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

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

February 16, 2021

Abstract

This supplemental appendix provides the following supporting materials. Section SA contains proofs of the theoretical results in Section 4 of the paper on the size of the new conditional specification test. Section SB provides additional theoretical results on the power of the new test. Section SC provides comparison with some power envelopes through simulations. Section SD collects extra details of the empirical application.

*Department of Economics, University of Pennsylvania; Email: xucheng@econ.upenn.edu.

†Finance Department, Wharton School, University of Pennsylvania; Email: wdou@wharton.upenn.edu.

‡Department of Economics, University of California, Los Angeles; Email: zhipeng.liao@econ.ecla.edu.

SA Proofs of Theoretical Results on Size of the New Test

We first discuss the assumptions imposed in the paper to establish results on the asymptotic size of the new specification test.

Assumption 1(i) requires that the rescaled moment condition is well approximated by a Gaussian limit. Assumption 1(ii) follows from the uniform law of large numbers. Assumption 1(iii) includes standard regularity conditions on uniformly bounded moment functions and their derivatives.

In our long-run risk example, Gaussian approximation is innocuous even if the root of the latent autoregressive process could be arbitrarily close to unity, different from the classical near unit root analysis (e.g., Phillips, 1987; Mikusheva, 2007). See Section E in Cheng, Dou, and Liao (2020) for the details.

Assumptions 2(i) and 2(iii) require that we have uniformly consistent estimators of the covariance function $\Omega(\cdot,\cdot)$ and its partial derivatives. Uniform consistency can be obtained by strengthening a pointwise consistent covariance matrix estimator with smoothness conditions. Assumptions 2(ii) and 2(iv) impose continuity and uniform upper bounds on the covariance function $\Omega(\cdot,\cdot)$ and its partial derivatives. Both Assumptions 1 and 2 are imposed on $\mathcal{P}$, not only on $\mathcal{P}_0$, because they are useful for both the size and power analysis of the proposed conditional specification test.

Assumption 3 is used to show consistency and asymptotic normality of $\hat{\theta}$ under the null hypothesis. Assumptions 3(i) and 3(ii) provide the identification uniqueness condition of the unknown parameter θ_0 using all valid moments under the null hypothesis. Assumption 3(iii) includes standard full rank conditions for the Jacobian matrix and the covariance matrix when all moments are used.

Assumption 4 is similar to Assumption 3 which is imposed on $\mathbb{E}\left[\bar{g}(\theta)\right]$ for the strong identification of θ_0 using all moments. This assumption is needed to show that the test statistic $\mathcal{T}$ converges to a chi-square distribution and the critical value $c_\alpha(\hat{d})$ converges in probability to the $1-\alpha$ quantile of this chi-square distribution under strong identification.

Throughout the proofs, we use K to denote a positive constant that may change from line to line. For any $x \in \mathbb{R}^{k_0}$ and any $k_0 \times k_0$ symmetric positive definite matrix A, $\|x\|_A \equiv (x'A^{-1}x)^{1/2}$. Let $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the smallest and the largest eigenvalues of a real symmetric matrix A, respectively. Proof of all auxiliary Lemmas, i.e., Lemmas SA1 – SA7 below, are given in Cheng, Dou, and Liao (2020).

Proof of Lemma 1. Since Mv^* and $m^*(\cdot)$ are mean-zero Gaussian, part (i) follows from $E\left[m^*(\cdot)(Mv^*)'\right] = 0$. For part (ii), by the law of iterated expectation and the definition of $c^*_\alpha(d^*)$,

$$P\left(L(v^*, d^*) > c^*_\alpha(d^*)\right) = E\left[P\left(L(v^*, d^*) > c^*_\alpha(d^*)|d^*\right)\right] \le \alpha. \tag{SA.1}$$

Part (iii) follows from $P\left(L(\upsilon^{*},d^{*})>c_{\alpha}^{*}(d^{*})|d^{*}\right)=\alpha$ under the specified continuity condition. *Q.E.D.*

The following results hold for the CUE $\hat{\theta}$ in (3.2) that uses the full moments and some estimators based on $\hat{\theta}$, regardless of the identification strength in the baseline moments.

Lemma SA1. *Under Assumptions 1, 2 and 3, the following results hold uniformly over $\mathbb{P}\in\mathcal{P}_0$:*
(a) $n^{1/2}(\hat{\theta}-\theta_0)=\left(Q'\Omega^{-1}Q\right)^{-1}Q'\Omega^{-1}g(\theta_0)+o_p(1)=O_p(1)$;
(b) $g(\hat{\theta})=\Omega^{1/2}M\Omega^{-1/2}g(\theta_0)+o_p(1)=O_p(1)$;
(c) $\hat{\Omega}=\Omega+o_p(1)$ *where* $\hat{\Omega}\equiv\hat{\Omega}(\hat{\theta})$;
(d) $\hat{M}=M+o_p(1)$;
(e) $\sup_{\theta\in\Theta}||\hat{V}(\theta)-V(\theta)||=o_p(1)$, *where* $\sup_{\theta\in\Theta}\|V(\theta)\|\leq c_{\lambda}^{-1}C_{\Omega}$.

