

Is Your Macro Model Matching the Targets?*

Joan Alegre

Juan Carlos Escanciano

Universidad Carlos III de Madrid

Universidad Carlos III de Madrid

September 20, 2026

Abstract

Testing whether a macroeconomic model can match its empirical targets is difficult when those targets come from multiple databases, structural parameters are not point identified, and the covariance matrix of the target estimators may be singular. We develop a minimum distance correct specification test that is robust to all three complications. Because cross-database covariances are unavailable, the test uses the worst-case critical value over all positive-semidefinite covariance completions consistent with the observable within-database blocks. This is the smallest critical value that guarantees asymptotic size control for every admissible dependence structure. We develop a residual-space reparameterization that reduces the dimensionality of the optimization problem and facilitates computation of the worst-case critical value. We illustrate the test through Monte Carlo simulations based on the New Keynesian model of [An and Schorfheide \(2007\)](#) and an empirical application to the model of [Nakamura and Steinsson \(2018\)](#).

Keywords: Minimum Distance estimation; Partial identification; Multiple databases; Specification testing.

JEL classification: C10, C12

*Research funded by Ministerio de Ciencia e Innovación, grant ECO2017-86675-P, MCI/AEI/FEDER/UE, grant PGC 2018-096732-B-I00, grant PID2021-127794NB-I00 and Comunidad de Madrid, grants EPUC3M11 (VPRICIT) and H2019/HUM-589.

1 Introduction

A common practice in macroeconomics is to use a structural model to match a set of empirically estimated targets. Researchers often fix some structural parameters and choose the remaining ones to minimize the distance between the empirical targets and their model-implied counterparts. This procedure can be viewed as minimum distance estimation (Gregory and Smith (1990); Hansen and Heckman (1996); Cocci and Plagborg-Møller (2025)). An important question is whether the resulting restricted model can match the targets at the population level. The restricted model is correctly specified with respect to these targets if there is at least one value of the remaining parameters for which the model-implied targets coincide with the population targets.

There is a well-established literature on specification testing in minimum distance and related moment-condition models (e.g., Sargan (1958); Hansen (1982); Shapiro (1986); Andrews and Mikusheva (2016)). Yet three features common in macroeconomic applications can make standard methods difficult to use. First, different targets are sometimes estimated from different databases, and the cross-covariances between their estimators may be unavailable (e.g., Cocci and Plagborg-Møller (2025); Vohra (2026); Cha et al. (2026)). Second, the structural parameters being estimated may not be point identified (e.g., Iskrev (2010); Komunjer and Ng (2011); Qu and Tkachenko (2012, 2017, 2023); Kociecki and Kolasa (2018, 2023)). Third, in applications such as Impulse Response Function Matching (IRFM), the asymptotic covariance matrix of the estimated targets may be singular (e.g., Inoue and Kilian (2016); Guerron-Quintana et al. (2017)). This paper proposes a correct specification test that is robust to all three features and can be applied to a broad class of structural economic models estimated by minimum distance.

Our test combines a standard Minimum distance statistic with a worst-case critical value. Under regularity conditions similar to those in Shapiro (1986), the statistic converges under the null to a generalized chi-square distribution that depends on the full asymptotic covariance matrix of the estimated targets. Since only the within-database covariance blocks are consistently estimable, we consider all positive-semidefinite covariance matrices that agree with these blocks. We define the worst-case critical value as the largest $(1 - \tau)$ -quantile across the resulting family of null distributions. For this statistic, it is the smallest critical value that controls the asymptotic rejection probability at level τ for every admissible cross-database covariance structure.

We illustrate the test through a Monte Carlo simulation and an empirical application. In simulations based on the New Keynesian model of An and Schorfheide (2007), our test is substantially less conservative than the procedure based on the Székely–Bakirov upper bound, while test sizes are controlled. Power increases with the magnitude of misspecification and the sample size. We then revisit the monetary policy and information-effects model of Nakamura and Steinsson (2018), combining interest-rate, inflation, and expected-growth responses estimated from two databases. The test does not reject the joint fit of the restricted model at the 5%

level, without assuming independence across databases or requiring point identification of the structural parameters. The application also illustrates the computational gains from our reparameterization, which reduces the critical-value search from 128 unknown cross-covariances to 56 residual cross-correlations.

This paper is related to several strands of literature. It builds most directly on work on inference with incomplete covariance information. The closest paper to ours is [Cocci and Plagborg-Møller \(2025\)](#). Their framework imposes point-identification conditions on the estimated parameters, while ours allows for partial identification. Another important difference concerns the construction of the worst-case critical value. For their joint over-identification test, [Cocci and Plagborg-Møller \(2025\)](#) use the Székely–Bakirov inequality to obtain an upper bound on this critical value, computed through semidefinite programming. This bound can introduce additional conservativeness. This paper instead proposes a reparameterization of the covariance matrix underlying the generalized chi-square distribution that yields an equivalent lower-dimensional maximization problem for the sharp worst-case critical value.

A second connection is to the literature on inference with nuisance parameters that are not point identified (e.g., [Davies, 1977](#); [Shapiro, 1986](#); [Andrews and Ploberger, 1994](#); [Hansen, 1996](#)). Our analysis builds on [Shapiro \(1986\)](#). Shapiro shows that point identification of the structural nuisance parameter is not required for minimum distance correct specification testing. What matters instead is the local geometry of the set of model-implied reduced-form parameters. Following Shapiro, we assume that the set of model-implied targets is locally a smooth manifold, with dimension given by the Jacobian rank. Relative to Shapiro, however, we allow the asymptotic covariance matrix of the reduced-form estimators to be singular and only partially known because the cross-database covariance blocks are unavailable, which motivates our sharp worst-case critical value.

Our setting also intersects the literature on inference under weak identification (e.g., [Staiger and Stock \(1994\)](#); [Kleibergen \(2005\)](#); [Stock and Wright \(2000\)](#); [Antoine and Renault \(2009\)](#); [Andrews and Mikusheva \(2016\)](#); [Andrews and Guggenberger \(2019\)](#); [Forneron \(2024\)](#)). Weak-identification procedures seek size control uniformly over sequences in which identification strength may drift from strong to weak or fail in the limit. In contrast, we study a simpler identification setting: under the regularity conditions described above, parameter combinations are either strongly identified or exactly unidentified, without allowing identification strength to deteriorate as the sample size grows. However, we address a problem outside the scope of standard weak-identification methods: incomplete covariance information across databases.

The remainder of the paper is organized as follows. Section [2](#) introduces the minimum distance framework with partially identified structural parameters, possibly singular asymptotic covariance matrices, and targets estimated from multiple databases with unknown cross-database covariances. Section [3](#) develops the specification test, derives its generalized chi-square null

distribution, constructs the sharp worst-case critical value, and section 4 studies the power under fixed and local alternatives. Section 5 provides an exact residual-space reparameterization of the worst-case critical-value problem that reduces the dimension of the required optimization. Section 6 presents the feasible test, including estimation of the relevant ranks and residual covariance factors. Section 7 studies the finite-sample performance of the procedure in an impulse-response-matching exercise based on the model of An and Schorfheide (2007). Section 8 applies the test to Nakamura and Steinsson (2018). Section 9 concludes. Appendix A provides additional implementation details and a complete algorithm, while Appendix B contains the proofs.

2 Framework

This section describes a setting characterized by three departures from the standard minimum distance framework that commonly arise in applications. First, the estimated structural parameters may be partially identified. Second, the asymptotic covariance matrices of the reduced-form estimators may be singular. Third, the reduced-form parameters may be estimated from several databases, so that their within-database covariance matrices are consistently estimable while their between-database covariances are unknown.

Let θ_0 be a vector of reduced-form parameters used as targets in minimum distance estimation:

$$\theta_0 = \left(\theta_0^{(1)'} , \dots , \theta_0^{(B)'} \right)' \in \mathbb{R}^m, \quad \theta_0^{(b)} \in \mathbb{R}^{m_b}, \quad \sum_{b=1}^B m_b = m,$$

where the reduced-form parameters in block b are estimated using database b . Let the corresponding estimators be

$$\hat{\theta} = \left(\hat{\theta}^{(1)'} , \dots , \hat{\theta}^{(B)'} \right)'.$$

As in Cocci and Plagborg-Møller (2025) we study settings in which the estimators obtained from the different databases admit a joint Gaussian limit.

Assumption 1 (Reduced-form asymptotics)

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} Z, \quad Z \sim N(0, \Sigma_0),$$

where $\Sigma_0 \succeq 0$ may be singular.¹

¹The common normalization n is used for notational simplicity. Differences in the sample sizes of the databases can be incorporated into the corresponding blocks of Σ_0 .

Partition the asymptotic covariance matrix conformably as

$$\Sigma_0 = \begin{pmatrix} \Sigma_{1,0} & \cdots & \Sigma_{1B,0} \\ \vdots & \ddots & \vdots \\ \Sigma_{B1,0} & \cdots & \Sigma_{B,0} \end{pmatrix}.$$

For each database, the within-database covariance matrix $\Sigma_{b,0} \succeq 0$ is consistently estimable. In contrast, the cross-database blocks $\Sigma_{bc,0}$, $b \neq c$, are unknown. Importantly, we do not assume that estimators obtained from different databases are independent. We assume only that their joint asymptotic distribution exists, while leaving the cross-database covariance structure unrestricted.

Let

$$\alpha \in \mathcal{A} \subseteq \mathbb{R}^q, \quad \beta \in \mathcal{B} \subseteq \mathbb{R}^p,$$

denote the structural parameters. The structural model implies the mapping

$$g : \mathcal{A} \times \mathcal{B} \longrightarrow \mathbb{R}^m,$$

where α denotes the estimated parameters and β the fixed parameters. Consider the case in which the fixed parameters are calibrated to $\bar{\beta} \in \mathcal{B}$. The set of reduced-form parameters that can be generated by the restricted model is

$$\Theta_0(\bar{\beta}) \equiv \{g(\alpha, \bar{\beta}) : \alpha \in \mathcal{A}\}. \quad (1)$$

A central feature of our framework is that the estimated parameter α need not be point identified. Under the null, the identified set

$$\mathcal{A}_0(\theta_0, \bar{\beta}) \equiv \{\alpha \in \mathcal{A} : \theta_0 = g(\alpha, \bar{\beta})\}$$

may contain more than one element. This is a major departure from current methods for inference with incomplete cross-database covariance blocks, such as [Cocci and Plagborg-Møller \(2025\)](#), where point identification of the estimated parameters is required.

The following conditions replace point identification with local regularity of the restricted-model set: they ensure that $\Theta_0(\bar{\beta})$ is locally an r_α -dimensional C^2 manifold around θ_0 .

Assumption 2 (Local manifold regularity) *There exist $\alpha_0 \in \mathcal{A}_0(\theta_0, \bar{\beta}) \cap \text{int}(\mathcal{A})$, a neighborhood U of α_0 , and a neighborhood V of θ_0 , such that*

1. *The mapping $g(\cdot, \bar{\beta})$ is twice continuously differentiable on U .*
2. *For any $\alpha \in U$, $\text{rank}\{\nabla_\alpha g(\alpha, \bar{\beta})\} = r_\alpha \leq q$.*

$$3. \Theta_0(\bar{\beta}) \cap V = \{g(\alpha, \bar{\beta}) : \alpha \in U\}.$$

Define the Jacobian

$$G_0 \equiv \nabla_{\alpha} g(\alpha_0, \bar{\beta}) = \frac{\partial g(\alpha_0, \bar{\beta})}{\partial \alpha'} \in \mathbb{R}^{m \times q}. \quad (2)$$

The rank of this Jacobian, $r_{\alpha} \equiv \text{rank}(G_0)$, determines the local dimension of the manifold around θ_0 and can be interpreted as the maximum number of parameters that can be locally identified. [Shapiro \(1986\)](#) develops a specification test that is robust to lack of identification under primitive conditions that imply Assumption [2](#).

3 Test

In this section we develop a conservative test for testing correct specification of the restricted model:

$$H_0 : \exists \alpha \in \mathcal{A} \text{ s.t. } \theta_0 = g(\alpha, \bar{\beta}) \quad \text{vs.} \quad H_1 : \nexists \alpha \in \mathcal{A} \text{ s.t. } \theta_0 = g(\alpha, \bar{\beta}). \quad (3)$$

We first characterize the asymptotic null distribution of the test statistic. We then construct critical values that control size for every covariance matrix that agrees with the known within-database covariance blocks. Finally, we study the power under fixed and local alternatives.

3.1 Test Statistic and Null Distribution

Consider the standard minimum distance statistic for testing these hypotheses:

$$T_n \equiv n \min_{\alpha \in \mathcal{A}} \left(\hat{\theta} - g(\alpha, \bar{\beta}) \right)' \widehat{W} \left(\hat{\theta} - g(\alpha, \bar{\beta}) \right), \quad (4)$$

where $\widehat{W}$ is an estimated weighting matrix satisfying the following condition.

Assumption 3 (Weighting matrix) *The matrix $\widehat{W}$ is symmetric and*

$$\|\widehat{W} - W\|_{\text{op}} \xrightarrow{p} 0,$$

where $W = W' \succ 0$ is deterministic.

