$$

where $A \otimes B$ denotes the Kronecker product of two real matrices A and B. The parameter λ determines the correlation between the two moments in $g_{1,\infty}(\theta_0)$, while ρ mainly controls the correlations between moments in $g_{0,\infty}(\theta_0)$ and $g_{1,\infty}(\theta_0)$. In the case that $Q_{0,\infty}$ is a non-degenerate normal random vector, we let

$$\Omega_{gg} = I_{k_0}, \qquad \Omega_{0g} = \Omega'_{g0} = \lambda I_{k_0}, \quad \text{and} \quad \Omega_{1g} = \Omega'_{g1} = \lambda 1_{1\times q} \otimes I_{k_1}, \tag{SE.16}$$

where we also use λ to control the correlation between $Q_{0,\infty}$ and $g_\infty(\theta_0)$.

Throughout this simulation, we let $\theta_0 = 0$, $\Theta = [-1, 1]$, $Q_0 = c_g 1_{k_0\times 1}$, $Q_1 = (j^{-1})_{j=1,\dots,k_1}$, $\mu_g = c_\mu 1_{k_0\times 1}$, $\eta = a 1_{k_1\times 1}$, and $\lambda = 0.1$. We consider a benchmark case and three deviations from the benchmark, which are defined as follows:

Benchmark case: $\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, non-random $Q_{0,\infty}$, $q = k_0/k_1 = 1$ or 2;
Deviation case 1: $\rho = 0.2$ or 0.8, and $q = 2$;
Deviation case 2: $c_g = 0.1$;
Deviation case 3: random $Q_{0,\infty}$.

In the benchmark case, we set $c_g = 0$ to model weak baseline moments whose derivatives Q_0 are 0 in the limiting experiment. The deviation cases enable us to investigate how the power properties of the conditional test change when: (1) the baseline moments and the asset pricing moments have correlation; (2) the baseline moments provide nontrivial identification when combined with the asset pricing moments; (3) the matrix $Q_{0,\infty}$ is random. The simulation results in the benchmark case and in the three deviation cases are presented in Figure S1 and Figures S2–S4, respectively. We plot the finite-sample rejection probability against a, where $\eta = (a, a)'$ under the alternative. All the results are calculated with 10,000 simulation replications.

Discussion. In all cases, the power of the proposed conditional specification test is between PE-2 and PE-3 (J test). PE-2 is the power of the UMP unbiased test with respect to a smaller subset of the general alternative hypothesis in (SE.3) and it is constructed using the true alternative value η, whereas the conditional specification test does not require such information. Simulation results show that the power function of the conditional specification test is rather close to PE-2 in many cases with a substantial improvement from PE-3. The benchmark case in Figure S1 shows that increasing the number of baseline moments significantly enlarges the power gain compared to PE-3 while roughly maintaining the same amount of power loss compared to PE-2. Figure S2 and Figure S3 show that increasing the correlation between the baseline moments and the asset-pricing moments,

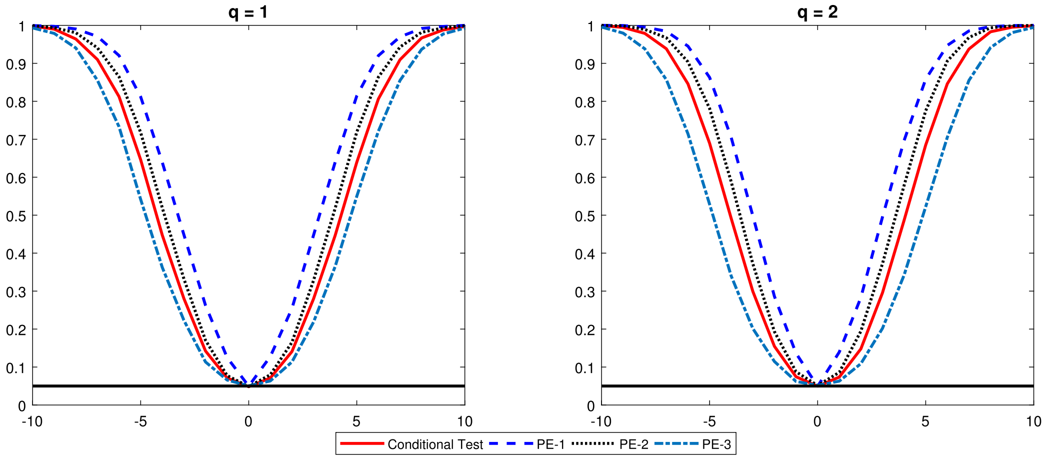

FIGURE S1.—Power comparison in the benchmark case. *Note*: In the benchmark case, we have $\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, non-random $Q_{0,\infty}$, $q = k_0/k_1 = 1$ or 2.

or increasing the identification strength of the baseline moments to the structural parameter, make all powers higher and reduce the power difference between the conditional specification test and PE-2. Figure S4 shows that reducing the signal-to-noise ratio in the baseline moments results in a larger gap between the power of the conditional specification test and PE-2. Nevertheless, we still see noticeable improvement over PE-3 in the two scenarios of this case.