We next present a few lemmas used in the proof of Theorem 1. For any $x\in\mathbb{R}^{k}$, continuous vector function $m_d:\Theta\mapsto\mathbb{R}^{k_0}$, continuous matrix function $V_d:\Theta\mapsto\mathbb{R}^{k_0\times k}$, $k\times k$ symmetric positive definite matrix Ω_d, symmetric and continuous matrix function $\Omega_{0,d}(\cdot):\Theta\mapsto\mathbb{R}^{k_0\times k_0}$ which is positive definite for any $\theta\in\Theta$, and $k\times k$ symmetric idempotent matrix M_d, let

$$\xi\equiv(x',d')',\ \text{where } d\equiv(m_d(\cdot)',\text{vec}(V_d(\cdot))',\text{vech}(\Omega_d)',\text{vech}(\Omega_{0,d}(\cdot))',\text{vech}(M_d)')'.$$

Define

$$R(\xi)\equiv\|x\|_{\Omega_d}^2-\min_{\theta\in\Theta}\|m_d(\theta)+V_d(\theta)x\|_{\Omega_{0,d}(\theta)}^2, \tag{SA.2}$$

and

$$L(\upsilon;d)\equiv\upsilon'M_d\upsilon-\min_{\theta\in\Theta}\left\|m_d(\theta)+V_d(\theta)\Omega_d^{1/2}M_d\upsilon\right\|_{\Omega_{0,d}(\theta)}^2. \tag{SA.3}$$

The test statistic $\mathcal{T}$ in (3.1) can be written as

$$\begin{aligned}\mathcal{T}=R(\hat{\xi}),\ \text{where }&\hat{\xi}\equiv(g(\hat{\theta})',\hat{d}')'\text{ and}\\ &\hat{d}\equiv(\hat{m}(\cdot)',\text{vec}(\hat{V}(\cdot))',\text{vech}(\hat{\Omega})',\text{vech}(\hat{\Omega}_0(\cdot))',\text{vech}(\hat{M})')'.\end{aligned} \tag{SA.4}$$

Given $\hat{d}$, the critical value $c_{\alpha}(\hat{d})$ is simulated using $L(\upsilon^{*};\hat{d})$ with independent draws of $\upsilon^{*}\sim N(0,I_k)$. To show the bounded and Lipschitz properties of functionals of ξ, we use the metric

$$\|\xi\|_s=\|x\|+\sup_{\theta\in\Theta}\|m_d(\theta)\|+\sup_{\theta\in\Theta}\|V_d(\theta)\|+\|\Omega_d\|+\sup_{\theta\in\Theta}\|\Omega_{0,d}(\theta)\|+\|M_d\|. \tag{SA.5}$$

Lemma SA2. *Under Assumptions 1, 2 and 3,*

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

where $\xi^{*}\equiv((\Omega^{1/2}M\upsilon^{*})',d^{*\prime})'$ and BL_1 denotes the set of functionals with Lipschitz constant and *supremum norm bounded above by 1.*

To use the weak convergence of $\hat{\xi}$ for studying the statistic $\mathcal{T}$, we follow Andrews and Mikusheva (2016) and define a truncated version of $R(\xi)$ as

$$R_C(\xi) \equiv R(\xi) t_C(x'\Omega^{-1}x) \tag{SA.6}$$

where $t_C(u) \equiv I\{u < C\} + (2C - u)C^{-1} I\{C \leq u < 2C\}$ for any $u \in \mathbb{R}$ and some $C \geq 1$. Similarly, to study the critical value $c_\alpha(\hat{d})$, we define a truncated version of $L(\upsilon; d)$ as

$$\overline{L}_C(\upsilon; d) \equiv L_C(\upsilon; d) I\{\|\upsilon\|^2 \leq C\}, \text{ where } L_C(\upsilon; d) \equiv L(\upsilon; d) t_C(\upsilon' M_d \upsilon). \tag{SA.7}$$

Compared with $R_C(\xi)$, the truncation in $\overline{L}_C(\upsilon; d)$ has an extra term $I\{\|\upsilon\|^2 \leq C\}$, which is needed to show that $\overline{L}_C(\upsilon; d)$ is Lipschitz in M_d. Since M_d may not have full rank, the truncation with $t_C(\upsilon' M_d \upsilon)$ is insufficient to bound $\|\upsilon\|$. Thus, truncation with $\|\upsilon\|^2 \leq C$ is added in $\overline{L}_C(\upsilon; d)$.

