Supplemental Appendix to
Macro-Finance Decoupling: Robust Evaluations of
Macro Asset Pricing Models

Xu Cheng*, Winston Wei Dou†, Zhipeng Liao‡

August 23, 2021

Abstract

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 (2021). Section SA provides the proofs of several lemmas on the asymptotic convergence of the random components in the test statistic $T$ 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.

*Department of Economics, University of Pennsylvania; Email: xucheng@econ.upenn.edu.
†Finance Department, The Wharton School, University of Pennsylvania; Email: wdou@wharton.upenn.edu.
‡Department of Economics, University of California, Los Angeles; Email: zhipeng.liao@econ.ucla.edu.
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 (2021) under their Assumptions 1, 2, 3 and 4. The proofs of Lemmas A3, A4 and A5 in the appendix to Cheng, Dou, and Liao (2021) are in Sections SB and SC of this supplemental appendix. Throughout this supplemental appendix, 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 $C_{1,n} \equiv \sup_{\theta \in \Theta} |\bar{g}(\theta)'(\hat{\Omega}(\theta))^{-1}\bar{g}(\theta) - G(\theta)'(\Omega(\theta))^{-1}G(\theta)|$ where $G(\theta) \equiv E[g(\theta)]$. Then by Assumptions 1(i, iii), 2(i) and 3(iii),

$$C_{1,n} = o_p(1), \quad \text{(SA.1)}$$
uniformly over $\mathbb{P} \in \mathcal{P}_0$. Consider any $\varepsilon > 0$. By the definition of $\hat{\theta}$ and $\theta_0$,

$$\sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{P}\left(||\hat{\theta} - \theta_0|| \geq \varepsilon\right) \leq \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{P}\left(\min_{\theta \in B_{c}(\theta_0)} \bar{g}(\theta)'(\hat{\Omega}(\theta))^{-1}\bar{g}(\theta) \leq \bar{g}(\theta_0)'(\hat{\Omega}(\theta_0))^{-1}\bar{g}(\theta_0)\right) \leq \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{P}\left(\min_{\theta \in B_{c}(\theta_0)} G(\theta)'(\Omega(\theta))^{-1}G(\theta) \leq 2C_{1,n}\right) \leq \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{P}\left(c_1^{-1}\delta_2^2 \leq 2C_{1,n}\right), \quad \text{(SA.2)}$$

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

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

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

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

Applying the first order expansion, we get

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

where $\tilde{\theta}$ is the mean value between $\theta_0$ and $\hat{\theta}$ and it may vary across rows. By Assumption 1(ii), the consistency of $\tilde{\theta}$ and (SA.4),

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

2
Similarly, we can show that
\[ q(\hat{\theta}) = Q(\hat{\theta}) + o_p(1) = Q + o_p(1) = O_p(1), \text{ uniformly over } \mathbb{P} \in \mathcal{P}_0. \]  
(SA.7)

By Assumption 2(i, ii) and the consistency of \( \hat{\theta} \),
\[ \hat{\Omega} \equiv \hat{\Omega}(\hat{\theta}) = \Omega + o_p(1) \]  
(SA.8)
uniformly over \( \mathbb{P} \in \mathcal{P}_0 \). Applying the chain rule, we get the first order condition of \( \hat{\theta} \):
\[
0_{d_\theta \times 1} = 2q(\hat{\theta})'\hat{\Omega}^{-1}g(\hat{\theta}) - \left( n^{-1/2}g(\hat{\theta})'\hat{\Omega}^{-1}\frac{\partial \hat{\Omega}(\hat{\theta})}{\partial \theta_j} \hat{\Omega}^{-1}g(\hat{\theta}) \right)_{j=1,\ldots,d_\theta}, \tag{SA.9}
\]
where \((a_j)_{j=1,\ldots,d_\theta} \equiv (a_1,\ldots,a_{d_\theta})'\) for any real numbers \(a_1,\ldots,a_{d_\theta}\). By Assumptions 1(i, iii), 2(iv) and 3(iii), the consistency of \( \hat{\theta} \), (SA.5), (SA.6), (SA.7) and (SA.8),
\[
q(\hat{\theta})'\hat{\Omega}^{-1}g(\hat{\theta}) = q(\hat{\theta})'\hat{\Omega}^{-1}g(\theta_0) + q(\hat{\theta})'\hat{\Omega}^{-1}q(\hat{\theta})n^{1/2}(\hat{\theta} - \theta_0) \\
= Q'\Omega^{-1}g(\theta_0) + (Q'\Omega^{-1}Q + o_p(1)) n^{1/2}(\hat{\theta} - \theta_0) + o_p(1) \tag{SA.10}
\]
uniformly over \( \mathbb{P} \in \mathcal{P}_0 \). Similarly, we can show that for \( j = 1,\ldots,d_\theta \),
\[
n^{-1/2}g(\hat{\theta})'\hat{\Omega}^{-1}\frac{\partial \hat{\Omega}(\hat{\theta})}{\partial \theta_j} \hat{\Omega}^{-1}g(\hat{\theta}) = n^{1/2}(\hat{\theta} - \theta_0)o_p(1) + o_p(1) \text{ uniformly over } \mathbb{P} \in \mathcal{P}_0. \tag{SA.11}
\]

Combining the results in (SA.9), (SA.10), (SA.11), and applying Assumptions 1(iii) and 3(iii), we deduce that
\[
n^{1/2}(\hat{\theta} - \theta_0) = -(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}g(\theta_0) + o_p(1) = O_p(1) \tag{SA.12}
\]
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
\[
q(\hat{\theta})'\hat{\Omega}^{-1}q(\hat{\theta}) = Q'\Omega^{-1}Q + o_p(1) \text{ uniformly over } \mathbb{P} \in \mathcal{P}_0, \tag{SA.14}
\]
which together with Assumption 1(iii), 2(iv), and 3(iii) implies that
\[
\lim_{n \to \infty} \inf_{P \in \mathcal{P}_0} \mathbb{P} \left( K^{-1} \leq \lambda_{\min}(q(\hat{\theta})'(\hat{\Omega})^{-1}q(\hat{\theta})) \leq \lambda_{\max}(q(\hat{\theta})'(\hat{\Omega})^{-1}q(\hat{\theta})) \leq K \right) = 1. \tag{SA.15}
\]
Let \( \| \cdot \|_S \) denote the matrix operator norm. By Exercise 7.2.18 in Horn and Johnson (1990),
\[
\left\| \hat{\Omega}^{1/2} - \Omega^{1/2} \right\|_S \leq \left\| \hat{\Omega} - \Omega \right\| \left\| \Omega^{-1} \right\|_S, \tag{SA.16}
\]
which together with Assumption 3(iii) and (SA.8) implies that
\[
\left\| \hat{\Omega}^{1/2} - \Omega^{1/2} \right\|_S = o_p(1) \text{ uniformly over } P \in \mathcal{P}_0. \tag{SA.17}
\]
By (SA.17) and the relation between the operator norm and the Frobenius norm,
\[
\left\| \hat{\Omega}^{1/2} - \Omega^{1/2} \right\| = o_p(1) \text{ uniformly over } P \in \mathcal{P}_0. \tag{SA.18}
\]
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 \( P \in \mathcal{P}_0 \). The claim in the lemma follows by (SA.14), (SA.15) and (SA.19).