Theorem 3.1 (Statistic's distribution) *Suppose Assumptions [1–3](#) hold. Then under H_0*

$$T_n \xrightarrow{d} \sum_{j=1}^{r_Q} \lambda_j \chi_{1,j}^2,$$

where $\lambda_1, \dots, \lambda_{r_Q}$ are the strictly positive eigenvalues of $R_W^{1/2} \Sigma_0 R_W^{1/2}$, with:

$$R_W := W - WG_0(G_0'WG_0)^\dagger G_0'W,$$

where $\dagger$ denotes the Moore-Penrose g -inverse. The $\chi_{1,j}^2$'s are mutually independent, and $r_Q = \text{rank} \left(R_W^{1/2} \Sigma_0 R_W^{1/2} \right)$. When $r_Q = 0$, the limiting variable is identically zero.

Proof. See Appendix B ■

The limiting distribution is therefore a central generalized chi-square distribution. Both the number of nonzero terms and their weights may depend on the unknown cross-database covariances of Σ_0 . Singularity of Σ_0 may generate additional zero eigenvalues but does not invalidate the representation.

Remark 3.1 (Relation to Cocci and Plagborg-Møller (2025)) *Under the point-identification conditions of Cocci and Plagborg-Møller (2025), G_0 has full column rank and $G_0'WG_0$ is non-singular. Hence, $(G_0'WG_0)^\dagger = (G_0'WG_0)^{-1}$, and the null distribution in Theorem 3.1 coincides with the Gaussian residual quadratic form underlying their joint overidentification test.*

Remark 3.2 (Relation to Shapiro (1986)) *The quadratic-form null limit in Theorem 3.1 is a specialization of the result in Shapiro (1986) for regular overparameterized models. To implement the chi-square test emphasized there, however, the full asymptotic covariance matrix must be available so that an appropriate discrepancy weighting can be constructed. In our setting, only the within-database blocks of Σ_0 are consistently estimable, so the unknown cross-database covariances leave the null distribution and its critical value only partially determined. This motivates the worst-case critical-value construction developed below.*

The following subsection constructs critical values that remain valid over all positive-semidefinite covariance completions compatible with the known within-database blocks.

3.2 Worst-Case Critical Value

The null distribution derived above depends on the full asymptotic covariance matrix Σ_0 . Because only its within-database blocks are identified, the available covariance information does not determine a unique critical value. We therefore consider all covariance matrices that are compatible with the known within-database blocks and compute the critical value of the test by taking the worst-case $(1 - \tau)$ -quantile over all admissible covariance completions.

Define the covariance-completion set

$$\mathcal{S}(\Sigma_1, \dots, \Sigma_B) \equiv \left\{ \tilde{\Sigma} \in \mathbb{S}_+^m : \tilde{\Sigma}_{bb} = \Sigma_b, \quad b = 1, \dots, B \right\}, \quad (5)$$

where $\mathbb{S}_+^m$ denotes the set of $m \times m$ symmetric positive-semidefinite matrices.

For any $A \in \mathbb{S}_+^m$, define

$$q_{1-\tau}(A) \equiv \inf \{c \in \mathbb{R} : \Pr(\varepsilon' A \varepsilon \leq c) \geq 1 - \tau\}, \quad \varepsilon \sim N(0, I_m).$$

Thus, $q_{1-\tau}(A)$ is the $(1 - \tau)$ -quantile of the generalized chi-square distribution generated by A . The sharp worst-case asymptotic critical value is therefore

$$c_{1-\tau}^*(R_W) \equiv \sup_{\tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B)} q_{1-\tau} \left(R_W^{1/2} \tilde{\Sigma} R_W^{1/2} \right). \quad (6)$$

The critical value $c_{1-\tau}^*(R_W)$ is the smallest critical value that controls the asymptotic rejection probability at level τ for every admissible covariance completion.

The supremum problem in (6) is generally a high-dimensional, nonconcave maximization problem and may therefore be difficult to solve. Previous work, such as [Cocci and Plagborg-Møller \(2025\)](#), addresses this difficulty by deriving upper bounds on the sharp worst-case critical value. The drawback of this strategy is that it may produce an even more conservative test. To avoid relying on such upper bounds, Section 5 reparameterizes the original high-dimensional problem as an equivalent lower-dimensional maximization problem.

Remark 3.3 (Size Loss Due to Conservativeness) *The actual asymptotic null rejection probability under Σ_0 is*

$$\tau_{\text{act}}(\Sigma_0; R_W) \equiv 1 - F_{Q_0}(c_{1-\tau}^*(R_W)).$$

We measure pointwise conservativeness by the relative size-loss index

$$\mathcal{L}_\tau(\Sigma_0; R_W) \equiv 1 - \frac{\tau_{\text{act}}(\Sigma_0; R_W)}{\tau} \quad (7)$$

The index satisfies

$$0 \leq \mathcal{L}_\tau(\Sigma_0; R_W) \leq 1.$$

A value of zero means that the test uses the entire nominal size, whereas a value close to one means that it almost never rejects under the null.

4 Power

Define the asymptotic power function, whenever the limit exists, as

$$\pi_\tau(\theta) \equiv \lim_{n \rightarrow \infty} \mathbb{P}_\theta\{T_n > c_{1-\tau}^*(R_W)\},$$

where $\mathbb{P}_\theta$ denotes a fixed data-generating process with population target θ . For the fixed alternatives considered below, we assume that $\hat{\theta}_n \xrightarrow{p} \theta$ under $\mathbb{P}_\theta$.

Define the set of positively separated alternatives as

$$\Theta_1 \equiv \left\{ \theta \in \mathbb{R}^m : \inf_{\mu \in \Theta_0(\bar{\beta})} (\theta - \mu)' W (\theta - \mu) > 0 \right\}.$$

The following theorem establishes consistency against every fixed alternative in this set.

Theorem 4.1 (Consistency against fixed alternatives) *Fix $\theta_1 \in \Theta_1$. Suppose that $\hat{\theta}_n \xrightarrow{p} \theta_1$ and Assumption 3 holds. Then, for every $\tau \in (0, 1)$ such that $c_{1-\tau}^*(R_W) < \infty$,*

$$\pi_\tau(\theta_1) = 1.$$

Proof. See Appendix B. ■

If $\Theta_0(\bar{\beta})$ is closed, then $\Theta_1 = \mathbb{R}^m \setminus \Theta_0(\bar{\beta})$, so the test is consistent against every fixed alternative.

We next study power to local alternatives. Fix $\theta_0 \in \Theta_0(\bar{\beta})$ and consider a sequence of data-generating processes $\mathbb{P}_n$ with population targets

$$\theta_n = \theta_0 + \frac{h}{\sqrt{n}} + o(n^{-1/2}), \quad h \in \mathbb{R}^m.$$

Such a sequence is a local alternative when $\theta_n \notin \Theta_0(\bar{\beta})$ for all sufficiently large n . Define the asymptotic local power function, whenever the limit exists, as

$$\pi_\tau^{\text{loc}}(h) \equiv \lim_{n \rightarrow \infty} \mathbb{P}_n\{T_n(\bar{\beta}) > c_{1-\tau}^*(R_W)\},$$

where R_W and c are evaluated at the baseline θ_0 . Along these sequences, assume that

$$\sqrt{n}(\hat{\theta}_n - \theta_n) \xrightarrow{d} Z \sim N(0, \Sigma_0) \quad \text{under } \mathbb{P}_n,$$

with the same limiting covariance Σ_0 as at the null baseline.

Theorem 4.2 (Local power) *Suppose the preceding Gaussian limit holds, Assumption 2 holds at θ_0 , and Assumption 3 holds. If $r_Q > 0$, then*

$$T_n(\bar{\beta}) \xrightarrow{d} (Z + h)' R_W (Z + h), \quad \pi_\tau^{\text{loc}}(h) = \Pr\{(Z + h)' R_W (Z + h) > c\}.$$

Moreover, $\pi_\tau^{\text{loc}}(h) \geq \pi_\tau^{\text{loc}}(0)$, with equality if and only if $h \in \mathcal{T}_0 \equiv \text{col}(G_0)$.

Proof. See Appendix B. ■

Since $\mathcal{T}_0 = \ker(R_W)$, tangent movements can be absorbed to first order by re-estimating α and leave the limiting rejection probability unchanged. Every non-tangent direction instead increases rejection above the actual asymptotic null probability $\pi_\tau^{\text{loc}}(0) \leq \tau$.

4.1 Local Power Due to Miscalibration

Our test concerns whether the population targets belong to $\Theta_0(\bar{\beta})$, rather than whether the calibrated parameters equal their true values. A miscalibration therefore induces misspecification only when its effect on the targets cannot be accommodated by re-estimating α .

Fix a baseline $\theta_0 = g(\alpha_0, \bar{\beta}) \in \Theta_0(\bar{\beta})$ and consider a sequence of data-generating processes $\mathbb{P}_n$ with population targets

$$\theta_n = g(\alpha_n, \beta_n), \quad \alpha_n = \alpha_0 + \frac{a}{\sqrt{n}}, \quad \beta_n = \bar{\beta} + \frac{d}{\sqrt{n}}, \quad (8)$$

for fixed admissible directions a and d , while the model used for testing continues to fix $\beta = \bar{\beta}$.

Define

$$D_0 = \nabla_{\beta} g(\alpha_0, \bar{\beta}).$$

The first-order movement in the targets is $G_0 a + D_0 d$. Since re-estimation of α absorbs movements in $\text{col}(G_0)$, only the component of $D_0 d$ outside this space can generate first-order local power.

Similarly to the previous section, define the asymptotic local power along this sequence as

$$\pi_\tau^{\text{cal}}(a, d) \equiv \lim_{n \rightarrow \infty} \mathbb{P}_n \{T_n(\bar{\beta}) > c_{1-\tau}^*(R_W)\}.$$

Lemma 4.3 (Local power against miscalibration) *Suppose the conditions of Theorem 4.2 hold along the sequence above and g is jointly differentiable at $(\alpha_0, \bar{\beta})$. Then*

$$\pi_\tau^{\text{cal}}(a, d) = \pi_\tau^{\text{loc}}(D_0 d) = \Pr\{(Z + D_0 d)' R_W (Z + D_0 d) > c_{1-\tau}^*(R_W)\},$$

so the limiting rejection probability does not depend on a . Moreover,

$$\pi_\tau^{\text{loc}}(D_0 d) > \pi_\tau^{\text{loc}}(0) \iff D_0 d \notin \text{col}(G_0).$$

Consequently, directions in

$$\ker(R_W D_0) = \ker(D_0' R_W D_0)$$

have no first-order local power, and their limiting rejection probability equals the actual asymptotic null rejection probability $\pi_\tau^{\text{loc}}(0) \leq \tau$.

Proof. See Appendix B. ■

5 Residual-Space Reparameterization

One important limitation of the approach described above is that computing the sharp worst-case critical value requires solving a potentially high-dimensional maximization problem. Cocco and Plagborg-Møller (2025) address this difficulty by deriving an upper bound on the sharp critical value. However, using such an upper bound makes the test more conservative and may reduce its local power. In this section, we develop an exact reparameterization of the problem that reduces the dimension of the maximization problem. For the detailed discussion of the following result check Appendix A.1

Write $R_W = H'H$, where $H = (H_1, \dots, H_B) \in \mathbb{R}^{(m-r_\alpha) \times m}$ has full row rank. For any admissible covariance completion $\tilde{\Sigma}$, let $Z = (Z'_1, \dots, Z'_B)' \sim N(0, \tilde{\Sigma})$ and define

$$Y_b \equiv H_b Z_b, \quad Y \equiv \sum_{b=1}^B Y_b = HZ.$$

Since $Z'R_W Z = Y'Y$, the limiting statistic depends only on the aggregate residual Y . Each database contributes a residual covariance matrix

$$\Omega_b \equiv \text{Var}(Y_b) = H_b \Sigma_b H_b' = L_b L_b', \quad s_b \equiv \text{rank}(\Omega_b),$$

where $L_b \in \mathbb{R}^{(m-r_\alpha) \times s_b}$ is a thin square root.

Writing $Y_b = L_b U_b$ with $\text{Var}(U_b) = I_{s_b}$, the unknown dependence across databases is described by the covariance matrix C of $(U'_1, \dots, U'_B)'$. Its admissible block-correlation set is

$$\mathcal{C} \equiv \left\{ C \in \mathbb{S}_+^s : C_{bb} = I_{s_b}, \quad b = 1, \dots, B \right\}, \quad s \equiv \sum_{b=1}^B s_b.$$

For $L \equiv (L_1, \dots, L_B)$, the aggregate residual covariance is therefore $\text{Var}(Y) = LCL'$.

Theorem 5.1 (Correlation reparameterization) *The sharp worst-case critical value admits the exact representation*

$$c_{1-\tau}^*(R_W) = \max_{C \in \mathcal{C}} q_{1-\tau}(LCL'). \quad (9)$$

Proof. See Appendix B. ■

Remark 5.1 (Dimensionality reduction) *The original problem contains $\sum_{b < c} m_b m_c$ unknown cross-covariance entries, whereas (9) contains at most*

$$\sum_{b < c} s_b s_c, \quad s_b \leq \min \{m - r_\alpha, \text{rank}(\Sigma_b)\},$$

unknown cross-correlations.