Lemma SA3. *Suppose that $\hat{\Omega}$ is symmetric and positive definite and $\hat{\Omega}_0$ is the leading $k_0 \times k_0$ submatrix of $\hat{\Omega}$. Then $\mathcal{T} \geq 0$. Moreover, if $\hat{Q}'\hat{Q}$ is nonsingular, $L(\upsilon; \hat{d}) \geq 0$ for any $x \in \mathbb{R}^k$.*

Lemma SA4. *Given $R(\xi) \geq 0$, the functional $R_C(\xi)$ is bounded and Lipschitz in ξ.*

Lemma SA5. *Let $c_{\alpha,C}(d) \equiv \inf\left\{c : P^*\left(\upsilon^* : \overline{L}_C(\upsilon^*; d) > c\right) \leq \alpha\right\}$. Given $L(\upsilon; d) \geq 0$, $c_{\alpha,C}(d)$ is bounded and Lipschitz in d.*

The extra truncation $\|\upsilon\|^2 \leq C$ in $\overline{L}_C(\upsilon; d)$ causes a discrepancy between $c_{\alpha,C}(d^*)$ and the conditional $1-\alpha$ quantile of $R_C(\xi^*)$ given d^*. Lemma SA6 below shows that we can choose C large enough such that the discrepancy is negligible, which is one of the key elements to show the uniform size control of the conditional specification test.

Lemma SA6. *For any $\varepsilon \in (0, 1)$ and any $\delta > 0$, there is a finite constant C_δ such that for any $C \geq C_\delta$: $P\left(R_C(\xi^*) > c_{\alpha,C}(d^*) + \varepsilon\right) \leq \alpha + \delta/4$.*

Proof of Theorem 1. The proof strategy follows from that for Theorem 1 of Andrews and Mikusheva (2016). The major differences are as follows. (i) The test statistic and the critical value are defined with different functions, $R(\hat{\xi})$ and $L(\upsilon^*; \hat{d})$, respectively. These two functions have to be truncated differently too, as in (SA.6) and (SA.7), respectively, to yield the bounded Lipschitz property. (ii) The additional truncation to $L(\upsilon^*; \hat{d})$ causes a discrepancy between $c_{\alpha,C}(d^*)$ and the conditional $1-\alpha$ quantile of $R_C(\xi^*)$ given d^*. Lemma SA6 is used to address these problems.

For notational simplicity, we assume that $\inf_{\theta\in\Theta} \lambda_{\min}(\hat{\Omega}(\theta)) \geq K^{-1}$, $\lambda_{\min}(\hat{Q}'\hat{Q}) \geq K^{-1}$ and $\sup_{\theta\in\Theta} \lambda_{\max}(\hat{\Omega}(\theta)) \leq K$ in the proof. This assumption is innocuous since the above properties hold with probability approaching 1 (wpa1) in view of Assumptions 1(ii), 2(i, iv) and 3(iii), and the consistency of $\hat{\theta}$ under the null. Suppose that the claim of the theorem does not hold. Then

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in\mathcal{P}_0} \mathbb{P}\left(R(\hat{\xi}) > c_\alpha(\hat{d}) + \varepsilon\right) > \alpha, \tag{SA.8}$$

which implies that there exists $\delta > 0$ and a divergent sequence n_i (indexed by i) such that

$$\mathbb{P}_{n_i}\left(R(\hat{\xi}) > c_\alpha(\hat{d}) + \varepsilon\right) > \alpha + \delta \text{ for all } i. \tag{SA.9}$$

For any $u \in \mathbb{R}$ and any i, by the union bound of probability,

$$\mathbb{P}_{n_i}\left(R(\hat{\xi}) > u\right) \leq \mathbb{P}_{n_i}\left(R(\hat{\xi}) > u,\, g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) \leq C\right) + \mathbb{P}_{n_i}\left(g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) > C\right). \tag{SA.10}$$

By the definition of $\hat{\theta}$, $g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) \leq g(\theta_0)'(\hat{\Omega}(\theta_0))^{-1}g(\theta_0)$ which together with Assumptions 1(i), 2(i) and 3 implies that $g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) = O_p(1)$ uniformly over $\mathbb{P} \in \mathcal{P}_0$. Therefore, there exists a large constant $C_{1,\delta}$ such that for all large n_i,

$$\mathbb{P}_{n_i}\left(g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) > C_{1,\delta}\right) \leq \delta/4, \tag{SA.11}$$

which together with (SA.9) and (SA.10) implies that

$$\mathbb{P}_{n_i}\left(R(\hat{\xi}) > c_\alpha(\hat{d}) + \varepsilon,\, g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) \leq C\right) > \alpha + 3\delta/4, \tag{SA.12}$$

for any $C \geq C_{1,\delta}$. By definition,

$$I\left\{R_C(\hat{\xi}) > u\right\} \geq I\left\{R(\hat{\xi}) > u\right\} I\left\{g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}) \leq C\right\} \text{ for any } u \in \mathbb{R}, \tag{SA.13}$$

where $R_C(\hat{\xi}) \equiv R(\hat{\xi})t_C(g(\hat{\theta})'(\hat{\Omega}^{-1})g(\hat{\theta}))$ and $t_C(u) = 1$ for $u \leq C$ following its definition. By (SA.12) and (SA.13), we have for any $C \geq C_{1,\delta}$,

$$\mathbb{P}_{n_i}\left(R_C(\hat{\xi}) > c_\alpha(\hat{d}) + \varepsilon\right) > \alpha + 3\delta/4. \tag{SA.14}$$