(e) By Assumption 2(i, ii) and the consistency of \( \hat{\theta} \),
\[
\sup_{\theta \in \Theta} \left\| \hat{\Omega}(\theta, \hat{\theta}) - \Omega(\theta, \theta_0) \right\| = o_p(1) \text{ uniformly over } P \in \mathcal{P}_0. \tag{SA.20}
\]
By (SA.13) and (SA.20),
\[
\sup_{\theta \in \Theta} \left\| (\hat{\Omega}(\theta, \hat{\theta}) - \Omega(\theta, \theta_0)) \Omega^{-1} \right\| \leq (\lambda_{\min}(\hat{\Omega}))^{-1} \sup_{\theta \in \Theta} \left\| \hat{\Omega}(\theta, \hat{\theta}) - \Omega(\theta, \theta_0) \right\| = o_p(1) \tag{SA.21}
\]
uniformly over \( P \in \mathcal{P}_0 \), where the inequality is by the Cauchy-Schwarz inequality. Similarly,
\[
\sup_{\theta \in \Theta} \left\| \Omega(\theta, \theta_0)(\hat{\Omega}^{-1} - \Omega^{-1}) \right\| = \sup_{\theta \in \Theta} \left\| \Omega(\theta, \theta_0)\hat{\Omega}^{-1}(\hat{\Omega} - \Omega)\Omega^{-1} \right\| \\
\leq (\lambda_{\min}(\hat{\Omega})\lambda_{\min}(\Omega))^{-1} \sup_{\theta, \tilde{\theta} \in \Theta} \left\| \Omega(\theta, \tilde{\theta}) \right\| \left\| \hat{\Omega} - \Omega \right\| = o_p(1) \tag{SA.22}
\]
uniformly over \( P \in \mathcal{P}_0 \), 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)
Proof of Lemma A6. The desirable result follows from (SA.27), (SA.28), and the triangle inequality.

\[
\sup_{\theta \in \Theta} \left\| \hat{V}(\theta) - V(\theta) \right\|
\leq \sup_{\theta \in \Theta} \left\| (\hat{\Omega}(\theta, \hat{\theta}) - \Omega(\theta, \theta_0))\hat{\Omega}^{-1} \right\|
+ \sup_{\theta \in \Theta} \left\| \Omega(\theta, \theta_0)(\hat{\Omega}^{-1} - \Omega^{-1}) \right\| = o_p(1)
\]  

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

\[
\sup_{\theta \in \Theta} \left\| V(\theta) \right\| \leq \sup_{\theta \in \Theta} \left\| S_0 \Omega(\theta, \theta_0)\Omega^{-1} \right\| \leq (\lambda_{\min}(\Omega))^{-1} \sup_{\theta \in \Theta} \left\| \Omega(\theta, \theta_0) \right\| \leq c_\lambda^{-1} C_{\Omega}
\]  

which finishes the proof. Q.E.D.

Proof of Lemma A2. To link \( \hat{\xi} \) and \( \xi^* \), we first define

\[
\hat{\xi} \equiv (\hat{v}', \hat{m} (\cdot)', \text{vec}(V (\cdot))', \text{vech}(\Omega (\cdot))', \text{vech}(M (\cdot))')',
\]  

where \( \hat{v} \equiv \Omega^{1/2} M \Omega^{-1/2} g(\theta_0) \) and \( \hat{m} (\cdot) \equiv g_0 (\cdot) - V (\cdot) \hat{v} \). The difference between \( \hat{\xi} \) and \( \xi^* \) lies in the first two elements, where \( \hat{v} \) and \( \hat{m} (\cdot) \) in \( \hat{\xi} \) involve the empirical process \( g (\cdot) \), but \( \Omega^{1/2} M \psi^* \) and \( m^* (\cdot) \) in \( \xi^* \) 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

\[
\lim_{n \to \infty} \sup_{P \in \mathcal{P}_0} \sup_{f \in BL_1} \left\| \mathbb{E}[f(g) - \mathbb{E}[g]] - E \left[ f(\psi) \right] \right\| = 0.
\]  

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

\[
\lim_{n \to \infty} \sup_{P \in \mathcal{P}_0} \sup_{f \in BL_1} \left\| \mathbb{E}[f(\hat{\xi}_1)] - E \left[ f(\xi^*) \right] \right\| = 0.
\]  

Next, note that the difference between \( \hat{\xi} \) and \( \hat{\xi}_1 \) 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 (2021). Thus, we have \( \hat{\xi} = \hat{\xi}_1 + o_p(1) \) uniformly over \( P \in \mathcal{P}_0 \) following Lemma A1(c, d, e) in Cheng, Dou, and Liao (2021), which implies

\[
\lim_{n \to \infty} \sup_{P \in \mathcal{P}_0} \sup_{f \in BL_1} \left\| \mathbb{E}[f(\hat{\xi})] - E \left[ f(\xi^*) \right] \right\| = 0.
\]  

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

Proof of Lemma A6. Since \( 0 \leq L(v; d^*) \leq v' M v \) for any \( v \in \mathbb{R}^k \) and \( u t_C (u) \leq C \) for any \( u \geq 0 \),
we have
\[
|L_C(v; d^*) - \mathcal{L}_C(v; d^*)| = L(v; d^*)t_C(v'Mv)I\{\|v\|^2 > C\} \leq CI\{\|v\|^2 > C\} \quad \text{(SA.29)}
\]
for any \(v \in \mathbb{R}^k\), which implies that
\[
P\left(|L_C(v^*, d^*) - \mathcal{L}_C(v^*, d^*)| > \varepsilon\right) \leq P\left(I\{\|v^*\|^2 > C\} > \varepsilon/C\right) \leq P\left(\|v^*\|^2 > C\right). \quad \text{(SA.30)}
\]

Since \(\|v^*\|^2\) follows the chi-square distribution with degree of freedom \(k\), there exists a constant \(C_\delta\) such that \(P(\|v^*\|^2 > C_\delta) \leq \delta/4\) which together with (SA.30) implies that for any \(C \geq C_\delta\)
\[
P\left(|L_C(v^*, d^*) - \mathcal{L}_C(v^*, d^*)| > \varepsilon\right) \leq \delta/4. \quad \text{(SA.31)}
\]

By the union bound of probability and (SA.31), we have for any \(C \geq C_\delta\),
\[
P(L_C(v^*, d^*) > c_{a,C}(d^*) + \varepsilon) \leq P(\mathcal{L}_C(v^*, d^*) + |L_C(v^*, d^*) - \mathcal{L}_C(v^*, d^*)| > c_{a,C}(d^*) + \varepsilon) \leq P(\mathcal{L}_C(v^*, d^*) > c_{a,C}(d^*)) + \delta/4 \leq \alpha + \delta/4, \quad \text{(SA.32)}
\]
where the last inequality is by the definition of \(c_{a,C}(d^*)\). Since \(R_C(\xi^*) = L_C(v^*, d^*)\) by definition, the claim of the lemma follows from (SA.32).