At first order, re-estimating α absorbs movements of the targets along the model's tangent space. The reparameterization retains only the variation that cannot be absorbed by these adjustments and searches over its unknown cross-database correlations. It therefore removes covariance information irrelevant to the limiting statistic without changing the worst-case critical value.

6 Feasible Test

The feasible test replaces the population objects in Section 5 with sample counterparts. Compute T_n from the unpenalized minimum distance problem in (4), and select a regular estimate $\hat{\alpha}$ for evaluating $\hat{G} \equiv \nabla_\alpha g(\hat{\alpha}, \bar{\beta})$, as described in Appendix A.2.

A hard-thresholded singular-value decomposition of $\widehat{W}^{1/2} \hat{G}$ yields an estimated tangent rank $\hat{r}_\alpha$ and an orthonormal basis $\hat{U}_R$ for the corresponding weighted residual space. Define

$$\hat{H} \equiv \hat{U}_R' \widehat{W}^{1/2} = (\hat{H}_1, \dots, \hat{H}_B) \in \mathbb{R}^{(m - \hat{r}_\alpha) \times m}.$$

For each database, form $\hat{\Omega}_b \equiv \hat{H}_b \hat{\Sigma}_b \hat{H}_b'$. Hard thresholding its eigenvalues yields an estimated rank $\hat{s}_b$ and a thin factor $\hat{L}_b \in \mathbb{R}^{(m - \hat{r}_\alpha) \times \hat{s}_b}$.

Collect the nonempty factors in $\hat{L} \equiv (\hat{L}_b : \hat{s}_b > 0)$, let $\hat{s} \equiv \sum_{b=1}^B \hat{s}_b$, and define

$$\hat{\mathcal{C}} \equiv \{C \in \mathbb{S}_+^{\hat{s}} : C_{bb} = I_{\hat{s}_b} \text{ for every retained block } b\}.$$

When $m - \hat{r}_\alpha > 0$ and $\hat{s} > 0$, the estimated worst-case critical value is

$$\hat{c}_{1-\tau} \equiv \max_{C \in \hat{\mathcal{C}}} q_{1-\tau}(\hat{L} C \hat{L}').$$

Otherwise, set $\hat{c}_{1-\tau} = 0$ and do not reject. The feasible test is therefore

$$\hat{\phi}_n(\bar{\beta}) \equiv \mathbf{1}\{m - \hat{r}_\alpha > 0, \hat{s} > 0\} \mathbf{1}\{T_n > \hat{c}_{1-\tau}\}.$$

Under the conditions of Proposition A.4 in Appendix A.4,

$$\widehat{c}_{1-\tau} \xrightarrow{p} c_{1-\tau}^*(R_W), \quad \limsup_{n \rightarrow \infty} \Pr\{\widehat{\phi}_n(\bar{\beta}) = 1\} \leq \tau$$

for every admissible null covariance completion. Appendix A.4 states the conditions for this result, and Appendix A.5 summarizes the implementation.

7 Monte Carlo Simulation: An and Schorfheide (2007)

We study the finite-sample behavior of the test in an impulse-response-matching exercise based on the linearized New Keynesian model of An and Schorfheide (2007). The example combines twelve targets, two databases with unknown cross-database covariance, and one locally unidentified structural parameter. We examine rejection under correct specification and power against local departures from the model.

Let x_t denote detrended output, π_t inflation, R_t the nominal interest rate, z_t the productivity-growth disturbance, and g_t the government-spending disturbance. The no-spillover model is

$$x_t = E_t x_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau} (R_t - E_t \pi_{t+1} - E_t z_{t+1}), \quad (10)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(x_t - g_t), \quad (11)$$

$$R_t = \rho_m R_{t-1} + (1 - \rho_m) \{\psi_1 \pi_t + \psi_2 (x_t - g_t)\} + \sigma_m \varepsilon_{m,t}, \quad (12)$$

$$z_t = \rho_z z_{t-1} + \sigma_z \varepsilon_{z,t}, \quad g_t = \rho_g g_{t-1} + \sigma_g \varepsilon_{g,t}, \quad (13)$$

where $(\varepsilon_{m,t}, \varepsilon_{z,t}, \varepsilon_{g,t})' \stackrel{\text{i.i.d.}}{\sim} N(0, I_3)$, and $\varepsilon_{m,t}$, $\varepsilon_{z,t}$, and $\varepsilon_{g,t}$ denote the monetary policy, productivity, and government spending structural shocks, respectively. The data-generating process (DGP) is parameterized as follows:

$$(\tau, \beta, \kappa, \psi_1, \psi_2) = (2, 0.9975, 0.33, 1.5, 0.125),$$

$$(\rho_m, \rho_z, \rho_g) = (0.75, 0.90, 0.95), \quad (\sigma_m, \sigma_z, \sigma_g) = (0.20, 0.30, 0.60).$$

Using this parameterization, we compute impulse responses to one-standard-deviation structural innovations. For $j \in \{m, z, g\}$, let R_h^j , x_h^j , and π_h^j denote the responses of (R_t, x_t, π_t) at horizon h to shock j . The twelve targets are divided evenly between two databases:

$$\theta_0^{(1)} = (R_0^m, x_0^m, \pi_0^m, R_1^m, x_1^g, \pi_1^g)', \quad (14)$$

$$\theta_0^{(2)} = (R_0^z, x_0^z, \pi_0^z, R_1^z, x_1^z, \pi_1^z)'. \quad (15)$$

The first database contains responses to monetary policy and demand shocks, whereas the

second contains responses to a productivity shock. The six targets within each database are jointly estimated, but the covariance between the two estimators is unavailable.

Among the structural parameters, only

$$\alpha = (\psi_1, \psi_2, \rho_m, \sigma_m)'$$

is estimated by impulse-response matching; the remaining structural parameters are fixed at their benchmark values. We use identity weighting matrix. The Jacobian of the twelve model-implied responses has rank $r_\alpha = 3$. Thus, at most three independent combinations of the four parameters can be locally identified, leaving one degree of local underidentification. The specification residual has dimension $m - r_\alpha = 12 - 3 = 9$.

The within-database covariance matrices are generated from the same impulse-response estimators used to construct the targets.

We compare four tests with nominal level 5%, all based on the same MD statistic. The full-information oracle observes Σ_0 and uses the corresponding generalized chi-square critical value. The proposed test observes only Σ_1 and Σ_2 and numerically maximizes the generalized chi-square quantile over the admissible residual correlations. The SB test uses the same rank-adjusted residual space but replaces the sharp quantile by the conservative Székely–Bakirov trace bound. Finally, the joint overidentification procedure of [Cocci and Plagborg-Møller \(2025\)](#) assumes point identification to compute the projection matrix of the weighted jacobian space, and uses the Székely–Bakirov bound instead of solving the whole worst case critical value. Because the population Jacobian has rank three rather than four, the Cocci and Plagborg overidentification test is included only as a benchmark.

This comparison separates two sources of conservativeness. The difference between the oracle and the sharp robust test measures the cost of not knowing the cross-database covariance. The additional difference between the sharp and SB-bound tests measures the cost of replacing the sharp worst-case quantile with a general upper bound.

Table 1 reports finite-sample rejection frequencies under correct specification. The parameter ρ scales residual dependence along the constructed covariance path: $\rho = 0$ corresponds to a DGP with independent databases, whereas $\rho = 1$ is the DGP with the least-favorable cross-covariance structure. The table reports the cases $\rho \in \{0, 0.5, 0.8, 1\}$.

Table 1: Finite-sample rejection probabilities under the null

Correlation scale ρ	N	Rejection probability (%)			
		Oracle	Ours	SB bound	CPM
0	100	2.9	1.2	0.5	0.2
0	500	5.1	2.9	0.8	0.4
0	1000	4.6	2.2	0.6	0.5
0.5	100	3.8	3.3	1.6	0.9
0.5	500	4.8	2.9	0.8	0.3
0.5	1000	4.2	2.9	1.3	0.6
0.8	100	5.1	4.2	1.5	1.0
0.8	500	4.6	4.3	1.9	0.7
0.8	1000	4.8	4.2	1.9	1.1
1	100	4.8	4.8	2.3	1.1
1	500	4.7	4.7	2.5	0.9
1	1000	4.9	4.9	2	1.5

Notes: Entries are rejection frequencies in percentage points, rounded to one decimal. The oracle observes the complete covariance matrix. “Ours” denotes the test based on the residual-space reparameterization. SB denotes the test based on the Székely–Bakirov upper bound. CPM denotes the joint overidentification test of [Cocci and Plagborg-Møller \(2025\)](#). The nominal significance level is 5%.

The oracle is generally closest to the 5% nominal level. Among the procedures that do not observe the cross-database covariance, our test is the least conservative. Its rejection frequency is as low as 1.2% at $N = 100$ under independence, but generally increases as ρ approaches the least-favorable completion. This is the expected cost of robust size control: the critical value must protect against every admissible covariance completion and is therefore relatively large under more benign dependence structures. The CPM benchmark is the most conservative. Since its maintained point-identification condition fails in this design, it is included only as an out-of-assumption benchmark.

Figure 1 reports the asymptotic rejection probability across different correlation scenarios, where $\rho = 0$ corresponds to independence across databases and $\rho = 1$ to the least-favorable cross-covariance structure.

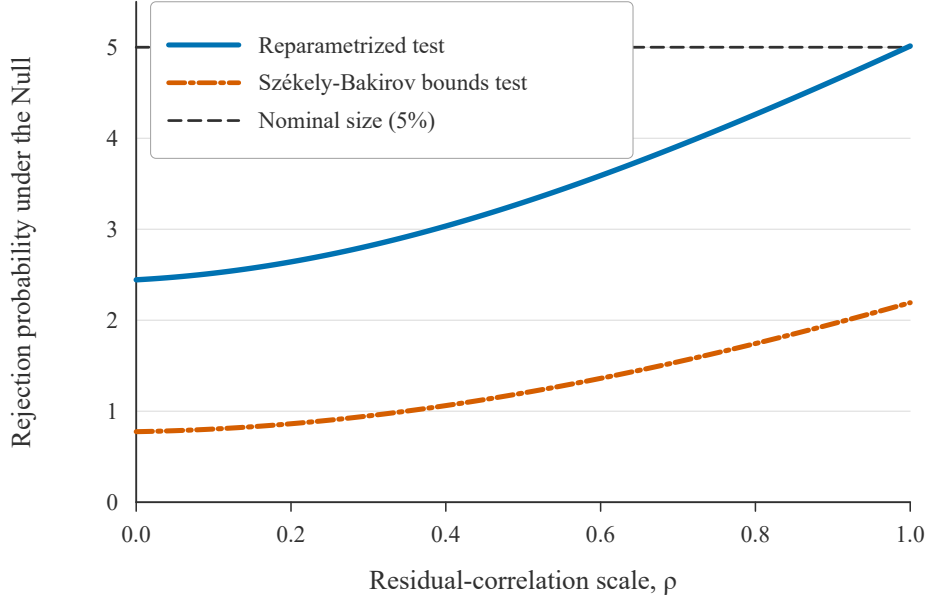


Figure 1: Asymptotic size along the residual-correlation path.

The size of our test rises toward the nominal level as the true covariance approaches the least-favorable covariance. The SB procedure remains well below 5%, illustrating the additional size loss caused by replacing the sharp worst-case quantile with an upper bound. The oracle is omitted because its asymptotic rejection probability equals the nominal level by construction.

7.1 Power under alternative forms of misspecification

We next examine finite-sample power against three forms of misspecification induced by departures of the fixed parameters from their calibrated values. Let

$$\theta_0 = g(\alpha_0, \bar{\beta}) \in \Theta_0(\bar{\beta}).$$

For each alternative value $\beta \neq \bar{\beta}$, the DGP is characterized by

$$\theta(\beta) = g(\alpha_0, \beta),$$

while the model used for testing continues to fix the calibrated parameters at $\bar{\beta}$. We evaluate the rejection probability for alternatives satisfying

$$\theta(\beta) \notin \Theta_0(\bar{\beta}).$$

We consider the case where the true Σ_0 is associated with $\rho = 0.8$ and study three forms of

misspecification. The first varies the government-spending persistence, ρ_g ; the second varies the Phillips-curve slope, κ ; and the third varies the productivity persistence parameter, ρ_z . These departures are chosen along directions that generate misspecification according to the condition discussed in Subsection 4.1.

Figure 2 reports rejection probabilities for $N = 100, 500$, and 1000 as the magnitude of each departure increases. For ease of interpretation, we show the power functions over the space of calibrated parameters rather than in terms of the targets.