Since $L(v, \hat{d}) \geq 0$ for any $v \in \mathbb{R}^k$ by Lemma SA3 and $t_C(u) \leq 1$ for any $u \in \mathbb{R}$, we have $L_C(v, \hat{d}) \leq L(v, \hat{d})$ for any $v \in \mathbb{R}^k$, which further implies that $c_{\alpha,C}(\hat{d}) \leq c_\alpha(\hat{d})$. Therefore, by (SA.14) we deduce that for any $C \geq C_{1,\delta}$,

$$\mathbb{P}_{n_i}\left(R_C(\hat{\xi}) - c_{\alpha,C}(\hat{d}) \geq \varepsilon\right) > \alpha + 3\delta/4. \tag{SA.15}$$

Let $U_{C,n}$ be a random variable which has the same distribution as $R_C(\hat{\xi}) - c_{\alpha,C}(\hat{d})$ under the law $\mathbb{P}_n$. Let $U_{\infty,C,n}$ be a random variable which has the same distribution as $R_C(\xi^*) - c_{\alpha,C}(d^*)$. By Lemma SA4 and Lemma SA5, $R_C(\xi) - c_{\alpha,C}(d)$ is bounded and Lipschitz in ξ. Therefore, by Lemma SA2,

$$\lim_{n\to\infty} \sup_{f\in BL_1} \|\mathbb{E}\left[f(U_{C,n})\right] - E\left[f(U_{\infty,C,n})\right]\| = 0. \tag{SA.16}$$

Since $U_{\infty,C,n}$ is bounded for any n, by Prokhorov's theorem, there exists a subsequence n_j (of n_i) and a random variable U_C such that $U_{\infty,C,n_j} \to_d U_C$, which together with (SA.16) implies that $U_{C,n_j} \to_d U_C$. Since (SA.15) can be written as $\mathbb{P}_{n_i}\left(U_{C,n_i} \geq \varepsilon\right) > \alpha + 3\delta/4$, by Portmanteau theorem,

$$\begin{aligned}\liminf_{n_j\to\infty} P\left(U_{\infty,C,n_j} > \varepsilon/2\right) &\geq P\left(U_C > \varepsilon/2\right) \geq P\left(U_C \geq \varepsilon\right) \\ &\geq \limsup_{n_j\to\infty} \mathbb{P}_{n_j}\left(U_{C,n_j} \geq \varepsilon\right) \geq \alpha + 3\delta/4, \quad \text{for any } C \geq C_{1,\delta}.\end{aligned} \tag{SA.17}$$

We next show that for all large C, $P\left(U_{\infty,C,n_j} > \varepsilon/2\right) \leq \alpha + \delta/4$ for any n_j, which contradicts (SA.17), and hence the claim of the theorem holds. To this end, for $C \geq C_{2,\delta}$ in Lemma SA6,

$$P\left(U_{\infty,C,n_j} > \varepsilon/2\right) = P\left(R_C(\xi^*) > c_{\alpha,C}(d^*) + \varepsilon/2\right) \leq \alpha + \delta/4, \tag{SA.18}$$

where the equality holds because U_{∞,C,n_j} and $R_C(\xi^*) - c_{\alpha,C}(d^*)$ have the same distribution and the inequality follows from Lemma SA6. *Q.E.D.*

Let $\hat{\theta}^* \equiv \arg\min_{\theta\in\Theta} \|\hat{m}(\theta) + \hat{V}(\theta)\hat{\Omega}^{1/2}\hat{M}\upsilon^*\|^2_{\hat{\Omega}_0(\theta)}$ and $\tilde{M}_0 \equiv (\Omega_0^{-1/2}S_0\Omega^{1/2})'M_0(\Omega_0^{-1/2}S_0\Omega^{1/2})$.

Lemma SA7. *Under Assumptions 1, 2, 3 and 4, we have uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$:*
(a) $n^{1/2}(\hat{\theta}^* - \theta_0) = -(Q'\Omega^{-1}Q)^{-1}Q'\Omega^{-1}g(\theta_0) - \left(Q_0'\Omega_0^{-1}Q_0\right)^{-1} Q_0'\Omega_0^{-1}S_0\Omega^{1/2}M\upsilon^* + o_p(1)$;
(b) $L(\upsilon^*, \hat{d}) = \upsilon^{*\prime}(M - \tilde{M}_0)\upsilon^* + o_p(1)$;
(c) $\upsilon^{*\prime}(M - \tilde{M}_0)\upsilon^* \sim \chi^2_{k_1}$.