**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 (2021), and (SA.18), we have uniformly over \(\mathbb{P} \in \mathcal{P}_0\),
\[
\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta}) = \Omega^{1/2}M\Omega^{-1/2}(\Omega^{1/2}v^* - g(\theta_0)) + o_p(1) = O_p(1), \quad \text{(SA.33)}
\]
which together with Lemma A1(c) in Cheng, Dou, and Liao (2021) implies that
\[
\sup_{\theta \in \Theta} \left\|\hat{V}(\theta)(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta}))\right\| = O_p(1) \text{ uniformly over } \mathbb{P} \in \mathcal{P}_0. \quad \text{(SA.34)}
\]

Therefore by the triangle inequality, Assumption 1 and (SA.34),
\[
\sup_{\theta \in \Theta} \left\|n^{-1/2}(\hat{m}(\theta) + \hat{V}(\theta)(\hat{\Omega}^{1/2}\hat{M}v^*) - \mathbb{E}[g_0(\theta)])\right\|
\leq \sup_{\theta \in \Theta} \left\|n^{-1/2}g_0(\theta) - \mathbb{E}[g_0(\theta)]\right\| + \sup_{\theta \in \Theta} \left\|n^{-1/2}\hat{V}(\theta)(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta}))\right\| = o_p(1) \quad \text{(SA.35)}
\]
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
\[
\hat{\theta}^* = \theta_0 + o_p(1) \quad \text{(SA.36)}
\]
uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$. Applying the chain rule, we get the first order condition of $\hat{\theta}^*$:

$$
0_{d_\theta \times 1} = 2 \left( \frac{\partial}{\partial \theta^*} (\hat{m}(\hat{\theta}) + \hat{V}_S(\hat{\theta})v^*) \right)' \left( \Omega_0^* \right)^{-1} (\hat{m}(\hat{\theta}) + \hat{V}_S(\hat{\theta})v^*) - \left( \left( \frac{\partial}{\partial \theta^*} (\hat{m}(\hat{\theta}) + \hat{V}_S(\hat{\theta})v^*) \right)' \left( \Omega_0^* \right)^{-1} (\hat{m}(\hat{\theta}) + \hat{V}_S(\hat{\theta})v^*) \right)_{j=1, \ldots, d_\theta},
$$

(SA.37)

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

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

(SA.38)

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

$$
\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^* = g_0(\hat{\theta}^*) + \hat{V}(\hat{\theta}^*)(\hat{\Omega}_0^*/2\hat{\Omega} - g(\hat{\theta})) = O_p(1) + n^{1/2} (\hat{\theta}^* - \theta_0)O_p(1),
$$

(SA.39)

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

$$
n^{-1/2} (\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*)'(\hat{\Omega}_0^*)^{-1} \frac{\partial \hat{\Omega}_0(\hat{\theta}^*)}{\partial \theta_j} (\hat{\Omega}_0^*)^{-1} (\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*)
$$

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

(SA.40)

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

$$
\max_{1 \leq j \leq d_\theta} \sup_{\theta \in \Theta} \left\| \hat{V}(\theta)/\partial \theta_j \right\| = O_p(1) \text{ uniformly over } \mathbb{P} \in \mathcal{P}_0,
$$

(SA.41)

which combined with (SA.33) implies that

$$
\frac{\partial}{\partial \theta^*} \hat{V}(\theta)^*(\hat{\Omega}_0^*/2\hat{\Omega} - g(\hat{\theta})) = O_p(1)
$$

(SA.42)

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

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

(SA.43)
Collecting the results in (SA.42) and (SA.43), we have
\[
\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta}(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*)
= \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta}g_0(\hat{\theta}^*) + n^{-1/2} \frac{\partial}{\partial \theta'} \hat{V}(\hat{\theta}^*)(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta})) = Q_0 + o_p(1) \tag{SA.44}
\]
uniformly over \( P \in \mathcal{P}_0 \cap \mathcal{P}_{00} \). Applying the first order expansion to get
\[
\hat{m}(\hat{\theta}^*) + \hat{V}(\hat{\theta}^*)v^*
= g_0(\hat{\theta}^*) + \hat{V}(\hat{\theta}^*)(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta}))
\]
\[
+ \frac{\partial g_0(\hat{\theta}^*)}{\partial \theta'}(\hat{\theta}^* - \theta_0) + \frac{\partial \hat{V}(\hat{\theta}^*)}{\partial \theta'}(\hat{\theta}^* - \theta_0)(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta}))
= g_0(\theta_0) + \hat{V}(\theta_0)(\hat{\Omega}^{1/2}\hat{M}v^* - g(\hat{\theta})) + (Q_0 + o_p(1))n^{1/2}(\hat{\theta}^* - \theta_0) + o_p(1) \tag{SA.45}
\]
uniformly over \( P \in \mathcal{P}_0 \). By (SA.45) and (SA.46), we have uniformly over \( P \in \mathcal{P}_0 \cap \mathcal{P}_{00} \),
\[
\hat{m}(\hat{\theta}^*) + \hat{V}(\hat{\theta}^*)v^* = (Q_0 + o_p(1))n^{1/2}(\hat{\theta}^* - \theta_0)
+ Q_0(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}g(\theta_0) + S_0\Omega^{1/2}Mv^* + o_p(1). \tag{SA.47}
\]
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
\[
n^{1/2}(\hat{\theta}^* - \theta_0) = -(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}g(\theta_0) - (Q_0'\Omega_0^{-1}Q_0)^{-1}Q_0'\Omega_0^{-1}S_0\Omega^{1/2}Mv^* + o_p(1) \tag{SA.48}
\]
uniformly over \( P \in \mathcal{P}_0 \cap \mathcal{P}_{00} \), which proves part (a) of the lemma.

(b) By (SA.48), Assumptions 1(iii), 2(iv), 3(iii) and 4(ii),
\[
n^{1/2}(\hat{\theta}^* - \theta_0) = O_p(1) \text{ uniformly over } P \in \mathcal{P}_0 \cap \mathcal{P}_{00}. \tag{SA.49}
\]
By Assumptions 2(iv), 3(iii) and 4(ii), (SA.47) and (SA.48), we have uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$,

$$\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^* = -Q_0(Q_0'\Omega_0^{-1}Q_0)^{-1}Q_0'\Omega_0^{-1}S_0\Omega_1^{1/2}Mv^* + S_0\Omega_1^{1/2}Mv^* + o_p(1)$$

$$= \Omega_0^{1/2}M_0\Omega_0^{-1/2}S_0\Omega_1^{1/2}Mv^* + o_p(1)$$

$$= \Omega_0^{1/2}M_0\Omega_0^{-1/2}S_0\Omega_1^{1/2}v^* + o_p(1) = O_p(1), \quad (SA.50)$$

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 $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$,

$$(\hat{m}(\hat{\theta}^*) + \hat{V}_S(\hat{\theta}^*)v^*)' (\hat{\Omega}_0^{1/2} - \hat{V}_S(\hat{\theta}^*)v^*) = v^* M_0 v^* + o_p(1), \quad (SA.51)$$

where $M_0 \equiv (\Omega_0^{-1/2}S_0\Omega_1^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2})$. Since $v^* = O_p(1)$, Lemma A1(d) implies that

$$v^* M_0 v^* = v^* M v^* + o_p(1) \quad (SA.52)$$

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,

$$\hat{M}_0^2 = (\Omega_0^{-1/2}S_0\Omega_1^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2})(\Omega_0^{-1/2}S_0\Omega_1^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2})$$

$$= (\Omega_0^{-1/2}S_0\Omega_1^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2}) = \hat{M}_0 \quad (SA.53)$$

and

$$\hat{M}_0 M = \hat{M}_0 - (\Omega_0^{-1/2}S_0\Omega_1^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2})\Omega^{-1/2}Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1/2}$$

$$= \hat{M}_0, \quad (SA.54)$$

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

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

and

$$tr(\hat{M}_0) = tr((\Omega_0^{-1/2}S_0\Omega_1^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2}))$$

$$= tr(M_0(\Omega_0^{-1/2}S_0\Omega_1^{1/2})) = tr(M_0) = k_0 - d_\theta, \quad (SA.56)$$

which implies that $tr(M - \hat{M}_0) = k - k_0 = k_1$. Therefore, $M - \hat{M}_0$ 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 (2021), 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 $\xi$ 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 $S_\xi \equiv \{\xi : R(\xi) \geq 0$ 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 $S_\xi$. Next, we want to show that for any $\xi_1, \xi_2 \in S_\xi$,

$$|R(\xi_1) - R(\xi_2)| \leq C_R \|\xi_1 - \xi_2\|_s$$  \hfill (SB.57)

for some constant $C_R$. By the triangle inequality,

$$|R(\xi_1) - R(\xi_2)| \leq R(\xi_1) + R(\xi_2) \leq x'_2 \Omega_d^{-1} x_2 + x'_1 \Omega_d^{-1} x_1 = 2C,$$  \hfill (SB.58)