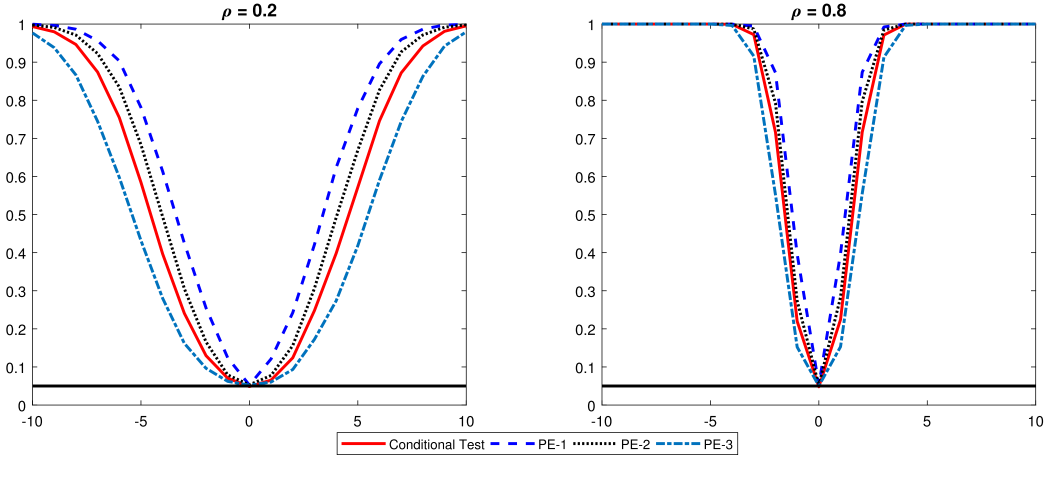

FIGURE S2.—Power comparison in the deviation case 1. *Note*: In the deviation case 1, we have $\rho = 0.2$ or 0.8, $c_g = 0$, $c_\mu = 1$, non-random $Q_{0,\infty}$, $q = 2$.

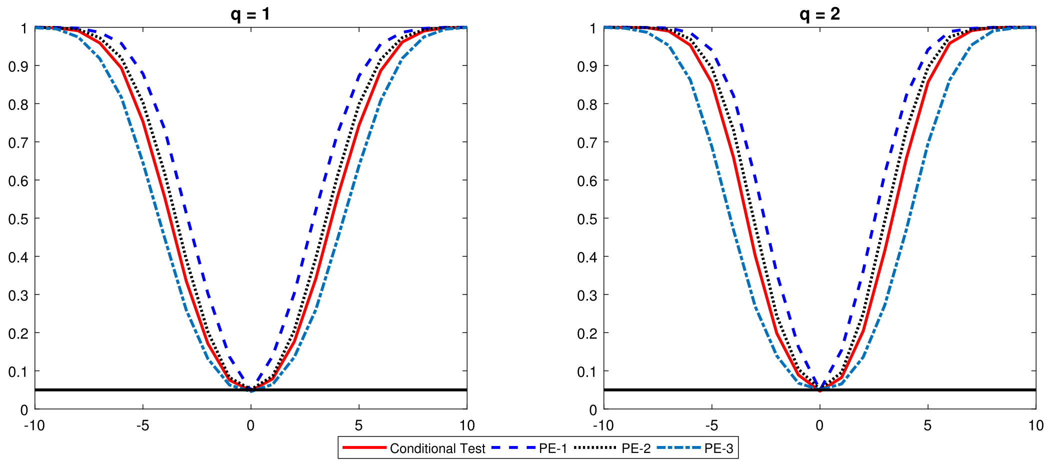

FIGURE S3.—Power comparison in the deviation case 2. *Note*: In the deviation case 2, we have $c_g = 0.1$, $\rho = 0.4$, $c_\mu = 1$, non-random $Q_{0,\infty}$, $q = 1$ or 2.

SF. ADDITIONAL DETAILS OF THE EMPIRICAL APPLICATION

We have eight baseline moment conditions $\mathbb{E}[\bar{g}_0(\theta)] = 0_{8\times 1}$ when $\theta = \theta_0$, where $\theta \equiv (\theta_1, \dots, \theta_4)$ is the reparameterized parameter defined as

$$\theta_1 \equiv \frac{p}{\alpha - \gamma}, \quad \theta_2 \equiv \frac{\sigma_p^2}{1-\rho^2}, \quad \theta_3 \equiv \rho, \quad \text{and} \quad \theta_4 \equiv \gamma. \tag{SF.1}$$

In the model, $\bar{g}_0(\theta)$ only depends on a subvector of θ. We have six asset pricing moment conditions $\mathbb{E}[\bar{g}_1(\theta_0)] = 0_{6\times 1}$, where $\bar{g}_1(\theta)$ depends on all the components in θ.