Proof of Theorem 2. (i) Under Assumptions 1 – 3, Lemma SA1 gives

$$g(\hat{\theta}) = \Omega^{1/2}M\Omega^{-1/2}g(\theta_0) + o_p(1) \quad \text{and} \quad \hat{\Omega} \equiv \hat{\Omega}(\hat{\theta}) = \Omega + o_p(1), \tag{SA.19}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0$. Let $\hat{\theta}_0 \equiv \arg\min_{\theta\in\Theta} g_0(\theta)'(\hat{\Omega}_0(\theta))^{-1}g_0(\theta)$. Adding Assumption 4, we have

$$g_0(\hat{\theta}_0) = \Omega_0^{1/2}M_0\Omega_0^{-1/2}g_0(\theta_0) + o_p(1) \quad \text{and} \quad \hat{\Omega}_0 \equiv \hat{\Omega}_0(\hat{\theta}_0) = \Omega_0 + o_p(1) \tag{SA.20}$$

uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$, where $M_0 \equiv I_{k_0} - \Omega_0^{-1/2}Q_0(Q_0'\Omega_0^{-1}Q_0)^{-1}Q_0'\Omega_0^{-1/2\prime}$. Therefore, $\mathcal{T} \to_d \chi^2_{k_1}$ uniformly over $\mathbb{P} \in \mathcal{P}_0$ by the standard arguments in the literature (e.g., Eichenbaum, Hansen, and Singleton, 1988; Hall, 2005, Section 5).

We next prove part (ii). The critical value is simulated from

$$L(\upsilon^*, \hat{d}) = \upsilon^{*\prime}\hat{M}\upsilon^* - \left\|\hat{m}(\hat{\theta}^*) + \hat{V}(\hat{\theta}^*)\hat{\Omega}^{1/2}\hat{M}\upsilon^*\right\|^2_{\hat{\Omega}_0(\hat{\theta}^*)}. \tag{SA.21}$$

By Lemma SA7(b, c), we have uniformly over $\mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}$,

$$L(\upsilon^*, \hat{d}) = L^* + o_p(1), \text{ where } L^* \equiv \upsilon^{*\prime}(M - \tilde{M}_0)\upsilon^* \sim \chi^2_{k_1}. \tag{SA.22}$$

By (SA.22), there exists a positive sequence $\delta_n = o(1)$ such that for any $\varepsilon > 0$,

$$\mathbb{P}^*\left(|L(\upsilon^*, \hat{d}) - L^*| \geq \varepsilon/2\right) = o(\delta_n), \text{ uniformly over } \mathbb{P} \in \mathcal{P}_0 \cap \mathcal{P}_{00}, \tag{SA.23}$$

where $\mathbb{P}^* \equiv P^* \otimes \mathbb{P}$ denotes the product measure of υ^* and the data. Due to the independence between P^* and $\mathbb{P}$, for any $\varepsilon > 0$ and for all large n,

$$\mathbb{P}^*\left(|L(\upsilon^*, \hat{d}) - L^*| \geq \varepsilon/2 \middle| \hat{d}\right) \leq \delta_n \text{ wpa1}. \tag{SA.24}$$

Note that $c_\alpha(\hat{d})$ is the $1-\alpha$ conditional quantile of $L(\upsilon^*, \hat{d})$ given $\hat{d}$ and $L^* \sim \chi^2_{k_1}$ is independent of $\hat{d}$. Therefore, (SA.24) implies

$$q_{1-\alpha-\delta_n}(\chi^2_{k_1}) - \varepsilon/2 \leq c_\alpha(\hat{d}) \leq q_{1-\alpha+\delta_n}(\chi^2_{k_1}) + \varepsilon/2 \text{ wpa1} \tag{SA.25}$$

because by (SA.24) and the union bound of (conditional) probability, we have

$$\begin{aligned}\mathbb{P}^*\left(L(\upsilon^*, \hat{d}) > q_{1-\alpha+\delta_n}(\chi^2_{k_1}) + \varepsilon/2 \middle| \hat{d}\right) &\leq \mathbb{P}^*\left(L^* > q_{1-\alpha+\delta_n}(\chi^2_{k_1}) \middle| \hat{d}\right) + \delta_n = \alpha, \\ \mathbb{P}^*\left(L^* > c_\alpha(\hat{d}) + \varepsilon/2 \middle| \hat{d}\right) &\leq \mathbb{P}^*\left(L(\upsilon^*, \hat{d}) > c_\alpha(\hat{d}) \middle| \hat{d}\right) + \delta_n \leq \alpha + \delta_n.\end{aligned} \tag{SA.26}$$

Since $\delta_n = o(1)$ and $\chi^2_{k_1}$ is continuous with a strictly increasing quantile function, for all large n,

$$q_{1-\alpha-\delta_n}(\chi^2_{k_1}) - \varepsilon/2 \leq q_{1-\alpha}(\chi^2_{k_1}) \leq q_{1-\alpha+\delta_n}(\chi^2_{k_1}) + \varepsilon/2, \tag{SA.27}$$

which together with (SA.25) implies that, for any $\varepsilon > 0$, $|c_\alpha(\hat{d}) - q_{1-\alpha}(\chi^2_{k_1})| \leq \varepsilon$ wpa1. *Q.E.D.*

SB 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 SB1. *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* $\lambda_{\min}(\hat{Q}'\hat{Q}) \geq c_\lambda$ *wpa1.*

Assumption SB1(i) implies that there are globally misspecified moments in $\mathbb{E}[\bar{g}_1(\theta_0)] = 0_{k_1\times 1}$. Assumption SB1(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 SB1(ii) is the counterpart of Assumption 3(iii) under the alternative.