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,

$$|R(\xi_1) - R(\xi_2)| \leq \left| x'_1 \Omega_{d,1}^{-1} x_1 - x'_2 \Omega_{d,2}^{-1} x_2 \right|$$

$$+ \left| \min_{\theta \in \Theta} A_1(\theta)'(\Omega_{0,d,1}(\theta))^{-1} A_1(\theta) - \min_{\theta \in \Theta} A_2(\theta)'(\Omega_{0,d,2}(\theta))^{-1} A_2(\theta) \right|. \hfill (SB.59)$$

By the triangle inequality, the Cauchy-Schwarz inequality, $x'_1 \Omega_{d,1}^{-1} x_1 \leq C$ and $x'_2 \Omega_{d,2}^{-1} x_2 \leq C$,

$$\left| x'_1 \Omega_{d,1}^{-1} x_1 - x'_2 \Omega_{d,2}^{-1} x_2 \right| \leq \left| (x_1 - x_2)' \Omega_{d,1}^{-1} x_1 \right| + \left| x'_2 \Omega_{d,2}^{-1} (\Omega_{d,1} - \Omega_{d,2}) \Omega_{d,1}^{-1} x_1 \right| + \left| x'_2 \Omega_{d,2}^{-1} x_2 - x_2 \right|$$

$$\leq \left[ (x_1\Omega_{d,1}^{-1} x_1)^{1/2} + (x'_2 \Omega_{d,2}^{-1} x_2)^{1/2} \right] \frac{1}{\lambda_{\min}(\Omega_{d,1})^{1/2}} \|x_1 - x_2\|$$

$$+ \left[ (x'_1 \Omega_{d,1}^{-1} x_1)^{1/2} (x'_2 \Omega_{d,2}^{-1} x_2)^{1/2} \right] \frac{1}{\lambda_{\min}(\Omega_{d,1})^{1/2}} \|\Omega_{d,1} - \Omega_{d,2}\|$$

$$\leq 2e_\chi^{1/2} C^{1/2} \|x_1 - x_2\| + e_\chi C \|\Omega_{d,1} - \Omega_{d,2}\| \leq C_1 \|\xi_1 - \xi_2\|_s \hfill (SB.60)$$

for some constant $C_1$. 

10
Let $\theta_j$ 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

$$
\left|\min_{\theta \in \Theta} A_1(\theta)'(\Omega_{0,d,1}(\theta))^{-1}A_1(\theta) - \min_{\theta \in \Theta} A_2(\theta)'(\Omega_{0,d,2}(\theta))^{-1}A_2(\theta)\right|
\leq \max_{\{\theta_1, \theta_2\}} |A_1(\theta)'(\Omega_{0,d,1}(\theta))^{-1}A_1(\theta) - A_2(\theta)'(\Omega_{0,d,2}(\theta))^{-1}A_2(\theta)|
\leq \max_{\{\theta_1, \theta_2\}} \left| (A_1(\theta) - A_2(\theta))'(\Omega_{0,d,1}(\theta))^{-1}A_1(\theta) \right|
+ \max_{\{\theta_1, \theta_2\}} \left| A_2(\theta)'(\Omega_{0,d,1}(\theta))^{-1}(A_1(\theta) - A_2(\theta)) \right|
+ \max_{\{\theta_1, \theta_2\}} \left| A_2(\theta)'((\Omega_{0,d,1}(\theta))^{-1} - (\Omega_{0,d,2}(\theta))^{-1})A_2(\theta) \right|.
$$

(SB.61)

We next investigate the three terms after the second inequality of (SB.61) one by one. By the triangle inequality and the Cauchy-Schwarz inequality,

$$
\max_{\{\theta_1, \theta_2\}} \|A_1(\theta) - A_2(\theta)\|
\leq \sup_{\theta \in \Theta} \|m_{d,1}(\theta) - m_{d,2}(\theta)\| + \|x_1\| \sup_{\theta \in \Theta} \|V_{d,1}(\theta) - V_{d,2}(\theta)\| + \sup_{\theta \in \Theta} \|V_{d,2}(\theta)\| \|x_1 - x_2\|
\leq \|\xi_1 - \xi_2\|_s + \lambda_{\max}(\Omega_{d,1})(x_1'\Omega_{d,1}^{-1}x_1)^{1/2}\|\xi_1 - \xi_2\|_s + C_V\|\xi_1 - \xi_2\|_s,
$$

(SB.62)

where $C_V \equiv c_\lambda^{-1}C_\Omega$ and the second inequality is by the definition of $\|\xi_1 - \xi_2\|_s$ and $\sup_{\theta \in \Theta} \|V_{d,2}(\theta)\| \leq c_\lambda^{-1}C_\Omega$ (which is proved in Lemma A1(e) of Cheng, Dou, and Liao (2021)). Therefore, by Assumption 2(iv) and (SB.62),

$$
\max_{\{\theta_1, \theta_2\}} \|A_1(\theta) - A_2(\theta)\| \leq (1 + (CC_\Omega)^{1/2} + C_V)\|\xi_1 - \xi_2\|_s,
$$

(SB.63)

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

$$
\max_{\{\theta_1, \theta_2\}} \|A_1(\theta)\| \leq \|A_1(\theta_1)\| + \|A_1(\theta_2)\|
\leq \|A_1(\theta_1)\| + \|A_2(\theta_2)\| + \|A_1(\theta_2) - A_2(\theta_2)\|
\leq C_\Omega^{1/2}\left( (A_1(\theta_1)'(\Omega_{d,1,0}(\theta_1))^{-1}A_1(\theta_1))^{1/2} + (A_2(\theta_2)'(\Omega_{d,2,0}(\theta_2))^{-1}A_2(\theta_2))^{1/2} \right)
+ (1 + (CC_\Omega)^{1/2} + C_V)\|\xi_1 - \xi_2\|_s
\leq 2(C_\Omega C)^{1/2} + (1 + (CC_\Omega)^{1/2} + C_V)C.
$$

(SB.64)

By the same arguments, the inequality in (SB.64) applies to $\max_{\{\theta_1, \theta_2\}} \|A_2(\theta)\|$. By the Cauchy-
Schwarz inequality, Assumption 3(iii), (SB.63) and (SB.64),

\[
\max_{\{\theta_1, \theta_2\}} \left| (A_1(\theta) - A_2(\theta))' (\Omega_{d,1,0}(\theta))^{-1} A_1(\theta) \right|
\]
\[
\leq \max_{\{\theta_1, \theta_2\}} (\lambda_{\min}(\Omega_{0,d,1}(\theta)))^{-1} \|A_1(\theta) - A_2(\theta)\| \|A_1(\theta)\| \leq C_2 \|\xi_1 - \xi_2\|_s \tag{SB.65}
\]

for some constant \(C_2\). Similarly, we can show that

\[
\max_{\{\theta_1, \theta_2\}} \left| A_2(\theta)'(\Omega_{0,d,1}(\theta))^{-1} (A_1(\theta) - A_2(\theta)) \right| \leq C_2 \|\xi_1 - \xi_2\|_s \tag{SB.66}
\]