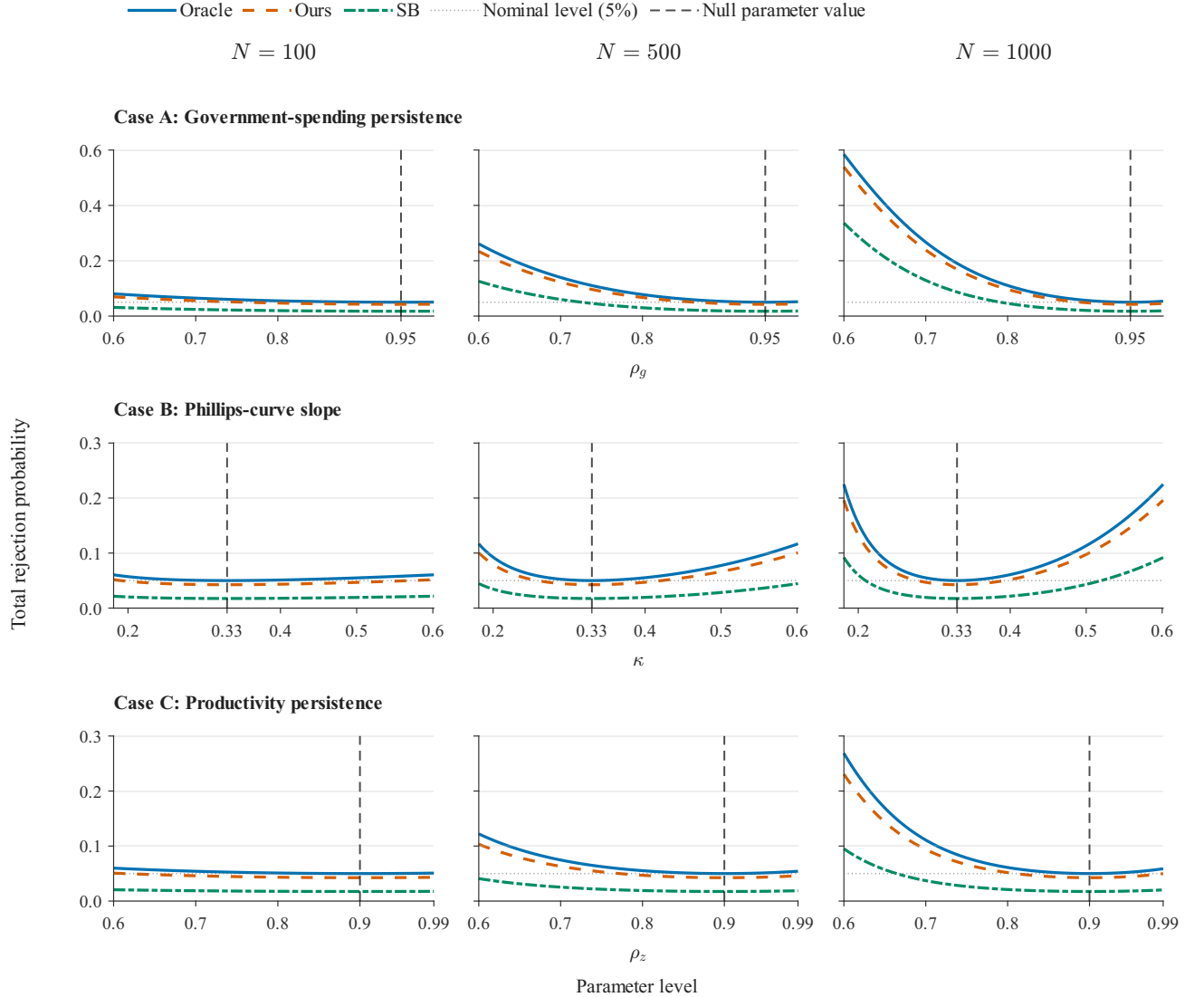


Figure 2: Power under alternative forms of misspecification.

Two patterns are clear. First, power increases with both the magnitude of misspecification and sample size: departures that are difficult to distinguish from sampling variation at $N = 100$ become increasingly visible as the sample grows. Second, our test generally remains substantially closer to the full-information Oracle than the SB procedure, showing that using the sharp worst-

case critical value can recover an important part of the power lost when the cross-database covariance is unknown. The differences across the four alternatives are also informative.

8 Application: Monetary Policy and Information effects

[Nakamura and Steinsson \(2018\)](#) provide evidence of a monetary policy information effect: Fed announcements change private-sector beliefs not only about future monetary policy but also about the economic outlook. They identify policy surprises using interest-rate movements within a 30-minute window around scheduled Fed announcements and examine how growth forecasts respond. They find that unexpected interest-rate increases are associated with higher expected growth rather than lower growth, suggesting that these announcements signal a stronger economy. They then estimate a New Keynesian model by matching model-implied impulse responses to those estimated in the data. To explain their findings, they incorporate this information effect into the model alongside the conventional contractionary effect of higher interest rates.

8.1 Description of the Two Datasets

[Nakamura and Steinsson \(2018\)](#) use two datasets to estimate the responses that are later targeted in their IRFM estimation.

The first combines Treasury and TIPS data to estimate sixteen responses: real yields, real instantaneous forwards, and the corresponding break-even inflation measures at maturities of two, three, five, and ten years. Responses are OLS slopes of daily price changes on policy news measured in 30-minute FOMC windows. Depending on the maturity, responses are estimated using either 74 announcements over 2004–2014 or 106 announcements over 2000–2014. Therefore, from this database we estimate a total of 16 responses used as targets for the IRFM estimation.

The second uses Blue Chip Economic Indicators to estimate eight responses of revisions to real-GDP-growth expectations, from the current quarter through seven quarters ahead. The sample spans 1995–2014, with the number of observations declining from 120 at short horizons to 13 at the longest horizon. Hence, from this database we obtain 8 targets. The IRFM estimation therefore uses 24 targets in total.

8.2 New Keynesian Model with Information Effects

Following [Nakamura and Steinsson \(2018\)](#), we consider a New Keynesian model with internal consumption habits and sticky prices with indexation. The model is designed to capture both the conventional transmission of monetary policy and the information contained in Federal

Reserve announcements. In log-linearized form, private-sector behavior and monetary policy are characterized by

$$\lambda_{x,t} = E_t \lambda_{x,t+1} + i_t - E_t \pi_{t+1} - r_t^n, \quad (16)$$

$$\lambda_{x,t} = -\sigma_c [(1 + \rho b^2)x_t - b x_{t-1} - \rho b E_t x_{t+1}], \quad (17)$$

$$\Delta \pi_t = \rho E_t \Delta \pi_{t+1} + \kappa \zeta (\omega x_t - \lambda_{x,t}), \quad (18)$$

$$i_t - E_t \pi_{t+1} = \bar{r}_t + \phi_\pi \pi_t, \quad (19)$$

$$\bar{r}_t = (\rho_1 + \rho_2) \bar{r}_{t-1} - \rho_1 \rho_2 \bar{r}_{t-2} + \varepsilon_t, \quad (20)$$

where

$$x_t = y_t - y_t^n, \quad \lambda_{x,t} = \lambda_t - \lambda_t^n, \quad \Delta \pi_t = \pi_t - \pi_{t-1},$$

and

$$\sigma_c = [\sigma(1 - b)(1 - \rho b)]^{-1}.$$

The superscript n denotes the corresponding flexible-price, or natural, allocation. In particular, r_t^n denotes the natural real interest rate. The Euler equation therefore implies that the conventional monetary-policy channel operates through the real interest-rate gap, $i_t - E_t \pi_{t+1} - r_t^n$.

Nakamura and Steinsson augment this conventional mechanism by allowing monetary announcements to change private-sector beliefs about economic fundamentals. In particular, they assume that, in response to a monetary announcement,

$$E_t r_{t+j}^n = \gamma E_t \bar{r}_{t+j}, \quad j \geq 0, \quad (21)$$

where γ measures the strength of the information effect. Thus, a fraction γ of the announcement-induced change in the expected policy component is interpreted by private agents as news about the natural rate of interest, while the remaining fraction $1 - \gamma$ represents a conventional change in the interest-rate gap. When $\gamma = 0$, monetary announcements convey no information about the natural rate.

8.3 Impulse Response Function Matching and Specification Test

The New Keynesian model incorporates sticky prices, internal consumption habits, and a backward-looking term in the Phillips curve. Federal Reserve announcements affect both the expected path of interest rates and private-sector beliefs about the natural rate of interest, introducing an information effect alongside the conventional monetary policy channel. In their original estimation, they fix the values of eight structural parameters, described in Table 2.

The restricted model can be misspecified for two reasons. First, the unrestricted model may be unable to match all targets jointly for any value of the structural parameters. Second,

Table 2: Fixed parameters

Symbol	Definition	Value
ρ	Subjective discount factor	0.99
δ	Nominal rigidity	0.75
ω	Elasticity of marginal cost to output	2
θ	Elasticity of substitution across varieties	10
σ	Intertemporal elasticity of substitution	0.50
b	Consumption habit	0.90
ϕ_π	Endogenous inflation feedback in the monetary-policy rule	0.01
ι_π^a	Price indexation to lagged inflation	1

the unrestricted model may be correctly specified, but the restrictions imposed in Table 2 prevent any value of the remaining free parameters from matching all targets jointly. Our test assesses the specification of the restricted model without distinguishing between these two sources of misspecification. It does so without requiring point identification or knowledge of which additional parameters would need to be fixed to ensure local identification of the remaining parameters.

The estimated parameters are the two autoregressive roots of the policy process, ρ_1 and ρ_2 , the slope of the Phillips curve, $\kappa\zeta$, and the information-share parameter ψ . We estimate the model twice: once using standard MD and once using MD with a ridge penalty. The standard MD estimates are used only to compute the test statistic, while the ridge-MD estimates are used to evaluate the Jacobians required to compute the worst-case critical values of the test.

Table 3: Structural parameter estimates

Symbol	Definition	Unpenalized MD	Regular (ridge)
ρ_1	First autoregressive root of monetary shock	0.874	0.873
ρ_2	Second autoregressive root of monetary shock	0.870	0.867
$\kappa\zeta \times 10^5$	Phillips curve slope	14.68	14.43
γ	Information effect	0.75	0.74

We construct the test following the steps described in the algorithm in Appendix A.5. In particular, the weighting matrix is the same as that used in the original paper by Nakamura and Steinsson (2018). To estimate the ranks by hard thresholding, we use the simple threshold $\tau_G = \sqrt{\log(120)/120} = 0.19974$ for the Jacobian and $\tau_\Omega = \log(120)/120$ for the covariance matrices of the different block residuals. See Appendix A.3 for a full discussion of this choice. The rank estimation process yields two main results: the estimated rank of the Jacobians is

three rather than four, and the dimension of the worst-case critical-value optimization problem is reduced from 128 unknown cross-covariances to 56 unknown residual cross-correlations.

The following table summarizes the results of the correct specification test for the restricted model.

Table 4: Specification-test results at the 5% level

	Value
Test statistic	75.53
Worst-Case Critical Value	111.66
Székel–Bakirov upper bound	159.91

Thus, the null hypothesis that the restricted model is correctly specified is not rejected using the worst-case critical value computed via the residual-space reparameterization. The next subsection examines the robustness of this conclusion.

8.4 Robustness of the Testing Conclusions

To examine the robustness of the previous conclusion, we construct a linear path between the independence residual correlation matrix and the residual correlation matrix associated with the worst-case critical value. We then plot the critical values along this path.

Consider the reparametrized residual-correlation path

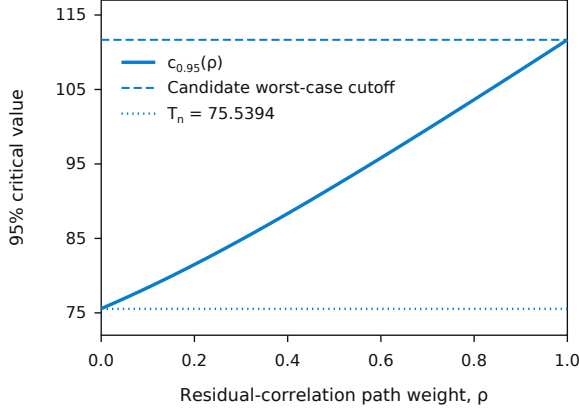
$$\hat{R}(\rho) = \begin{pmatrix} I_{\hat{s}_1} & \rho \hat{R}_{\hat{s}_1 \hat{s}_2} \\ \rho \hat{R}_{\hat{s}_1 \hat{s}_2}^\top & I_{\hat{s}_2} \end{pmatrix}, \quad \rho \in [0, 1],$$

where $\hat{s}_1 = 8$, $\hat{s}_2 = 7$ are the estimated rank of the variance/covariance matrix of the residuals of each block, and $\hat{R}_{\hat{s}_1 \hat{s}_2}$ is the worst case cross-correlation matrix obtained from the reparametrized critical-value search.

In this representation, $\rho = 0$ corresponds to independence between the residual components associated with the two data blocks, while $\rho = 1$ corresponds to the worst-case dependence structure. Panel A of Figure 3 plots the covariance-specific 95% quantile along this path. It rises from 75.5891 to 111.66. There is no rejection anywhere on the displayed path, although under independence the critical value exceeds the statistic by only 0.0496. This suggests that the non-rejection of the null hypothesis is robust, at least along the linear path between independence and the worst-case scenario.