Theorem SB1. *Suppose Assumptions 1, 2 and SB1 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 SB1. We first show that the test statistic, written as $R(\hat{\xi})$, diverges at rate n under global misspecification. By Assumptions 2(i, iv) and SB1(ii),

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

where

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

By Assumption SB1(i), there exists a constant $c_g > 0$ such that $\min_{\theta\in\Theta} \|\mathbb{E}[\bar{g}(\theta)]\|^2 \geq c_g$, which combined with (SB.1), (SB.2), and Assumptions 1(i) and SB1(ii) implies that

$$R(\hat{\xi}) \geq n \left(K^{-1} \min_{\theta\in\Theta} \|\mathbb{E}[\bar{g}(\theta)]\|^2 - o_p(1) \right) \geq n c_g K^{-1} \text{ wpa1}. \tag{SB.3}$$

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 SB1(ii) and $q_{1-\alpha}(\chi_k^2)$ is the $1-\alpha$ quantile of $\|v^*\|^2$. Therefore, by (SB.3) and $c_\alpha(\hat{d}) \leq q_{1-\alpha}(\chi_k^2)$ wpa1, we have

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

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 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 SB2. *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 SB2. *Suppose Assumptions 1, 2, 4 and SB2 hold. For any $\mathbb{P} \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}$, we have*

$$\mathbb{P}\left(\mathcal{T} > c_\alpha(\hat{d})\right) \to P\left(\chi^2_{k_1}(a'_\Omega M a_\Omega) > q_{1-\alpha}(\chi^2_{k_1})\right), \text{ as } n \to \infty,$$

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 SB2. Under Assumptions 1 and 2, the strong identification in baseline moments in Assumption 4, and the local misspecification in Assumption SB2, $\hat{\theta}$ and $\hat{\theta}_0$ are consistent by the standard arguments and results in (SA.19) and (SA.20) remain valid. Therefore,

$$R(\hat{\xi}) \to_d (\upsilon^* + a_\Omega)' M (\upsilon^* + a_\Omega) - \upsilon_0^{*\prime} M_0 \upsilon_0^*, \tag{SB.5}$$

where $a_\Omega \equiv \Omega^{-1/2}a$ and υ_0^* denotes the leading k_0 subvector of υ^*. By the standard arguments in the GMM literature (e.g., Hall, 2005, Section 5), we have

$$(\upsilon^* + a_\Omega)' M (\upsilon^* + a_\Omega) - \upsilon_0^{*\prime} M_0 \upsilon_0^* \sim \chi^2_{k_1}(a'_\Omega M a_\Omega). \tag{SB.6}$$

We next study $c_\alpha(\hat{d})$ under the local misspecification. Since $\hat{\theta}$ is $n^{1/2}$ consistent under the local misspecification, Lemma SA7 remains valid for any $\mathbb{P} \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}$. Therefore, for any $\mathbb{P} \in \mathcal{P}_{00} \cap \mathcal{P}_{1,A}$,

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

By (SB.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 (SB.5) proves the claim of the theorem. *Q.E.D.*

As long as $a'_\Omega M a_\Omega > 0$, we have $P\left(\chi^2_{k_1}(a'_\Omega M a_\Omega) > q_{1-\alpha}(\chi^2_{k_1})\right) > \alpha$. 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 SB2 give

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

When the baseline moments provide strong identification, the conditional specification test is asymptotically equivalent to the C test following Theorems 2, SB1 and SB2.[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.

SC Comparison to Some Power Envelopes

In this section, we derive some power envelopes in a Gaussian experiment as in Section 4.1 in the paper and the baseline moments may only weakly identify the structural parameter θ_0. 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

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

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, $G \equiv (G_0', G_1')'$, G_0 and G_1 are $k_0 \times d_\theta$ and $k_1 \times d_\theta$ ($k_1 \geq d_\theta$) matrices respectively. We assume that G_1 has full rank, and $\Omega_0(\cdot,\cdot)$, Ω, G and the covariance between $g_{0,\infty}(\cdot)$ and $g_\infty(\theta_0)$ are known.