By the Cauchy-Schwarz inequality, Assumption 3(iii) and (SB.64),

\[
\max_{\{\theta_1, \theta_2\}} \left| A_2(\theta)'((\Omega_{0,d,1}(\theta))^{-1} - (\Omega_{0,d,2}(\theta))^{-1}) A_2(\theta) \right|
\]
\[
\leq \max_{\{\theta_1, \theta_2\}} \|A_2(\theta)\|^2 \|\Omega_{0,d,1}(\theta) - \Omega_{0,d,2}(\theta)\| \leq C_3 \sup_{\theta \in \Theta} \|\Omega_{0,d,1}(\theta) - \Omega_{0,d,2}(\theta)\| \tag{SB.67}
\]

for some constant \(C_3\). Collecting the results in (SB.61), (SB.65), (SB.66) and (SB.67), we get

\[
\left| \min_{\theta \in \Theta} A_1(\theta)'(\Omega_{0,d,1}(\theta))^{-1} A_1(\theta) - \min_{\theta \in \Theta} A_2(\theta)'(\Omega_{0,d,2}(\theta))^{-1} A_2(\theta) \right| \leq (2C_2 + C_3) \|\xi_1 - \xi_2\|_s \tag{SB.68}
\]

which together with (SB.59) and (SB.60) 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

\[
L(v; d) \equiv v'M_d v - \min_{\theta \in \Theta} (m_d(\theta) + V_{d,S}(\theta) v)' (\Omega_{d,0}(\theta))^{-1} (m_d(\theta) + V_{d,S}(\theta) v), \tag{SB.69}
\]

where \(V_{d,S}(\cdot) \equiv V_d(\cdot) \Omega_d^{1/2} M_d\). This functional form is analogous to \(R(\xi)\), with \(\Omega_d\) 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 \(\sup_{\theta \in \Theta} \|V_{d,S}(\theta)\| \leq C_{12}^{1/2} C_V\) (by Assumptions 2(iv) and 3(iii) and Lemma A1(e) of Cheng, Dou, and Liao (2021)), showing \(L(v; d)\) is Lipschitz in \(d\) is analogous to showing \(R(\xi)\) is Lipschitz in \(\xi\). The only difference is that \(M_d\) is not a full rank matrix, unlike \(\Omega_d\), which is the reason that we have to bound \(\|v\|\) directly instead of bounding \(v'M_d v\). Because (SB.60) in the proof of part (i) uses the full rank condition of \(\Omega_d\), we replace (SB.60) with the following argument to show \(v'M_d v\) is Lipschitz in \(M_d\). Given \(\|v\| \leq C\), we have

\[
\|v'M_{d,1} v - v'M_{d,2} v\| = \|v'(M_{d,1} - M_{d,2}) v\| \leq \|v\|^2 \|M_{d,1} - M_{d,2}\| \leq C^2 \|M_{d,1} - M_{d,2}\| \tag{SB.70}
\]

by the Cauchy-Schwarz inequality. The rest of the proof is analogous to those 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,

\[
0 \leq R_C(\xi) \leq (x'_\Omega^{-1}d)x't_C(x'_\Omega^{-1}d)x \leq C
\]

(SB.71)

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

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

\[
|R_C(\xi_1) - R_C(\xi_2)| \leq C_R \|\xi_1 - \xi_2\|_s,
\]

(SB.72)

where \( C_R \) is a constant. Without loss of generality, we assume that \( x'_2\Omega^{-1}_{d,2}x_2 \leq x'_1\Omega^{-1}_{d,1}x_1 \). By the triangle inequality,

\[
|R_C(\xi_1) - R_C(\xi_2)| \leq |(R(\xi_1) - R(\xi_2))t_C(x'_1\Omega^{-1}_{d,1}x_1)| + |t_C(x'_1\Omega^{-1}_{d,1}x_1) - t_C(x'_2\Omega^{-1}_{d,2}x_2))R(\xi_2)|.
\]

(SB.73)

We have (SB.72) holds with \( C_R = C_{R_1} + C_{R_2} \) if we can show that

\[
|(R(\xi_1) - R(\xi_2))t_C(x'_1\Omega^{-1}_{d,1}x_1)| \leq C_{R_1} \|\xi_1 - \xi_2\|_s
\]

(SB.74)

and

\[
|t_C(x'_1\Omega^{-1}_{d,1}x_1) - t_C(x'_2\Omega^{-1}_{d,2}x_2))R(\xi_2)| \leq C_{R_2} \|\xi_1 - \xi_2\|_s
\]

(SB.75)

for some constants \( C_{R_1} \) and \( C_{R_2} \).

We first consider (SB.74). First, note that it holds trivially if \( x'_1\Omega^{-1}_{d,1}x_1 > 2C \) because, in this case, \( t_C(x'_1\Omega^{-1}_{d,1}x_1) = 0 \). Second, given \( x'_2\Omega^{-1}_{d,2}x_2 \leq x'_1\Omega^{-1}_{d,1}x_1 \leq 2C \), we deduce that

\[
|\xi_1 - \xi_2| \leq C_{R_1} \|\xi_1 - \xi_2\|_s,
\]

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.75). First, note that it holds trivially if \( 2C < x'_2\Omega^{-1}_{d,2}x_2 \). In this case,

\[
t_C(x'_1\Omega^{-1}_{d,1}x_1) = t_C(x'_2\Omega^{-1}_{d,2}x_2) = 0
\]

(SB.77)

following the definition of \( t_C(u) \). Second, given \( x'_2\Omega^{-1}_{d,2}x_2 \leq 2C \), we have

\[
|t_C(x'_1\Omega^{-1}_{d,1}x_1) - t_C(x'_2\Omega^{-1}_{d,2}x_2))R(\xi_2)| \leq |R(\xi_2)| \leq x'_2\Omega^{-1}_{d,2}x_2 \leq 2C.
\]

(SB.78)
Thus, (SB.75) holds with \( C_{R_2} = 1 \) if \( \|\xi_1 - \xi_2\|_s > 2C \). Third, it remains to consider the case where \( x_2^d \Omega_{d,2}^{-1} x_2 \leq 2C \) and \( \|\xi_1 - \xi_2\|_s \leq 2C \). In this case, Lemma SC3 in Section SC implies that \( x_1^\prime \Omega_{d,1}^{-1} x_1 \leq C^\ast \) for some constant \( C^\ast \). In this case,

\[
\left| t_c(x_1^\prime \Omega_{d,1}^{-1} x_1) - t_c(x_2^\prime \Omega_{d,2}^{-1} x_2) \right| R(\xi_2) \leq t_c(x_1^\prime \Omega_{d,1}^{-1} x_1) - t_c(x_2^\prime \Omega_{d,2}^{-1} x_2) \leq 2C \left| x_1^\prime \Omega_{d,1}^{-1} x_1 - x_2^\prime \Omega_{d,2}^{-1} x_2 \right| \tag{SB.79}
\]

using \( 0 \leq R(\xi_2) \leq x_2^\prime \Omega_{d,2}^{-1} x_2 \leq 2C \) 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 \( |x_1^\prime \Omega_{d,1}^{-1} x_1 - x_2^\prime \Omega_{d,2}^{-1} x_2| \leq C_{R_2} \|\xi_1 - \xi_2\|_s \) for some constant \( C_{R_2} \) by the same arguments that show (SB.60) but with \( x_1^\prime \Omega_{d,1}^{-1} x_1 \leq 2C \) replaced by \( x_1^\prime \Omega_{d,1}^{-1} x_1 \leq C^\ast \). 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 \( t_c(x^\prime \Omega_{d}^{-1} x) \) replaced by \( t_c(v^\prime M_d v)I\{\|v\|^2 \leq C\} \) and Lemma SB1(i) replaced by Lemma SB1(ii), and hence is omitted. Q.E.D.