We consider the following calibrated values for the nuisance parameters:

$$(\delta, g_c, g_d, \sigma_c, \phi, \underline{v}, q) = (0.97, 0.02, 0.02, 0.02, 3.5, 0.07, 0.4). \tag{SF.2}$$

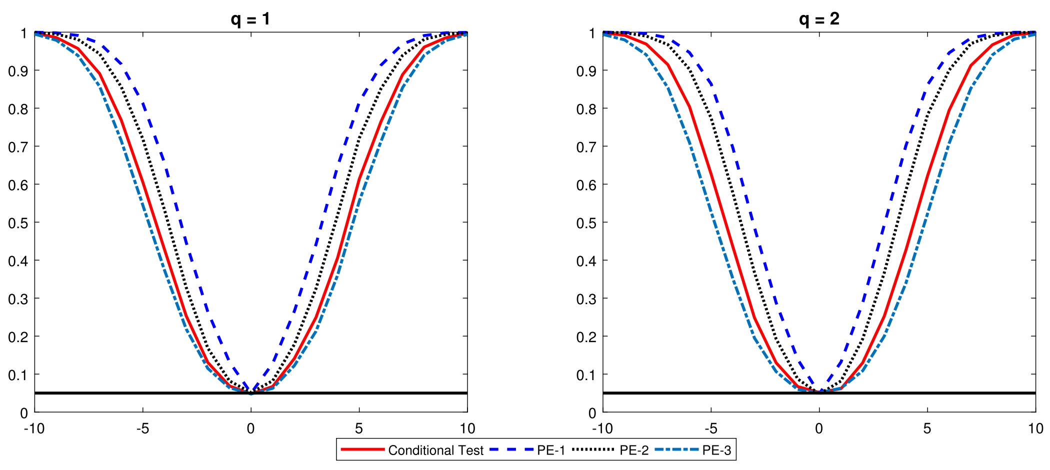

FIGURE S4.—Power comparison in the deviation case 3. *Note*: In the deviation case 4, we have random $Q_{0,\infty}$, $\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, and $q = 1$ or 2.

We consider $p \in \{0.3\%, 0.5\%, 0.7\%, 0.9\%, 1.1\%\}$, where $p = 0.7\%$ is our benchmark case and the other four values of p are used for the robustness check. The parameter space Θ for the unknown parameter is set to $\Theta \equiv \Theta_1 \times \Theta_2 \times \Theta_3 \times \Theta_4$, where

$$\begin{aligned} &\Theta_1 \equiv [0.001, 0.02], \qquad \Theta_2 \equiv [5, 12], \\ &\Theta_3 \equiv [0.95, 0.999], \quad \text{and} \quad \Theta_4 \equiv [3, 6]. \end{aligned} \tag{SF.3}$$

To compute the CUE estimator, the J statistic, and the statistic of the conditional specification test, we search through equally spaced grid points with step size (i.e., the distance between two adjacent points) 0.001 in Θ_1 and Θ_3, and step size 0.01 in Θ_2 and Θ_4.[2] The critical values of the conditional specification test are simulated using $B = 2500$ Gaussian random vectors.

To calculate the model uncertainty set for $p \in \{0.5\%, 0.7\%, 0.9\%\}$, we consider a smaller parameter space $\Theta_2 \equiv [5, 8]$ and a larger step size 0.1 of the grid points in Θ_2 and Θ_4 to reduce the computational cost. The parameter spaces Θ_j $(j = 1, 3, 4)$ and the grid points in Θ_1 and Θ_3 are unchanged. The reduced space Θ_2 still covers the CUE estimators of θ_2 for the three values of p considered. The model uncertainty sets of (η_1, η_3) and (η_3, η_4) are calculated through grid search with equally spaced grid points for η_j $(j = 1, 3, 4)$ with step size 0.001. The parameter spaces for η_j $(j = 1, 3, 4)$ are set large enough such that the model uncertainty sets from the J test are contained in the interior of these parameter spaces.

REFERENCES

ANDREWS, I., AND A. MIKUSHEVA (2016): "Conditional Inference with a Functional Nuisance Parameter," *Econometrica*, 84 (4), 1571–1612. [17,18]

CHENG, X., W. W. DOU, AND Z. LIAO (2022): "Macro-Finance Decoupling: Robust Evaluations of Macro Asset Pricing Models," *Econometrica*. [1,4,5,7,9-11,13,16,17,19]

HALL, A. R. (2005): *Generalized Method of Moments*. Oxford University Press. [17]

HORN, R. A., AND C. R. JOHNSON (1990): *Matrix Analysis*. Cambridge University Press. [3,15]

NEWEY, W. K. (1985): "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, 29 (3), 229–256. [17]

Co-editor Ulrich K. Müller handled this manuscript.

Manuscript received 26 May, 2020; final version accepted 27 August, 2021; available online 1 September, 2021.

[2]We have also considered a much larger parameter space with $\Theta_1 \equiv [0.001, 1]$, $\Theta_2 \equiv [1, 15]$, $\Theta_3 \equiv [0.9, 0.999]$, $\Theta_4 \equiv [1, 20]$, and step size 0.002, 0.2, 0.001, 0.2, respectively. The results on the CUE estimators and the J tests are similar to those reported in Table I of the paper.