Panel B holds the critical value at the worst case $c^* = 111.66$ and reports the loss of size due to using the worst case value as described in (7). Under independence, the loss of size is 81.6%, so that the size of the test using the worst-case critical value when the DGP has residual

A. Covariance-specific critical value



B. Size loss from using the worst-case critical value

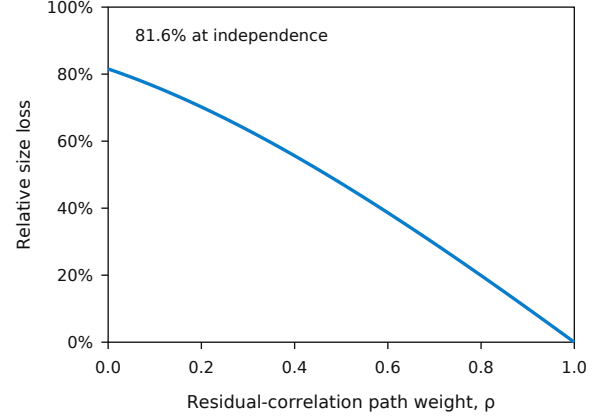


Figure 3: Dependence sensitivity in the reparametrized residual-correlation specification.

independent databases is 0.95%; it reaches 5% at the candidate endpoint by construction. These are asymptotic diagnostics, not finite-sample rejection frequencies.

9 Conclusion

This paper develops a minimum distance correct specification test motivated primarily by Macroeconomic applications. In these settings, empirical targets are often estimated from multiple databases with unavailable cross-database covariances, some structural parameters may not be point identified, and the asymptotic covariance matrix of the target estimators may be singular. These features arise naturally in applications such as impulse-response matching, but jointly make standard specification tests difficult to implement.

We obtain valid inference by using the sharp worst-case critical value over all covariance completions consistent with the estimable within-database blocks. An exact residual-space reparameterization preserves this critical value while reducing the dimension of the optimization problem. The resulting test controls asymptotic size, is consistent against fixed alternatives, and has non-trivial local power outside the model’s tangent space. The New Keynesian illustration shows the practical relevance of the procedure and the potential reduction in conservativeness relative to methods based on general upper bounds.

A Further discussions

A.1 Residual-space reparameterization

This subsection explains the two exact reductions underlying Theorem 5.1. Write

$$d \equiv m - r_\alpha.$$

First, because $R_W = H'H$, the limiting quadratic form can be written as the squared norm of the aggregate residual

$$Y = HZ = \sum_{b=1}^B H_b Z_b.$$

The original covariance-completion problem can therefore be expressed entirely in terms of the covariance matrix of the database-specific residual contributions $Y_b = H_b Z_b$. Lemma A.1 shows that this first reduction is exact.

Second, factor each residual covariance as $\Omega_b = L_b L_b'$ and write $Y_b = L_b U_b$. The standardized contribution U_b has covariance I_{s_b} , so all remaining unknown dependence across databases is summarized by the block-correlation matrix of the stacked vector of standardized contributions. Lemma A.2 shows that this second reduction is also exact.

Blocks with $s_b = 0$ are omitted from the standardized vector. If $s = \sum_{b=1}^B s_b = 0$, all residual contributions are identically zero and both critical-value representations equal zero. We therefore consider $s > 0$ below.

For any $\tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B)$, let

$$Z = (Z'_1, \dots, Z'_B)' \sim N(0, \tilde{\Sigma}), \quad Y_b = H_b Z_b, \quad Y = \sum_{b=1}^B Y_b.$$

Since $R_W = H'H$,

$$Z' R_W Z = Y' Y, \quad \text{Var}(Y) = H \tilde{\Sigma} H'.$$

Consequently,

$$q_{1-\tau} \left(R_W^{1/2} \tilde{\Sigma} R_W^{1/2} \right) = q_{1-\tau} \left(H \tilde{\Sigma} H' \right). \quad (22)$$

The covariance matrix of the stacked residual contributions has diagonal blocks

$$\Omega_b = H_b \Sigma_b H_b'.$$

Define

$$\mathcal{G}(\Omega_1, \dots, \Omega_B) \equiv \left\{ \Gamma \in \mathbb{S}_+^{dB} : \Gamma_{bb} = \Omega_b, \quad b = 1, \dots, B \right\},$$

where Γ is partitioned into $d \times d$ blocks, and let

$$S_B \equiv (I_d, \dots, I_d) \in \mathbb{R}^{d \times dB}.$$

For any $\Gamma \in \mathcal{G}(\Omega_1, \dots, \Omega_B)$, the covariance matrix of the aggregate residual is $S_B \Gamma S_B'$.

Lemma A.1 (Exact residual-space completion) *The set of admissible covariance matrices for the aggregate residual satisfies*

$$\left\{ H \tilde{\Sigma} H' : \tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B) \right\} = \{ S_B \Gamma S_B' : \Gamma \in \mathcal{G}(\Omega_1, \dots, \Omega_B) \}. \quad (23)$$

Proof. See Appendix B. ■

For each database, let

$$\Omega_b = L_b L_b', \quad L_b \in \mathbb{R}^{d \times s_b}, \quad s_b = \text{rank}(\Omega_b),$$

be a thin square root. Define

$$\mathcal{L} \equiv \text{diag}(L_1, \dots, L_B) \in \mathbb{R}^{dB \times s}, \quad L \equiv (L_1, \dots, L_B) \in \mathbb{R}^{d \times s}.$$

Thus, $\mathcal{L}$ maps standardized contributions into the stacked residual vector, whereas L maps them into the aggregate residual. In particular,

$$S_B \mathcal{L} = L.$$

Recall the block-correlation set

$$\mathcal{C} \equiv \{ C \in \mathbb{S}_+^s : C_{bb} = I_{s_b} \text{ for every } b \text{ with } s_b > 0 \}.$$

Lemma A.2 (Exact block-correlation representation) *The residual covariance-completion set admits the exact representation*

$$\mathcal{G}(\Omega_1, \dots, \Omega_B) = \{ \mathcal{L} C \mathcal{L}' : C \in \mathcal{C} \}. \quad (24)$$

Proof. See Appendix B. ■

Combining Lemmas A.1 and A.2 gives

$$\left\{ H \tilde{\Sigma} H' : \tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B) \right\} = \{ L C L' : C \in \mathcal{C} \}.$$

Together with (22), this reduces the original worst-case critical-value problem to

$$\sup_{C \in \mathcal{C}} q_{1-\tau}(LCL').$$

The proof of Theorem 5.1 shows that $\mathcal{C}$ is compact and that the objective is continuous, so the supremum is attained.

The dimension reduction follows from

$$s_b = \text{rank}(H_b \Sigma_b H_b') = \text{rank}(H_b \Sigma_b^{1/2}) \leq \min \{d, \text{rank}(\Sigma_b)\}.$$

Thus, an original cross-database covariance block of dimension $m_b \times m_c$ is replaced by a block of correlations of dimension $s_b \times s_c$. The reparameterization removes both variation along the model's tangent space and any additional rank deficiency in the residual contribution of each database.

A.2 Solution of Minimum Distance

Define the sample minimum distance criterion

$$Q_n(\alpha; \bar{\beta}) \equiv \{\widehat{\theta} - g(\alpha, \bar{\beta})\}' \widehat{W} \{\widehat{\theta} - g(\alpha, \bar{\beta})\},$$

and let

$$\widehat{q}_n(\bar{\beta}) \equiv \inf_{\alpha \in \mathcal{A}} Q_n(\alpha; \bar{\beta}).$$

The test statistic is

$$T_n = n \widehat{q}_n(\bar{\beta}).$$

The minimum distance criterion may have more than one minimizer when α is partially identified. The value $\widehat{q}_n(\bar{\beta})$, and hence the test statistic, remains well defined. Nevertheless, the subsequent estimation of the tangent space requires selecting a regular representative at which the Jacobian can be evaluated.

We select this representative using a vanishing ridge penalty. In particular, define

$$\widehat{\alpha} \equiv \widehat{\alpha}_{\lambda_n} \in \arg \min_{\alpha \in \mathcal{A}} \{Q_n(\alpha; \bar{\beta}) + \lambda_n \|\alpha\|^2\},$$

where $\lambda_n > 0$ satisfies

$$\lambda_n \longrightarrow 0, \quad n \lambda_n \longrightarrow \infty, \quad \sqrt{n} \lambda_n \longrightarrow 0.$$

Equivalently, one may set

$$\lambda_n = n^{-\kappa}, \quad \frac{1}{2} < \kappa < 1,$$

with $\lambda_n = n^{-3/4}$ providing a simple admissible choice.

The ridge estimator is used only to determine the point at which the Jacobian is evaluated. It does not enter the definition of $\hat{q}_n(\bar{\beta})$ or T_n . Thus, the unpenalized and penalized minimum distance problems are solved separately.

To state the required regularity conditions, define the minimum-norm member of the population identified set by

$$\alpha^\dagger \equiv \arg \min_{\alpha \in \mathcal{A}_0(\theta_0, \bar{\beta})} \|\alpha\|^2.$$

We assume that $\mathcal{A}$ is closed, that $g(\cdot, \bar{\beta})$ is continuous on $\mathcal{A}$, and that the penalized criterion is solved globally. The quadratic penalty then makes the objective coercive and guarantees the existence of a penalized solution. We further assume that $\alpha^\dagger$ is unique, belongs to $\text{int}(\mathcal{A})$, and is a regular representative of the restricted model.

In addition, the minimum-norm representation is assumed to be locally stable. More precisely, for μ in a relative neighborhood of θ_0 in $\Theta_0(\bar{\beta})$, define

$$a^\dagger(\mu) \equiv \arg \min_{\alpha \in \mathcal{A}} \{ \|\alpha\|^2 : g(\alpha, \bar{\beta}) = \mu \}.$$

We assume that this solution is locally unique and continuously differentiable at θ_0 , with

$$a^\dagger(\theta_0) = \alpha^\dagger.$$

This is a local second-order stability condition for the minimum-norm selection.

Under these conditions and the reduced-form asymptotics imposed in the main text,

$$\hat{\alpha} - \alpha^\dagger = O_p(n^{-1/2}).$$

Consequently, if $g(\cdot, \bar{\beta})$ is twice continuously differentiable in a neighborhood of $\alpha^\dagger$, then

$$\hat{G} \equiv \nabla_\alpha g(\hat{\alpha}, \bar{\beta})$$

satisfies

$$\|\hat{G} - G^\dagger\|_{\text{op}} = O_p(n^{-1/2}), \quad G^\dagger \equiv \nabla_\alpha g(\alpha^\dagger, \bar{\beta}).$$

The regularity requirement is therefore

$$\text{rank}(G^\dagger) = r_\alpha, \quad \text{col}(G^\dagger) = T_{\theta_0} \Theta_0(\bar{\beta}).$$

This construction does not require point identification of α . The identified set may contain several structural representatives, but its unique minimum-norm member must provide a regular local parameterization of the restricted-model set.

A.3 Rank estimation by hard thresholding

A.3.1 Generic hard-thresholding rule

The feasible procedure requires estimating the rank of several population matrices. This subsection first gives a generic hard-thresholding result and then specializes it to the weighted Jacobian and the residual covariance factors.

Let

$$A \in \mathbb{R}^{d_1 \times d_2}, \quad \hat{A} \in \mathbb{R}^{d_1 \times d_2}, \quad k_A \equiv \min\{d_1, d_2\}.$$

Let

$$\hat{\sigma}_{1,n} \geq \cdots \geq \hat{\sigma}_{k_A,n} \geq 0$$

denote the singular values of $\hat{A}$. Given a possibly data-dependent threshold $\tau_{A,n} > 0$, define

$$\hat{r}_A \equiv \sum_{j=1}^{k_A} \mathbf{1}\{\hat{\sigma}_{j,n} > \tau_{A,n}\}.$$

Lemma A.3 (Consistency of hard-thresholded rank estimation) *Let*

$$r_A \equiv \text{rank}(A).$$

Suppose that

$$\tau_{A,n} \xrightarrow{p} 0$$

and

$$\frac{\|\hat{A} - A\|_{\text{op}}}{\tau_{A,n}} \xrightarrow{p} 0.$$

Then

$$\Pr(\hat{r}_A = r_A) \longrightarrow 1.$$

Proof. See Appendix B. ■

If

$$\|\hat{A} - A\|_{\text{op}} = O_p(n^{-1/2}),$$

it is sufficient that

$$\tau_{A,n} \xrightarrow{p} 0, \quad \sqrt{n} \tau_{A,n} \xrightarrow{p} \infty.$$

A convenient specification is

$$\tau_{A,n} = \kappa_A a_{A,n} \sqrt{\frac{\log n}{n}}, \quad \kappa_A > 0,$$

where

$$a_{A,n} \xrightarrow{p} a_A \in (0, \infty).$$

The normalization $a_{A,n}$ places the threshold in the same numerical units as the singular values, while κ_A controls the amount of finite-sample thresholding. After placing A on a fixed numerical scale, one may set $a_{A,n} = 1$.