**Proof of Lemma A5.** Since \( 0 \leq L(v; d) \leq v^\prime M_d v \) for any \( v \in \mathbb{R}^k \) and \( u t_c(u) \leq C \) for any \( u \geq 0 \), we have

\[
\overline{L}_c(v; d) = L(v; d)t_c(v^\prime M_d v)I\{\|v\|^2 \leq C\} \leq L(v; d)t_c(v^\prime M_d v) \leq C, \tag{SB.80}
\]

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

\[
|\overline{L}_c(v; d_1) - \overline{L}_c(v; d_2)| \leq C_L \|d_1 - d_2\|_s. \tag{SB.81}
\]

Since \( \overline{L}_c(v; d_1) \geq \overline{L}_c(v; d_2) - C_L \|d_1 - d_2\|_s \) for any \( v \) and \( P^\ast (\overline{L}_c(v^\ast; d_1) > c_{\alpha,C}(d_1)) \leq \alpha \), we have \( P\left(\overline{L}_c(v^\ast; d_2) > c_{\alpha,C}(d_1) + C_L \|d_1 - d_2\|_s\right) \leq \alpha \), which implies that

\[
c_{\alpha,C}(d_2) \leq c_{\alpha,C}(d_1) + C_L \|d_1 - d_2\|_s. \tag{SB.82}
\]

Similarly, we also have

\[
c_{\alpha,C}(d_1) \leq c_{\alpha,C}(d_2) + C_L \|d_1 - d_2\|_s. \tag{SB.83}
\]

Combining (SB.82) and (SB.83), we get

\[
|c_{\alpha,C}(d_1) - c_{\alpha,C}(d_2)| \leq C_L \|d_1 - d_2\|_s, \tag{SB.84}
\]

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 (2021) and some other auxiliary results.

**Proof of Lemma A3.** For any square matrices $A_1$ and $A_2$, let $\text{diag}(A_1, A_2)$ 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 $g(\theta)'(\hat{\Omega}(\theta))^{-1}g(\theta)$,

$$R(\hat{x}) \geq g(\hat{\theta})'\hat{\Omega}^{-1}g(\hat{\theta}) - g(\hat{\theta})'(\hat{\Omega}^{-1} - \text{diag}(\hat{\Omega}^{-1}_0, 0_{k_1 \times k_1}))g(\hat{\theta}),$$

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,

$$\hat{\Omega}(\hat{\Omega}^{-1} - \text{diag}(\hat{\Omega}^{-1}_0, 0_{k_1 \times k_1}))\hat{\Omega} = \text{diag}(0_{k_0 \times k_0}, \hat{\Omega}_1 - \hat{\Omega}_{1,0}\hat{\Omega}_0^{-1}\hat{\Omega}_{0,1})$$

is a positive semi-definite matrix, which together with (SC.85) 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 $v \in \mathbb{R}^k$

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

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

$$L(x; \hat{d}) \geq v'\hat{M}\hat{\Omega}^{1/2} \left(\hat{\Omega}^{-1} - \text{diag}(\hat{\Omega}^{-1}_0, 0_{k_1 \times k_1})\right)\hat{\Omega}^{1/2}\hat{M}v \geq 0,$$

where the second inequality holds since the matrix in (SC.86) is positive semi-definite. Q.E.D.

**Lemma SC3.** For any $\xi_1$ and $\xi_2$ 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 on the constant $C$ of the lemma, $C_3$, and $c_\lambda$ 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),

$$x_1'\Omega_{d,1}^{-1}x_1 \leq 2x_2'\Omega_{d,2}^{-1}x_2 + 2(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2)$$

$$= 2x_2'\Omega_{d,2}^{-1}x_2 + 2x_2'(\Omega_{d,1}^{-1} - \Omega_{d,2}^{-1})x_2 + 2(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2)$$

$$\leq 2C + 2x_2'(\Omega_{d,1}^{-1} - \Omega_{d,2}^{-1})x_2 + 2(x_1 - x_2)'\Omega_{d,1}^{-1}(x_1 - x_2),$$

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 \( \|\xi_1 - \xi_2\|_s \leq C \),

\[
(x_1 - x_2)' \Omega_{d,1}^{-1} (x_1 - x_2) \leq (\lambda_{\min}(\Omega_{d,1}))^{-1} \|x_1 - x_2\|^2 \leq C_2^2 \lambda^{-1}. \tag{SC.90}
\]

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

\[
\left| x'_2 (\Omega_{d,1}^{-1} - \Omega_{d,2}^{-1}) x_2 \right|^2 = \left| x'_2 \Omega_{d,1}^{-1} (\Omega_{d,1} - \Omega_{d,2}) \Omega_{d,2}^{-1} x_2 \right|^2 \leq (x'_2 \Omega_{d,1}^{-1} x_2) (x'_2 \Omega_{d,2}^{-1} x_2) \|\Omega_{d,1} - \Omega_{d,2}\|^2 \\
\leq \frac{\lambda_{\max}(\Omega_{d,2}) (x'_2 \Omega_{d,2}^{-1} x_2)^2}{\lambda_{\min}(\Omega_{d,2})} \|\Omega_{d,1} - \Omega_{d,2}\|^2 \leq C_2 \lambda^{-3} C_\Omega, \tag{SC.91}
\]

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.89)-(SC.91). \( Q.E.D. \)

**Lemma SC4.** For any \( \xi_1 \) and \( \xi_2 \), 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

\[
\sup_{\theta \in \Theta} \|V_{d,S,1}(\theta) - V_{d,S,2}(\theta)\| \leq C_\Omega^{1/2} (1 + \lambda^{-1} C_\Omega^{1/2} + \lambda^{-1} C_\Omega) \|\xi_1 - \xi_2\|_s, \tag{SC.92}
\]

where \( C_\Omega \) and \( \lambda \) are in Assumptions 2(iv) and 3(iii) respectively.

**Proof of Lemma SC4.** By definition,

\[
V_{d,S,1}(\theta) - V_{d,S,2}(\theta) = [V_{d,1}(\theta) - V_{d,2}(\theta)] \Omega_{d,1}^{1/2} M_{d,1} \\
+ V_{d,2}(\theta)(\Omega_{d,1}^{1/2} - \Omega_{d,2}^{1/2}) M_{d,1} + V_{d,2}(\theta) \Omega_{d,2}^{1/2} (M_{d,1} - M_{d,2}). \tag{SC.93}
\]

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

\[
\sup_{\theta \in \Theta} \left\| [V_{d,1}(\theta) - V_{d,2}(\theta)] \Omega_{d,1}^{1/2} M_{d,1} \right\| \leq C_\Omega^{1/2} \sup_{\theta \in \Theta} \|V_{d,1}(\theta) - V_{d,2}(\theta)\|. \tag{SC.94}
\]

By the properties of \( V_{d,j}(\theta) \) and \( M_{d,j} \), and Exercise 7.2.18 in Horn and Johnson (1990),

\[
\sup_{\theta \in \Theta} \left\| V_{d,2}(\theta)(\Omega_{d,1}^{1/2} - \Omega_{d,2}^{1/2}) M_{d,1} \right\| \leq \sup_{\theta \in \Theta} \|V_{d,2}(\theta)\| \|\Omega_{d,1}^{1/2} - \Omega_{d,2}^{1/2}\|_S \\
\leq \lambda^{-1} C_\Omega \left\|\Omega_{d,1}^{-1}\right\|_S \|\Omega_{d,1} - \Omega_{d,2}\| \leq \lambda^{-2} C_\Omega \|\Omega_{d,1} - \Omega_{d,2}\|. \tag{SC.95}
\]

Similarly,

\[
\sup_{\theta \in \Theta} \left\| V_{d,2}(\theta) \Omega_{d,2}^{1/2} (M_{d,1} - M_{d,2}) \right\| \leq \lambda^{-1} C_\Omega^{3/2} \|M_{d,1} - M_{d,2}\|. \tag{SC.96}
\]

The desirable result follows by (SC.93)-(SC.96) 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 $P \in P_{1,\infty} \subset P$: (i) $\inf_{\theta \in \Theta} \|E[\hat{g}_1(\theta)]\| > c_{g_1}$ for some $c_{g_1} > 0$; (ii) $\lambda_{\min}(\hat{\Omega}_0(\theta_0)) \geq c_{\lambda}$, $\lambda_{\min}(\hat{\Omega}) \geq c_{\lambda}$ and $\lambda_{\min}(\hat{Q}'\hat{Q}) \geq c_{\lambda}$ wpa1.