We are interested in testing

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

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

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

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

Power Envelopes. Let $G^\perp$ denote the orthogonal complement of G. It is clear that $G'\Omega^{-1}$ and $G^{\perp\prime}$ are the left eigenvectors of $I_k - G(G'\Omega^{-1}G)^{-1}G'\Omega^{-1}$ with respect to the (left) eigenvalues 0

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

and 1 respectively. Let $D = (G^{\perp}, \Omega^{-1}G)'$, then D is non-singular. Moreover,

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

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

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

We next consider inference of η based only on Y.

Since the uniformly most powerful (UMP) test does not exists for (SC.3), we follow Andrews and Mikusheva (2016) and derive several power envelopes by reducing the alternative hypothesis (SC.3) and/or imposing restrictions on the class of tests. If the alternative hypothesis (SC.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

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

It is clear that the optimality of the test in (SC.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 (SC.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 (SC.3) which are proportional to a known vector η^* with $A(\eta^*) \neq 0$, i.e., $H_1 : \eta = a\eta^*$. 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

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

In the simulation study below, we let the test in (SC.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

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

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

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

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

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

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

$$\mathcal{T} = g_{\infty}(\hat{\theta})'\Omega^{-1}g_{\infty}(\hat{\theta}) - \min_{\theta\in\Theta}(m_{0,\infty}(\theta) + V(\theta)g_{\infty}(\hat{\theta}))'(\Omega_0(\theta))^{-1}(m_{0,\infty}(\theta) + V(\theta)g_{\infty}(\hat{\theta})) \tag{SC.11}$$

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 $d_\theta = 1$, $k_0 = qk_1$ and

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

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

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

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

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

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

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

We generate the covariance matrix Ω as follows

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

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 $g_{1,\infty}(\theta_0)$ for $k_1 = 2$, while ρ mainly controls the correlations between moments in $g_{0,\infty}(\theta_0)$ and $g_{1,\infty}(\theta_0)$. In the case that $G_{0,\infty}$ is a non-degenerate normal random vector, we let

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

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

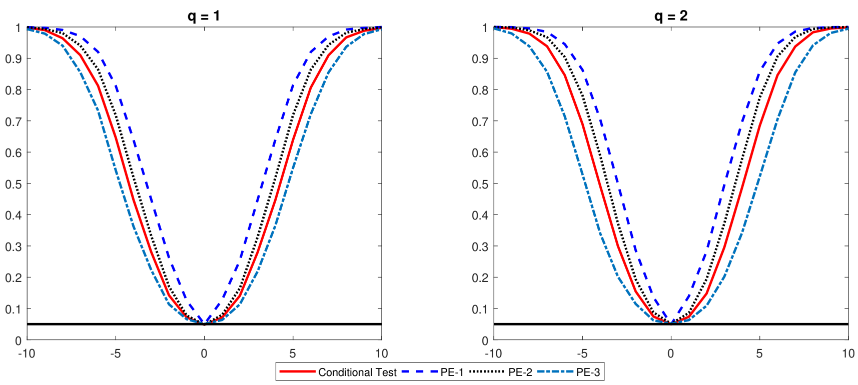
Throughout this simulation, we let $\theta_0 = 0$, $\Theta = [-1, 1]$, $G_0 = c_g 1_{k_0\times 1}$, $G_1 = (j^{-1})_{j=1,\ldots,k_1}$, $\mu_g = c_\mu 1_{k_0\times 1}$, $\eta = a 1_{k_1\times 1}$, $k_1 = 2$ and $\lambda = 0.1$. We consider a benchmark case and three deviations from the benchmark, which are defined as follows.

Benchmark case:	$\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, nonrandom $G_{0,\infty}$, $q = k_0/k_1 = 1$ 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 $G_{0,\infty}$.

In the benchmark case, we set $c_g = 0$ to model weak baseline moments whose derivatives G_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 $G_{0,\infty}$ 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.

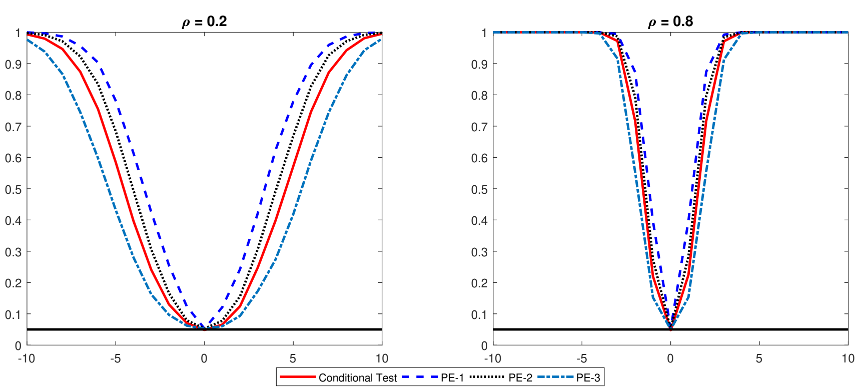
Discussion. In all cases, the power of the proposed conditional specification test is between PE-2 and PE-3 (J-test). PE-2 is the power of the UMP unbiased test with respect to a smaller subset of the general alternative hypothesis in (SC.3) and it is constructed using the true alternative value η, whereas 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 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