The conclusion of Lemma A.3 is unchanged under orthogonal changes of coordinates. In particular, it is sufficient that there exist orthogonal matrices $Q_{1,n}$ and $Q_{2,n}$ such that

$$\frac{\|\hat{A} - Q_{1,n} A Q_{2,n}'\|_{\text{op}}}{\tau_{A,n}} \xrightarrow{p} 0.$$

A.3.2 Estimation of the tangent rank

Let $D_\alpha \succ 0$ be a fixed diagonal matrix used to place the structural parameters on comparable numerical scales. The choice $D_\alpha = I_q$ corresponds to using the original parameterization. Define the normalized weighted Jacobian

$$\hat{X} \equiv \widehat{W}^{1/2} \widehat{G} D_\alpha.$$

Because D_α is nonsingular,

$$\text{rank}(W^{1/2} G D_\alpha) = \text{rank}(G)$$

and

$$\text{col}(W^{1/2} G D_\alpha) = \text{col}(W^{1/2} G).$$

Thus, the normalization does not change the population tangent space. Its purpose is only to reduce the dependence of the finite-sample rank decision on the units in which the parameters are reported. This follows the scale-invariance principle used in identification-category selection procedures such as [Forneron \(2024\)](#).

Let

$$\hat{\sigma}_{G,1} \geq \cdots \geq \hat{\sigma}_{G,\min\{m,q\}} \geq 0$$

denote the singular values of $\hat{X}$. Define

$$\tau_{G,n} = \kappa_G a_{G,n} \sqrt{\frac{\log n}{n}}, \quad \kappa_G > 0,$$

where

$$a_{G,n} \xrightarrow{p} a_G \in (0, \infty),$$

and estimate

$$\hat{r}_\alpha = \sum_{j=1}^{\min\{m,q\}} \mathbf{1} \{ \hat{\sigma}_{G,j} > \tau_{G,n} \}.$$

Suppose that, for some regular representative $G^\star$,

$$X^\star \equiv W^{1/2} G^\star D_\alpha$$

satisfies

$$\text{rank}(X^\star) = r_\alpha, \quad \text{col}(X^\star) = \text{col}(W^{1/2} G_0),$$

and

$$\|\hat{X} - X^\star\|_{\text{op}} = O_p(n^{-1/2}).$$

Then Lemma A.3 implies

$$\Pr(\hat{r}_\alpha = r_\alpha) \longrightarrow 1.$$

If $\hat{G}$ is computed numerically, its numerical differentiation error is required to be $o_p(\tau_{G,n})$.

Let

$$\hat{X} = \hat{U} \hat{D}_G \hat{V}'$$

be a full singular-value decomposition. Let $\hat{U}_R$ collect the left singular vectors associated with the singular values not retained in the estimated tangent space, and define

$$\hat{H} \equiv \hat{U}_R' \hat{W}^{1/2} = (\hat{H}_1, \dots, \hat{H}_B).$$

A.3.3 Estimation of the residual-block ranks

For database b , let

$$S_b \equiv \Sigma_b^{1/2}, \quad M_b \equiv H_b S_b.$$

Then

$$\Omega_b = H_b \Sigma_b H_b' = M_b M_b',$$

and therefore

$$s_b \equiv \text{rank}(\Omega_b) = \text{rank}(M_b).$$

Let

$$\hat{S}_b \equiv \hat{\Sigma}_b^{1/2}$$

be the symmetric positive-semidefinite square root of $\hat{\Sigma}_b$. Before computing this square root,

$\widehat{\Sigma}_b$ is numerically symmetrized and negative eigenvalues attributable to rounding error are set equal to zero. Define

$$\widehat{M}_b \equiv \widehat{H}_b \widehat{S}_b.$$

Let

$$\widehat{M}_b = \widehat{U}_b \widehat{D}_b \widehat{V}_b', \quad \widehat{D}_b = \text{diag} \left(\widehat{d}_{b,1}, \dots, \widehat{d}_{b,k_b} \right),$$

where

$$k_b = \min\{m - \widehat{r}_\alpha, m_b\}, \quad \widehat{d}_{b,1} \geq \dots \geq \widehat{d}_{b,k_b} \geq 0.$$

Define

$$\tau_{M,b,n} = \kappa_{M,b} a_{M,b,n} \sqrt{\frac{\log n}{n}}, \quad \kappa_{M,b} > 0,$$

where

$$a_{M,b,n} \xrightarrow{p} a_{M,b} \in (0, \infty).$$

Estimate the residual-block rank by

$$\widehat{s}_b = \sum_{j=1}^{k_b} \mathbf{1} \left\{ \widehat{d}_{b,j} > \tau_{M,b,n} \right\}.$$

Let $\widehat{U}_{b,+}$ and $\widehat{D}_{b,+}$ collect the retained left singular vectors and singular values. The corresponding thin residual factor is

$$\widehat{L}_b \equiv \widehat{U}_{b,+} \widehat{D}_{b,+}.$$

It satisfies

$$\widehat{L}_b \widehat{L}_b' = \sum_{j: \widehat{d}_{b,j} > \tau_{M,b,n}} \widehat{d}_{b,j}^2 \widehat{u}_{b,j} \widehat{u}_{b,j}'.$$

If $\widehat{s}_b = 0$, then $\widehat{L}_b$ is the $(m - \widehat{r}_\alpha) \times 0$ empty matrix.

Since

$$\widehat{\Omega}_b = \widehat{M}_b \widehat{M}_b', \quad \lambda_j(\widehat{\Omega}_b) = \widehat{d}_{b,j}^2,$$

the same rank decision can equivalently be implemented as

$$\lambda_j(\widehat{\Omega}_b) > \tau_{M,b,n}^2.$$

In particular,

$$\tau_{M,b,n}^2 = \kappa_{M,b}^2 a_{M,b,n}^2 \frac{\log n}{n}.$$

Thus, a factor singular-value threshold of order $\sqrt{\log n/n}$ corresponds to an eigenvalue threshold for $\widehat{\Omega}_b$ of order $\log n/n$.

For consistency, suppose that, on the event $\widehat{r}_\alpha = r_\alpha$, there exists an orthogonal matrix $Q_n \in \mathbb{R}^{(m-r_\alpha) \times (m-r_\alpha)}$ such that

$$\left\| \widehat{M}_b - Q_n M_b \right\|_{\text{op}} = O_p(n^{-1/2}).$$

Then Lemma A.3 implies

$$\Pr(\widehat{s}_b = s_b) \longrightarrow 1.$$

This factor-level perturbation condition must be verified for the chosen estimator of $\Sigma_b^{1/2}$; it does not follow solely from a root- n convergence result for $\widehat{\Omega}_b$.

If the preceding conditions hold for the weighted Jacobian and for every residual factor M_b , then

$$\Pr(\widehat{r}_\alpha = r_\alpha, \widehat{s}_b = s_b \text{ for every } b) \longrightarrow 1.$$

A.4 Validity of the feasible test

This subsection establishes the validity of the feasible test in Section 6, using the construction in Appendix A.5. Regular-solution selection and rank estimation are discussed in Appendices A.2 and A.3, respectively.

The result concerns the covariance-eigenvalue thresholding procedure used in the algorithm. We use

$$\tau_{G,n} = \kappa_G a_{G,n} \sqrt{\frac{\log n}{n}}, \quad \tau_{\Omega,b,n} = \kappa_{\Omega,b} a_{\Omega,b,n} \frac{\log n}{n},$$

where the tuning constants κ_G and $\kappa_{\Omega,b}$ are fixed and strictly positive, and the scale normalizations satisfy

$$a_{G,n} \xrightarrow{p} a_G \in (0, \infty), \quad a_{\Omega,b,n} \xrightarrow{p} a_{\Omega,b} \in (0, \infty), \quad b = 1, \dots, B.$$

This is the eigenvalue implementation of the factor-singular-value rule in Appendix A.3: $\tau_{\Omega,b,n} = \tau_{M,b,n}^2$, with $\kappa_{\Omega,b} = \kappa_{M,b}^2$ and $a_{\Omega,b,n} = a_{M,b,n}^2$. Its validity therefore uses the factor-level perturbation condition stated below.

Throughout this subsection, $\widehat{c}_{1-\tau}$ denotes the exact plug-in critical value defined by the maximization in Section 6.

Proposition A.4 (Consistency and validity of the feasible test) *Suppose Assumptions 1–3 hold under H_0 , and $0 < \tau < 1$. Assume that*

$$\|\widehat{\Sigma}_b - \Sigma_b\|_{\text{op}} = O_p(n^{-1/2}), \quad b = 1, \dots, B,$$

and

$$\|\widehat{G} - G^*\|_{\text{op}} = O_p(n^{-1/2}), \quad \|\widehat{W} - W\|_{\text{op}} = O_p(n^{-1/2}),$$

where G^* is a fixed regular representative satisfying

$$\text{rank}(G^*) = r_\alpha, \quad \text{col}(G^*) = \text{col}(G_0).$$

When $m - r_\alpha > 0$, define

$$M_b = H_b \Sigma_b^{1/2}, \quad \widehat{M}_b = \widehat{H}_b \widehat{\Sigma}_b^{1/2}.$$

Assume additionally that, on the event $\widehat{r}_\alpha = r_\alpha$, there exist orthogonal matrices

$$Q_n \in \mathbb{R}^{(m-r_\alpha) \times (m-r_\alpha)},$$

common to all database blocks, such that

$$\|\widehat{M}_b - Q_n M_b\|_{\text{op}} = O_p(n^{-1/2}), \quad b = 1, \dots, B.$$

Then

$$\widehat{c}_{1-\tau} \xrightarrow{p} c_{1-\tau}^*(R_W), \quad \limsup_{n \rightarrow \infty} \Pr\{\widehat{\phi}_n(\bar{\beta}) = 1\} \leq \tau.$$

Proof. See Appendix B. ■

A.5 Algorithm

Algorithm A.1: Construction of the feasible worst-case critical value

Inputs. Calibrated parameter $\bar{\beta}$; estimated parameter $\widehat{\alpha}$ following Section A.2; weighting matrix $\widehat{W} \in \mathbb{S}_{++}^m$; the within-database covariance estimators $\widehat{\Sigma}_b \in \mathbb{S}_+^{m_b}$, $b = 1, \dots, B$; the nominal level $\tau \in (0, 1)$; rank thresholds $\tau_{G,n}$ and $\{\tau_{\Omega,b,n}\}_{b=1}^B$.

The residual-covariance eigenvalue threshold $\tau_{\Omega,b,n}$ corresponds to the squared factor singular-value threshold discussed in Appendix A.3.

Select $\widehat{\alpha}$ as described in Appendix A.2. This representative is used only for Jacobian evaluation; compute T_n separately from the unpenalized minimum distance criterion in (4).

Each $\widehat{\Sigma}_b$ must be expressed under the common $\sqrt{n}$ normalization used in the test statistic. Thus, if the covariance estimator from database b is naturally expressed under a $\sqrt{n_b}$ normalization, use

$$\widehat{\Sigma}_b = \frac{n}{n_b} \widehat{\Sigma}_b^{(n_b)}.$$

A. Construction of the estimated residual factor $\widehat{H}$.

1. Evaluate the Jacobian at the selected regular representative:

$$\widehat{G} = \nabla_{\alpha} g(\widehat{\alpha}, \bar{\beta}),$$

and compute $\widehat{W}^{1/2}\widehat{G}$.

2. Compute a full singular-value decomposition

$$\widehat{W}^{1/2}\widehat{G} = \widehat{U}\widehat{D}\widehat{V}', \quad \widehat{U} = (\widehat{u}_1, \dots, \widehat{u}_m),$$

where the singular values satisfy

$$\widehat{\sigma}_1 \geq \dots \geq \widehat{\sigma}_{\min\{m, q\}} \geq 0.$$

Estimate the tangent rank by hard thresholding:

$$\widehat{r}_{\alpha} = \sum_{j=1}^{\min\{m, q\}} \mathbf{1}\{\widehat{\sigma}_j > \tau_{G, n}\}.$$

3. If $m - \widehat{r}_{\alpha} = 0$, set $\widehat{s}_b = 0$ for every b , $\widehat{s} = 0$, and $\widehat{c}_{1-\tau} = 0$, and proceed directly to the Output step.
4. Otherwise, collect the left singular vectors associated with the estimated residual space:

$$\widehat{U}_R = (\widehat{u}_{\widehat{r}_{\alpha}+1}, \dots, \widehat{u}_m).$$

Define

$$\widehat{H} = \widehat{U}_R' \widehat{W}^{1/2} \in \mathbb{R}^{(m-\widehat{r}_{\alpha}) \times m}.$$

Partition its columns conformably with the database blocks:

$$\widehat{H} = (\widehat{H}_1, \dots, \widehat{H}_B), \quad \widehat{H}_b \in \mathbb{R}^{(m-\widehat{r}_{\alpha}) \times m_b}.$$

B. Construction of $\widehat{\Omega}_b$, $\widehat{s}_b$, and $\widehat{L}_b$.

For each $b = 1, \dots, B$, perform the following steps.

1. Compute the estimated covariance matrix of the residual contribution from database b :

$$\widehat{\Omega}_b = \widehat{H}_b \widehat{\Sigma}_b \widehat{H}_b'.$$

Numerically symmetrize it:

$$\widehat{\Omega}_b \leftarrow \frac{\widehat{\Omega}_b + \widehat{\Omega}_b'}{2}.$$

2. Compute the spectral decomposition

$$\widehat{\Omega}_b = \widehat{E}_b \text{diag}(\widehat{\omega}_{b,1}, \dots, \widehat{\omega}_{b,m-\widehat{r}_\alpha}) \widehat{E}_b',$$

with the eigenvalues arranged in decreasing order.