Assumption SD1(i) implies that there are globally misspecified moments in $E[\hat{g}_1(\theta_0)] = 0_{k_1 \times 1}$. 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 $\theta_1$ 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 $P \in P_{1,\infty}$, $P(T > c_\alpha(\hat{d})) \to 1$ as $n \to \infty$.

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

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

where

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

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

$$R(\hat{\xi}) \geq n \left( K^{-1} \min_{\theta \in \Theta} \|E[\hat{g}(\theta)]\|^2 - o_p(1) \right) \geq nc_{g_1}K^{-1} \text{ wpa1}. \quad (SD.3)$$

The critical value satisfies $c_\alpha(\hat{d}) \leq q_{1-\alpha}(\chi^2_k)$ wpa1, because $L(v^*; \hat{d}) \leq v^*\hat{M}v^* \leq \|v^*\|^2$ wpa1 given that $\hat{M}$ is an idempotent matrix wpa1 under Assumption SD1(ii) and $q_{1-\alpha}(\chi^2_k)$ 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

\[
P\left(R(\hat{\xi}) > c_\alpha(\hat{d})\right) \geq P\left(nc_{q_1}K^{-1} > q_{1-\alpha}(\chi_k^2)\right) - o(1),
\]

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 the parameter \(\theta_0\) (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 \(P \in \mathcal{P}_{1,A} \subset \mathcal{P}:\)

(i) \(E[\hat{g}_1(\theta_0)] = an^{-1/2}\) for some \(a \in \mathbb{R}^{k_1}\) with \(\|a\| < \infty\);

(ii) Assumptions 3(ii) and 3(iii) hold for any \(P \in \mathcal{P}_{1,A}.

**Theorem SD2.** Suppose Assumptions 1, 2, 4 and SD2 hold. For any \(P \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}\), we have

\[
P\left(T > c_\alpha(\hat{d})\right) \to P\left(\chi^2_{k_1}(a_\Omega M a_\Omega) > q_{1-\alpha}(\chi^2_{k_1})\right), \quad \text{as } n \to \infty,
\]

where \(a_\Omega \equiv \Omega^{-1/2}a\) and \(\chi^2_{k_1}(a_\Omega M a_\Omega)\) 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 \(\theta_0\) are consistent by the standard arguments and results in (A.19) and (A.20) of Cheng, Dou, and Liao (2021) remain valid. Therefore,

\[
R(\hat{\xi}) \to_d (\Omega^{-1/2}v + a_\Omega)'M(\Omega^{-1/2}v + a_\Omega) - v_0'\Omega_0^{-1/2}M_0\Omega_0^{-1/2}v_0,
\]

where \(a_\Omega \equiv \Omega^{-1/2}a\), \(v\) is a multivariate normal random variable with mean zero and variance \(\Omega\), 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

\[
(\Omega^{-1/2}v + a_\Omega)'M(\Omega^{-1/2}v + a_\Omega) - v_0'\Omega_0^{-1/2}M_0\Omega_0^{-1/2}v_0 \sim \chi^2_{k_1}(a_\Omega M a_\Omega).
\]

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 (2021) remains valid for any \(P \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}\). Therefore, for any \(P \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A},

\[
L(v^*, \hat{d}) = v^*(M - \tilde{M}_0)v^* + o_p(1) \sim \chi^2_{k_1}.
\]
By (SD.7) and arguments analogous to those used to show Theorem 2(ii), we have $c_\alpha(d) \to_p q_{1-\alpha}(\chi^2_{k_1})$, 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 $E[g_0(\theta_0)] = 0_{k_0 \times 1}$ only depend on a subvector $\theta_{c,0}$ of $\theta_0$ with dimension $d_c$ and strongly identify $\theta_{c,0}$, arguments analogous to those used to show Theorem SD2 give

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

When the baseline moments provide strong identification, the conditional specification test is asymptotically equivalent to the $C$ test following Theorems 2, SD1 and SD2.\textsuperscript{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 (2021). 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

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

where $g_{\infty}(\theta_0) \equiv (g_{0,\infty}(\theta_0)', g_{1,\infty}(\theta_0)')'$ is normal with covariance matrix $\Omega$, $g_{0,\infty}(\theta_0)$ 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 $\Omega_0(\cdot, \cdot)$, $\Omega$, $Q$ and the covariance between $g_{0,\infty}(\cdot)$ and $g_{\infty}(\theta_0)$ are known.

We are interested in testing

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

\textsuperscript{1}See e.g., Hall (2005) for detailed derivations for the $C$ test.
while maintaining \( E[g_{0,\infty}(\theta_0)] = 0_{k_0 \times 1} \) 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 \( \theta_0 \) 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 \( Q^\perp \) 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,
\[
Dg_{\infty}(\hat{\theta}) = D(I_k - Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1})g_{\infty}(\theta_0) = \begin{pmatrix} Q^\perp g_{\infty}(\theta_0) \\ 0_{d_\theta \times k} \end{pmatrix}. \tag{SE.4}
\]

Based on (SE.4), observing \( g_{\infty}(\hat{\theta}) \) is equivalent to observing
\[
Y \equiv Q^\perp g_{\infty}(\theta_0) \sim N\left( A(\eta), Q^\perp\Omega Q^\perp \right), \text{ where } A(\eta) \equiv Q^\perp \begin{pmatrix} 0_{k_0 \times 1} \\ \eta \end{pmatrix}. \tag{SE.5}
\]

We next consider inference of \( \eta \) based only on \( Y \).

Since the uniformly most powerful (UMP) test does not exists 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 \( \eta^* \) with \( A(\eta^*) \neq 0 \), then the Neyman-Pearson lemma implies that the UMP test rejects \( H_0 \) if
\[
\frac{A(\eta^*)' (Q^\perp \Omega Q^\perp)^{-1} Y}{(A(\eta^*)' (Q^\perp \Omega Q^\perp)^{-1}A(\eta^*))^{1/2}} > z_{1-\alpha}, \tag{SE.6}
\]
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 \( \eta^* \) by construction. Its power may be low if the true value \( \eta \) under the alternative is different from \( \eta^* \). In the simulation study below, we let the test in (SE.6) depend on the true value under the alternative \( \eta \) (i.e., we replace \( \eta^* \) by \( \eta \)) and call its power (as a function of \( \eta \)) as PE-1. Next, we consider a subset of alternative hypothesis (SE.3) which are proportional to a known vector \( \eta^* \) with \( A(\eta^*) \neq 0 \), i.e., \( H_1 : \eta = a\eta^* \). Since \( \eta^* \) 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
\[
\left| \frac{A(\eta^*)' (Q^\perp \Omega Q^\perp)^{-1} Y}{(A(\eta^*)' (Q^\perp \Omega Q^\perp)^{-1}A(\eta^*))^{1/2}} \right| > z_{1-\alpha/2}. \tag{SE.7}
\]