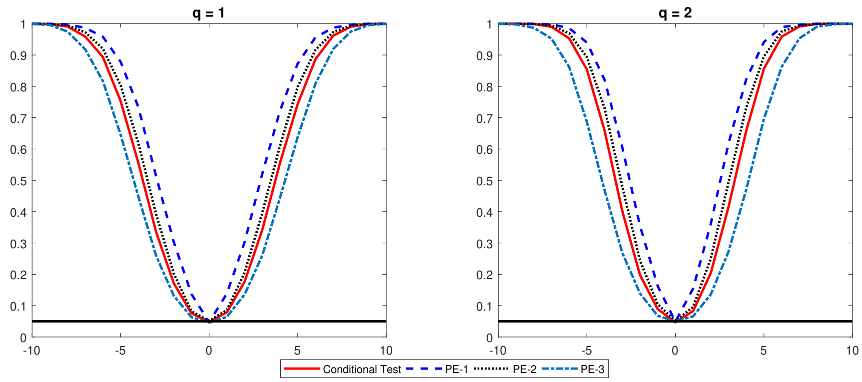
Note: In the Benchmark case, we have $\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, nonrandom $G_{0,\infty}$, $q = k_0/k_1 = 1$ or 2.

Figure 2: Power Comparison in the Deviation Case 1



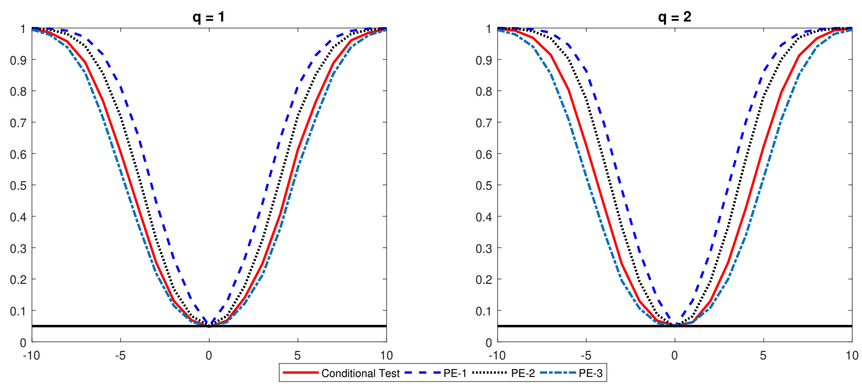
Note: In the deviation case 1, we have $\rho = 0.2$ or 0.8, $c_g = 0$, $c_\mu = 1$, non-random $G_{0,\infty}, q = 1$.

Figure 3: 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 $G_{0,\infty}$, $q = 1$ or 2.

Figure 4: Power Comparison in the Deviation Case 3



Note: In the deviation case 4, we have random $G_{0,\infty}$, $\rho = 0.4$, $c_g = 0$, $c_\mu = 1$, and $q = 1$ or 2.

SD Additional Details of the Empirical Application

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

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

In the model, $\bar{g}_0(\theta)$ only depends on a subvector of θ. We have 6 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

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

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

$$\Theta_1 \equiv [0.001, 0.02],\ \Theta_2 \equiv [5, 12],\ \ \Theta_3 \equiv [0.95, 0.999], \text{ and } \ \Theta_4 \equiv [3, 6]. \tag{SD.3}$$

To compute the CUE estimator, the J statistic and the test 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 slightly smaller parameter space $\Theta_2 \equiv [5, 8]$ for $\theta_{2,0}$ and keep Θ_j $(j = 1, 3, 4)$ unchanged to slightly reduce the computational cost. The reduced space Θ_2 still covers the the CUE estimators of $\theta_{2,0}$ 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.

Cheng, X., W. W. Dou, and Z. Liao (2020): "Supplemental Appendix to 'Macro-Finance Decoupling: Robust Evaluations of Macro Asset Pricing Models'," *working paper, http://ssrn.com/abstract=3609598, also available on the authors' personal websites.*

Eichenbaum, M. S., L. P. Hansen, and K. J. Singleton (1988): "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty," *The Quarterly Journal of Economics*, 103(1), 51–78.

Hall, A. R. (2005): *Generalized Method of Moments*. Oxford university press.

Mikusheva, A. (2007): "Uniform Inference in Autoregressive Models," *Econometrica*, 75(5), 1411–1452.

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

Phillips, P. C. B. (1987): "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74(3), 535–547.

[2]We have also considered a much larger parameter space with $\Theta_1 \equiv [0.001, 0.02]$, $\Theta_2 \equiv [1, 15]$, $\Theta_3 \equiv [0.90, 0.999]$, and $\Theta_4 \equiv [1, 20]$, and step size 0.001 in Θ_1 and Θ_3, and step size 0.1 in Θ_2 and Θ_4. The results on the CUE estimators and the J tests are very similar to those reported in Table 1 of the paper.