Truncate negative eigenvalues at zero:

$$\widehat{\omega}_{b,j}^+ = \max\{\widehat{\omega}_{b,j}, 0\}, \quad j = 1, \dots, m - \widehat{r}_\alpha.$$

3. Estimate the rank of the residual covariance block by hard thresholding:

$$\widehat{s}_b = \sum_{j=1}^{m-\widehat{r}_\alpha} \mathbf{1}\{\widehat{\omega}_{b,j}^+ > \tau_{\Omega,b,n}\}.$$

4. If $\widehat{s}_b = 0$, define $\widehat{L}_b$ as the $(m - \widehat{r}_\alpha) \times 0$ empty matrix.

Otherwise, let

$$\widehat{E}_{b,+} \in \mathbb{R}^{(m-\widehat{r}_\alpha) \times \widehat{s}_b}$$

collect the eigenvectors corresponding to the eigenvalues that satisfy the hard-thresholding condition, and let

$$\widehat{\Lambda}_{b,+} = \text{diag}(\widehat{\omega}_{b,j}^+ : \widehat{\omega}_{b,j}^+ > \tau_{\Omega,b,n}).$$

Construct the estimated thin square root

$$\widehat{L}_b = \widehat{E}_{b,+} \widehat{\Lambda}_{b,+}^{1/2}.$$

In either case, the hard-thresholded covariance estimate is

$$\widehat{\Omega}_b^{\text{HT}} \equiv \widehat{L}_b \widehat{L}_b'.$$

Define the set of databases with a nonzero estimated residual rank:

$$\widehat{\mathcal{B}}_+ = \{b : \widehat{s}_b > 0\}.$$

Set

$$\widehat{s} = \sum_{b \in \widehat{\mathcal{B}}_+} \widehat{s}_b$$

and concatenate the nonempty thin-square-root blocks:

$$\widehat{L} = \left(\widehat{L}_b : b \in \widehat{\mathcal{B}}_+ \right) \in \mathbb{R}^{(m-\widehat{r}_\alpha) \times \widehat{s}}.$$

If $\widehat{s} = 0$, set

$$\widehat{c}_{1-\tau} = 0,$$

and proceed directly to the Output step.

C. Construction of $\widehat{\mathcal{C}}$ and computation of the critical value.

1. Define the estimated residual-space correlation set

$$\widehat{\mathcal{C}} = \left\{ C \in \mathbb{S}_+^{\widehat{s}} : C_{bb} = I_{\widehat{s}_b} \text{ for every } b \in \widehat{\mathcal{B}}_+ \right\},$$

where the block ordering agrees with the ordering used in $\widehat{L}$.

2. For each candidate $C \in \widehat{\mathcal{C}}$, form

$$\widehat{Q}(C) = \widehat{L}C\widehat{L}'$$

and numerically symmetrize it:

$$\widehat{Q}(C) \leftarrow \frac{\widehat{Q}(C) + \widehat{Q}(C)'}{2}.$$

Compute the eigenvalues of $\widehat{Q}(C)$. Replace only negative eigenvalues attributable to floating-point error by zero. Do not apply the rank-selection thresholds again to $\widehat{Q}(C)$.

If all eigenvalues are zero, set

$$q_{1-\tau}(\widehat{Q}(C)) = 0.$$

Otherwise, denoting the nonnegative eigenvalues by $\widehat{\lambda}_j(C)$, compute

$$q_{1-\tau}(\widehat{Q}(C)) = \text{Quantile}_{1-\tau} \left(\sum_j \widehat{\lambda}_j(C) \chi_{1,j}^2 \right),$$

where the $\chi_{1,j}^2$ variables are mutually independent. Evaluate this quantile using a generalized chi-square distribution function and scalar inversion of that function.

3. Solve the global constrained maximization problem

$$\widehat{c}_{1-\tau} = \max_{C \in \widehat{\mathcal{C}}} q_{1-\tau} \left(\widehat{Q}(C) \right).$$

If $|\widehat{\mathcal{B}}_+| = 1$, then $\widehat{\mathcal{C}}$ is the singleton $\{I_{\widehat{s}}\}$, and no numerical maximization is required.

The numerical approximation must satisfy the accuracy requirement in Appendix A.4.

Output. The critical value $\widehat{c}_{1-\tau}$.

The feasible test is

$$\widehat{\phi}_n(\bar{\beta}) = \mathbf{1} \{m - \widehat{r}_\alpha > 0, \widehat{s} > 0\} \mathbf{1} \{T_n > \widehat{c}_{1-\tau}\}.$$

B Main Proofs

B.1 Null distribution proof

Proof of Theorem 3.1. Let

$$\widehat{\alpha} \in \arg \min_{\alpha \in \mathcal{A}} [\widehat{\theta} - g(\alpha, \bar{\beta})]' \widehat{W} [\widehat{\theta} - g(\alpha, \bar{\beta})],$$

and define the fitted reduced-form value

$$\widehat{\mu}_n \equiv g(\widehat{\alpha}, \bar{\beta}).$$

The first goal of the proof is to show that

$$\sqrt{n}(\widehat{\mu}_n - \theta_0) = D_0(D_0' W D_0)^{-1} D_0' W \sqrt{n}(\widehat{\theta} - \theta_0) + o_p(1).$$

We then use this expansion to show that the test statistic converges in distribution to $Z' R_W Z$, where $Z \sim N(0, \Sigma_0)$ and the closed-form expression for R_W is given below. This quadratic-form representation establishes the generalized chi-square limiting distribution.

Step 1: Localization of the fitted value. Under H_0 , $\theta_0 \in \Theta_0(\bar{\beta})$. Therefore, by the definition of $\widehat{\mu}_n$,

$$(\widehat{\theta} - \widehat{\mu}_n)' \widehat{W} (\widehat{\theta} - \widehat{\mu}_n) \leq (\widehat{\theta} - \theta_0)' \widehat{W} (\widehat{\theta} - \theta_0).$$

Since $\widehat{W} \xrightarrow{p} W \succ 0$,

$$\lambda_{\min}(\widehat{W})$$

is bounded away from zero with probability approaching one, while $\lambda_{\max}(\widehat{W}) = O_p(1)$. Assumption 1 therefore gives

$$\|\widehat{\theta} - \widehat{\mu}_n\| = O_p(n^{-1/2}),$$

and hence

$$\widehat{\mu}_n - \theta_0 = O_p(n^{-1/2}).$$

Consequently, with probability approaching one,

$$\widehat{\mu}_n \in \Theta_0(\bar{\beta}) \cap V.$$

Step 2: Work with the locally identified coordinates. By Assumption 2 and the constant-rank theorem, there exists a local C^2 chart

$$h : \Gamma \subset \mathbb{R}^{r_\alpha} \longrightarrow \Theta_0(\bar{\beta}) \cap V$$

such that, for some interior point γ_0 ,

$$h(\gamma_0) = \theta_0,$$

and

$$D_0 \equiv \nabla_\gamma h(\gamma_0) \in \mathbb{R}^{m \times r_\alpha}$$

has full column rank and satisfies

$$\text{col}(D_0) = \text{col}(G_0).$$

Because $\widehat{\mu}_n - \theta_0 = O_p(n^{-1/2})$, there exists $\widehat{\gamma}_n$ satisfying

$$\widehat{\mu}_n = h(\widehat{\gamma}_n), \quad \widehat{\gamma}_n \xrightarrow{p} \gamma_0.$$

Since h is a local chart and $D_0 = \nabla_\gamma h(\gamma_0)$ has full column rank, its local inverse on the restricted-model manifold is locally Lipschitz. Therefore,

$$\widehat{\gamma}_n - \gamma_0 = O_p(n^{-1/2}).$$

Step 3: First-order conditions in the effective coordinates. Because γ_0 is interior and

$$\widehat{\gamma}_n - \gamma_0 = O_p(n^{-1/2}),$$

the first-order conditions hold with probability approaching one:

$$\nabla_\gamma h(\widehat{\gamma}_n)' \widehat{W} [\widehat{\theta} - h(\widehat{\gamma}_n)] = 0.$$

Let

$$\tau_n \equiv \sqrt{n}(\hat{\gamma}_n - \gamma_0), \quad \Delta_n \equiv \sqrt{n}(\hat{\theta} - \theta_0).$$

A first-order Taylor expansion gives

$$\sqrt{n} \{h(\hat{\gamma}_n) - \theta_0\} = D_0 \tau_n + o_p(1)$$

and

$$\nabla_\gamma h(\hat{\gamma}_n) = D_0 + o_p(1).$$

Multiplying the first-order conditions by $\sqrt{n}$, using $\widehat{W} = W + o_p(1)$, and recalling that $\Delta_n = O_p(1)$ and $\tau_n = O_p(1)$, yields

$$D'_0 W (\Delta_n - D_0 \tau_n) = o_p(1).$$

Since $D'_0 W D_0$ is nonsingular,

$$\tau_n = (D'_0 W D_0)^{-1} D'_0 W \Delta_n + o_p(1).$$

It follows that

$$\sqrt{n}(\hat{\mu}_n - \theta_0) = D_0 (D'_0 W D_0)^{-1} D'_0 W \Delta_n + o_p(1).$$

Because

$$\text{col}(D_0) = \text{col}(G_0),$$

$$D_0 (D'_0 W D_0)^{-1} D'_0 W = G_0 (G'_0 W G_0)^\dagger G'_0 W.$$

Thus, the correct replacement for the invalid expansion of $\hat{\alpha} - \alpha_0$ is

$$\sqrt{n}(\hat{\mu}_n - \theta_0) = G_0 (G'_0 W G_0)^\dagger G'_0 W \Delta_n + o_p(1).$$

Step 4: Expansion of the statistic. Consequently,

$$\sqrt{n}(\hat{\theta} - \hat{\mu}_n) = [I_m - G_0 (G'_0 W G_0)^\dagger G'_0 W] \Delta_n + o_p(1).$$

Since the expression in brackets is $O_p(1)$ and $\widehat{W} - W = o_p(1)$,

$$\begin{aligned} T_n &= n(\hat{\theta} - \hat{\mu}_n)' \widehat{W} (\hat{\theta} - \hat{\mu}_n) \\ &= \Delta'_n [W - W G_0 (G'_0 W G_0)^\dagger G'_0 W] \Delta_n + o_p(1) \\ &= \Delta'_n R_W \Delta_n + o_p(1). \end{aligned}$$

Equivalently, writing

$$X_0 = W^{1/2} G_0,$$

$$R_W = W^{1/2} [I_m - X_0 (X_0' X_0)^\dagger X_0'] W^{1/2}.$$

This recovers exactly the matrix already defined in the main text.

Step 5: Generalized chi-square representation. By Assumption [1](#) and the continuous mapping theorem,

$$T_n \xrightarrow{d} Z' R_W Z, \quad Z \sim N(0, \Sigma_0).$$

Write

$$Z = \Sigma_0^{1/2} \varepsilon, \quad \varepsilon \sim N(0, I_m).$$

Then

$$Z' R_W Z = \varepsilon' \Sigma_0^{1/2} R_W \Sigma_0^{1/2} \varepsilon.$$

Let $\lambda_1, \dots, \lambda_{r_Q}$ be the strictly positive eigenvalues of

$$\Sigma_0^{1/2} R_W \Sigma_0^{1/2}.$$

An orthogonal spectral decomposition gives

$$Z' R_W Z \stackrel{d}{=} \sum_{j=1}^{r_Q} \lambda_j \chi_{1,j}^2.$$

Finally,

$$\Sigma_0^{1/2} R_W \Sigma_0^{1/2} = \left(R_W^{1/2} \Sigma_0^{1/2} \right)' \left(R_W^{1/2} \Sigma_0^{1/2} \right),$$

whereas

$$R_W^{1/2} \Sigma_0 R_W^{1/2} = \left(R_W^{1/2} \Sigma_0^{1/2} \right) \left(R_W^{1/2} \Sigma_0^{1/2} \right)'.$$

These two matrices have the same strictly positive eigenvalues. Therefore, the weights can equivalently be defined as the positive eigenvalues of

$$R_W^{1/2} \Sigma_0 R_W^{1/2}.$$

This proves the result. ■

B.2 Power proofs

Proof of Theorem 4.1. Write

$$Q_n(\mu) = (\hat{\theta}_n - \mu)' \widehat{W} (\hat{\theta}_n - \mu), \quad Q(\mu; \theta) = (\theta - \mu)' W (\theta - \mu).$$

Let

$$J_W(x) \equiv \inf_{\mu \in \Theta_0(\bar{\beta})} Q(\mu; x), \quad \|v\|_W \equiv (v' W v)^{1/2}.$$

By the triangle inequality, the distance to the model set satisfies

$$\left| \sqrt{J_W(\hat{\theta}_n)} - \sqrt{J_W(\theta)} \right| \leq \|\hat{\theta}_n - \theta\|_W \xrightarrow{p} 0.$$

Moreover, define

$$\varepsilon_n \equiv \frac{\|\widehat{W} - W\|_{\text{op}}}{\lambda_{\min}(W)} \xrightarrow{p} 0.$$

With probability approaching one, $\varepsilon_n < 1$, and

$$(1 - \varepsilon_n) J_W(\hat{\theta}_n) \leq \inf_{\mu \in \Theta_0(\bar{\beta})} Q_n(\mu) \leq (1 + \varepsilon_n) J_W(\hat{\theta}_n).$$

Since

$$\frac{T_n}{n} = \inf_{\mu \in \Theta_0(\bar{\beta})} Q_n(\mu),$$

it follows that

$$\frac{T_n}{n} \xrightarrow{p} J_W(\theta).$$

If $J_W(\theta) > 0$, then

$$T_n = n J_W(\theta) + o_p(n) \xrightarrow{p} \infty,$$

whereas the robust critical value $c_{1-\tau}^*(R_W)$ is finite by assumption. Hence the rejection probability converges to one.