In the simulation study below, we let the test in (SE.7) also depend on the true value \( \eta \) under the alternative and call its power (as a function of \( \eta \)) as PE-2. Both PE-1 and PE-2 are infeasible
because they require the knowledge of the true value \( \eta \) under the alternative. Finally, the following feasible test

\[ Y'(Q_{\perp}^{T}\Omega Q_{\perp})^{-1}Y > q_{1-\alpha}(\chi_{k-d_{q}}^{2}) \]  

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 \( T \) in the paper is the QLR statistic written as

\[ T \equiv g_{\infty}(\hat{\theta})'\Omega^{-1}g_{\infty}(\hat{\theta}) - \min_{\theta \in \Theta} g_{0,\infty}(\theta)'(\Omega_{0}(\theta))^{-1}g_{0,\infty}(\theta), \]  

where \( \Omega_{0}(\theta) \equiv \Omega_{0}(\theta, \theta) \). We apply the conditional inference based on this test statistic. Define

\[ m_{0,\infty}(\theta) \equiv g_{0,\infty}(\theta) - V(\theta)g_{\infty}(\hat{\theta}), \]  

where \( V(\theta) \equiv \text{Cov}(g_{0,\infty}(\theta), g_{\infty}(\hat{\theta})) \Omega^{-1} \) is a known function of \( \theta \). Then under the null hypothesis, \( \text{Cov}(m_{0,\infty}(\theta), g_{\infty}(\hat{\theta})) = 0 \) 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

\[ T = g_{\infty}(\hat{\theta})'\Omega^{-1}g_{\infty}(\hat{\theta}) - \min_{\theta \in \Theta} (m_{0,\infty}(\theta) + V(\theta)g_{\infty}(\hat{\theta}))'(\Omega_{0}(\theta))^{-1}(m_{0,\infty}(\theta) + V(\theta)g_{\infty}(\hat{\theta})) \]  

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 \( d_{q} = 1, k_{0} = qk_{1}, k_{1} = 2 \), and

\[ g_{0,\infty}(\theta) \equiv g_{0,\infty}(\theta_{0}) + (\theta - \theta_{0})Q_{0,\infty} \]  

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:

\[
\begin{pmatrix}
g_{0,\infty}(\theta_{0}) \\
g_{1,\infty}(\theta_{0}) \\
Q_{0,\infty}
\end{pmatrix}
\sim \mathcal{N}
\begin{pmatrix}
0_{k_{0} \times 1} \\
\eta \\
\mu_{g}
\end{pmatrix},
\Sigma
\]

where \( \Sigma \equiv \begin{pmatrix}
\Omega_{00} & \Omega_{01} & \Omega_{0g} \\
\Omega_{10} & \Omega_{11} & \Omega_{1g} \\
\Omega_{g0} & \Omega_{g1} & \Omega_{gg}
\end{pmatrix} \)  

where \( \mu_{g} \) 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 nonrandom vector as in the simple disaster risk model in Section 2 of Cheng, Dou, and Liao (2021). In this case, \( \Omega_{gg}, \Omega_{g1}, \Omega_{1g}, \Omega_{0g} \) and \( \Omega_{g0} \) 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:

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

\[ V(\theta) = \begin{pmatrix} \Omega_{00} & \Omega_{01} \\ \Omega_{01} & \Omega_{11} \end{pmatrix}^{-1} + (\theta - \theta_0) \begin{pmatrix} \Omega_{g0} & \Omega_{g1} \\ \Omega_{g1} & \Omega_{11} \end{pmatrix} \Omega_{gg}, \]

where \(\Omega = \begin{pmatrix} \Omega_{00} & \Omega_{01} \\ \Omega_{10} & \Omega_{11} \end{pmatrix} \).

(SE.14)

We generate the covariance matrix \(\Omega\) as follows

\[ \Omega = \begin{pmatrix} (1 + \rho^2)^{-1}(I_{k_0} + \rho^2 I_{q \times q} \otimes \Omega_u) & \rho I_{q \times q} \otimes \Omega_u \\ \rho I_{q \times q} \otimes \Omega_u & \Omega_u \end{pmatrix} \]

where \(\Omega_u = \begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix} \).

(SE.15)

where \(A \otimes B\) denotes the Kronecker product of two real matrices \(A\) and \(B\). The parameter \(\lambda\) determines the correlation between the two moments in \(g_{1,\infty}(\theta_0)\), while \(\rho\) 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

\[ \Omega_{gg} = I_{k_0}, \quad \Omega_{0g} = \Omega'_{g0} = \lambda I_{k_0} \quad \text{and} \quad \Omega_{1g} = \Omega'_{g1} = \lambda I_{q \times q} \otimes I_{k_1} \]

(SE.16)

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

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

Benchmark case: \(\rho = 0.4\), \(c_g = 0\), \(c_\mu = 1\), nonrandom \(Q_{0,\infty}\), \(q = k_0/k_1 = 1\) 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 non-trivial 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 3 deviation cases are presented in Figure 1 and Figures 2-4, 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 $\eta$, 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 1 shows that increasing the number of baseline moments significantly enlarges the power gain compared to PE-3 while roughly maintains the same amount of power loss compared to PE-2. Figure 2 and Figure 3 show that increasing the correlation between the baseline moments and the asset-pricing moments, 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 4 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 1: Power Comparison in the Benchmark Case

![Figure 1](image1.png)

Note: In the Benchmark case, we have $\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, nonrandom $Q_{0,\infty}$, $q = k_0/k_1 = 1$ or 2.

Figure 2: Power Comparison in the Deviation Case 1

![Figure 2](image2.png)

Note: In the deviation case 1, we have $\rho = 0.2$ or 0.8, $c_g = 0$, $c_\mu = 1$, non-random $Q_{0,\infty}$, $q = 2.$
We consider the following calibrated values for the nuisance parameters

\[(\delta, g_c, g_d, \sigma_c, \phi, \nu, q) = (0.97, 0.02, 0.02, 0.02, 3.5, 0.07, 0.4).\] (SF.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 \( \Theta \) for the unknown parameter is set to \( \Theta = \Theta_1 \times \Theta_2 \times \Theta_3 \times \Theta_4 \) where

\[
\Theta_1 = [0.001, 0.02], \; \Theta_2 = [5, 12], \; \Theta_3 = [0.95, 0.999], \; \text{and} \; \Theta_4 = [3, 6]. \quad \text{(SF.3)}
\]

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 \( \Theta_1 \) and \( \Theta_3 \), and step size 0.01 in \( \Theta_2 \) and \( \Theta_4 \).\(^2\) 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 \( \Theta_2 = [5, 8] \) and a larger step size 0.1 of the grid points in \( \Theta_2 \) and \( \Theta_4 \) to reduce the computational cost. The parameter spaces \( \Theta_j (j = 1, 3, 4) \) and the grid points in \( \Theta_1 \) and \( \Theta_3 \) are unchanged. The reduced space \( \Theta_2 \) still covers the CUE estimators of \( \theta_2 \) for the three values of \( p \) considered. The model uncertainty sets of \((\eta_1, \eta_3)\) and \((\eta_3, \eta_4)\) are calculated through grid search with equally spaced grid points for \( \eta_j \) \((j = 1, 3, 4)\) with step size 0.001. The parameter spaces for \( \eta_j \) \((j = 1, 3, 4)\) are set large enough such that the model uncertainty sets from the \( J \) test are contained in the interior of these parameter spaces.

References


\(^2\)We have also considered a much larger parameter space with \( \Theta_1 = [0.001, 1], \; \Theta_2 = [1.15], \; \Theta_3 = [0.9, 0.999], \; \Theta_4 = [1, 20] \), 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 1 of the paper.