Finally, because $W \succ 0$, its quadratic norm is equivalent to the Euclidean norm. Therefore

$$J_W(\theta) = 0 \iff \inf_{\mu \in \Theta_0(\bar{\beta})} \|\theta - \mu\|_2 = 0 \iff \theta \in \overline{\Theta_0(\bar{\beta})},$$

which gives the stated characterization. ■

Proof of Theorem 4.2. Let

$$\Delta_n \equiv \sqrt{n}(\hat{\theta}_n - \theta_0).$$

Under the local alternative,

$$\Delta_n = \sqrt{n}(\hat{\theta}_n - \theta_n) + \sqrt{n}(\theta_n - \theta_0) \xrightarrow{d} Z + h.$$

The local quadratic expansion gives

$$T_n = \inf_a (\Delta_n - G_0 a)' W (\Delta_n - G_0 a) + o_p(1) = \Delta_n' R_W \Delta_n + o_p(1).$$

Since $r_Q > 0$, the limiting quadratic form has a continuous distribution. The stated limiting distribution and local power function therefore follow from the continuous mapping theorem and convergence of the critical value.

To characterize the detectable directions, write

$$R_W = W^{1/2}(I - P_0)W^{1/2},$$

where P_0 is the orthogonal projection onto $\text{col}(W^{1/2}G_0)$. Hence

$$h' R_W h = \inf_a (h - G_0 a)' W (h - G_0 a).$$

Because $W \succ 0$, this expression is zero if and only if $h \in \text{col}(G_0) = \mathcal{T}_0$.

If $h \in \mathcal{T}_0$, then $R_W h = 0$ and therefore

$$(Z + h)' R_W (Z + h) = Z' R_W Z.$$

The local limit is thus identical to the null limit. By contrast, if $h \notin \mathcal{T}_0$, set

$$Y = R_W^{1/2} Z, \quad u = R_W^{1/2} h \neq 0.$$

Choose an orthogonal matrix V such that

$$V' \text{Var}(Y) V = \text{diag}(\lambda_1, \dots, \lambda_{r_Q}, 0, \dots, 0), \quad \lambda_j > 0.$$

Writing $v = V' u$, we obtain

$$(Z + h)' R_W (Z + h) = \|Y + u\|_2^2 \stackrel{d}{=} \sum_{j=1}^{r_Q} \lambda_j \chi_{1,j}^2 \left(\frac{v_j^2}{\lambda_j} \right) + \sum_{j=r_Q+1}^m v_j^2,$$

where the noncentral chi-square variables are mutually independent.

A noncentral χ_1^2 variable with positive noncentrality has a strictly larger upper-tail probability than its central counterpart at every positive threshold. Independent summation preserves this

ordering, and the ordering remains strict for the weighted sum because the remaining random summands have positive probability arbitrarily close to zero. A positive deterministic shift also strictly increases the upper-tail probability, since the central weighted sum has a continuous, strictly increasing distribution function on $(0, \infty)$.

Since $v \neq 0$, there is either a positive noncentrality or a positive deterministic shift. Therefore, at the finite positive critical value $c = c_{1-\tau}^*(R_W)$,

$$\pi_\tau^{\text{loc}}(h) = \Pr\{(Z + h)'R_W(Z + h) > c\} > \Pr\{Z'R_W Z > c\} = \pi_\tau^{\text{loc}}(0).$$

Together with the equality for $h \in \mathcal{T}_0$, this establishes the stated local-power inequality and its equality characterization. ■

Proof of Lemma 4.3. Joint differentiability gives

$$\sqrt{n}(\theta_n - \theta_0) = G_0 a + D_0 d + o(1).$$

The local quadratic expansion and $R_W G_0 = 0$ imply

$$T_n \Rightarrow (Z + G_0 a + D_0 d)'R_W(Z + G_0 a + D_0 d) = (Z + D_0 d)'R_W(Z + D_0 d).$$

Thus, the limiting statistic does not depend on a . Moreover,

$$d'D_0'R_W D_0 d = \min_b (D_0 d - G_0 b)'W(D_0 d - G_0 b).$$

Since $W \succ 0$, this minimum is zero exactly when $D_0 d \in \text{col}(G_0)$. In that case $R_W D_0 d = 0$, and the limiting statistic has its null law.

Otherwise, set $Y = R_W^{1/2} Z$ and $u = R_W^{1/2} D_0 d \neq 0$. Diagonalizing the covariance of Y expresses $\|Y + u\|^2$ as a sum of independent weighted noncentral chi-square variables, plus the squared component of u in the covariance nullspace. A nonzero u produces either a positive noncentrality or a positive deterministic shift. Since $r_Q > 0$, either change strictly increases the upper-tail probability at every finite positive threshold c .

Convergence of the critical value gives the stated rejection probabilities. The kernel characterization follows from $R_W \succeq 0$. ■

B.3 Residual reparametrization proofs

Proof of Lemma A.1. Consider first an admissible covariance completion $\tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B)$, and let

$$Z = (Z'_1, \dots, Z'_B)' \sim N(0, \tilde{\Sigma}).$$

Define $Y_b = H_b Z_b$ and

$$\Gamma \equiv \text{Var} \begin{pmatrix} Y_1 \\ \vdots \\ Y_B \end{pmatrix}.$$

Then $\Gamma \succeq 0$, and its diagonal blocks satisfy

$$\Gamma_{bb} = H_b \Sigma_b H_b' = \Omega_b.$$

Thus, $\Gamma \in \mathcal{G}(\Omega_1, \dots, \Omega_B)$. Moreover,

$$H \tilde{\Sigma} H' = \text{Var} \left(\sum_{b=1}^B Y_b \right) = S_B \Gamma S_B'.$$

This proves one inclusion in (23).

For the converse, take any $\Gamma \in \mathcal{G}(\Omega_1, \dots, \Omega_B)$, and let

$$(Y_1', \dots, Y_B')' \sim N(0, \Gamma).$$

For each b , define

$$A_b \equiv \Sigma_b H_b' \Omega_b^\dagger$$

and

$$\Psi_b \equiv \Sigma_b - \Sigma_b H_b' \Omega_b^\dagger H_b \Sigma_b. \quad (25)$$

The matrix Ψ_b is positive semidefinite. To see this, let

$$K_b \equiv H_b \Sigma_b^{1/2}, \quad \Omega_b = K_b K_b'.$$

Then

$$\Psi_b = \Sigma_b^{1/2} [I_{m_b} - K_b' (K_b K_b')^\dagger K_b] \Sigma_b^{1/2}.$$

The matrix

$$K_b' (K_b K_b')^\dagger K_b$$

is the orthogonal projection onto $\text{col}(K_b')$. Hence, the matrix in brackets is also an orthogonal projection and is positive semidefinite.

Let

$$e_b \sim N(0, \Psi_b), \quad b = 1, \dots, B,$$

be mutually independent and independent of $(Y'_1, \dots, Y'_B)'$, and define

$$Z_b \equiv A_b Y_b + e_b. \quad (26)$$

Using $\Omega_b^\dagger \Omega_b \Omega_b^\dagger = \Omega_b^\dagger$, we obtain

$$\begin{aligned} \text{Var}(Z_b) &= A_b \Omega_b A_b' + \Psi_b \\ &= \Sigma_b H_b' \Omega_b^\dagger H_b \Sigma_b + \Psi_b \\ &= \Sigma_b. \end{aligned}$$

It remains to verify that Z_b reproduces the desired residual contribution Y_b . Since

$$\text{Var}(Y_b) = \Omega_b,$$

the random vector Y_b belongs to $\text{col}(\Omega_b)$ almost surely. Therefore,

$$\Omega_b \Omega_b^\dagger Y_b = Y_b \quad \text{almost surely.}$$

Furthermore,

$$H_b \Psi_b H_b' = \Omega_b - \Omega_b \Omega_b^\dagger \Omega_b = 0,$$

which implies $H_b e_b = 0$ almost surely. Consequently,

$$\begin{aligned} H_b Z_b &= H_b A_b Y_b + H_b e_b \\ &= \Omega_b \Omega_b^\dagger Y_b \\ &= Y_b \quad \text{almost surely.} \end{aligned}$$

Let

$$Z = (Z'_1, \dots, Z'_B)', \quad \tilde{\Sigma} \equiv \text{Var}(Z).$$

Because $\tilde{\Sigma}$ is a covariance matrix, $\tilde{\Sigma} \succeq 0$, and the calculation above shows that $\tilde{\Sigma}_{bb} = \Sigma_b$. Hence,

$$\tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B).$$

Finally,

$$H \tilde{\Sigma} H' = \text{Var} \left(\sum_{b=1}^B H_b Z_b \right) = \text{Var} \left(\sum_{b=1}^B Y_b \right) = S_B \Gamma S_B'.$$

This proves the reverse inclusion. $\blacksquare$

Proof of Lemma A.2. Take first any $C \in \mathcal{C}$, and let

$$U = (U'_1, \dots, U'_B)' \sim N(0, C).$$

Define

$$Y_b = L_b U_b.$$

Then

$$\text{Var}(Y_b) = L_b C_{bb} L'_b = L_b L'_b = \Omega_b.$$

Therefore,

$$\Gamma \equiv \text{Var} \begin{pmatrix} Y_1 \\ \vdots \\ Y_B \end{pmatrix} = \mathcal{L} C \mathcal{L}'$$

belongs to $\mathcal{G}(\Omega_1, \dots, \Omega_B)$.

Conversely, take $\Gamma \in \mathcal{G}(\Omega_1, \dots, \Omega_B)$, and let

$$(Y'_1, \dots, Y'_B)' \sim N(0, \Gamma).$$

Define

$$U_b \equiv L_b^\dagger Y_b.$$

Because

$$\text{col}(\Omega_b) = \text{col}(L_b),$$

and Y_b belongs to this column space almost surely,

$$Y_b = L_b L_b^\dagger Y_b = L_b U_b \quad \text{almost surely.}$$

Moreover,

$$\begin{aligned} \text{Var}(U_b) &= L_b^\dagger \Omega_b (L_b^\dagger)' \\ &= L_b^\dagger L_b L'_b (L_b^\dagger)' \\ &= I_{s_b}. \end{aligned}$$

Therefore, if

$$C \equiv \text{Var} \begin{pmatrix} U_1 \\ \vdots \\ U_B \end{pmatrix},$$

then $C \succeq 0$ and $C_{bb} = I_{s_b}$, so $C \in \mathcal{C}$. Finally,

$$\Gamma = \mathcal{L}C\mathcal{L}'.$$

This establishes (24). ■

Proof of Theorem 5.1. Combining the projection identity (22) with Lemma A.1 gives

$$\begin{aligned} c_{1-\tau}^*(R_W) &= \sup_{\tilde{\Sigma} \in \mathcal{S}(\Sigma_1, \dots, \Sigma_B)} q_{1-\tau} \left(H \tilde{\Sigma} H' \right) \\ &= \sup_{\Gamma \in \mathcal{G}(\Omega_1, \dots, \Omega_B)} q_{1-\tau} (S_B \Gamma S_B'). \end{aligned}$$

By Lemma A.2, every feasible Γ can be written as $\Gamma = \mathcal{L}C\mathcal{L}'$, with $C \in \mathcal{C}$. Since $S_B\mathcal{L} = L$,

$$S_B \Gamma S_B' = S_B \mathcal{L}C\mathcal{L}' S_B' = LCL'.$$

It follows that

$$c_{1-\tau}^*(R_W) = \sup_{C \in \mathcal{C}} q_{1-\tau}(LCL'). \quad (27)$$

It remains to show that the supremum is attained. Let

$$s \equiv \sum_{b=1}^B s_b.$$

The set $\mathcal{C} \subseteq \mathbb{S}_+^s$ is closed. Moreover, every $C \in \mathcal{C}$ satisfies

$$\text{tr}(C) = \sum_{b=1}^B s_b = s.$$

Since $C \succeq 0$,

$$\|C\|_F \leq \text{tr}(C) = s,$$

so $\mathcal{C}$ is bounded and therefore compact.

The mapping

$$C \longmapsto LCL'$$

is continuous, and the generalized chi-square quantile $Q \mapsto q_{1-\tau}(Q)$ is continuous on $\mathbb{S}_+^d$. To see the latter, let $Q_k \rightarrow Q$ and use the same $\varepsilon \sim N(0, I_d)$ for all k . Then

$$\varepsilon' Q_k \varepsilon \longrightarrow \varepsilon' Q \varepsilon \quad \text{almost surely.}$$

When $Q \neq 0$, the limiting generalized chi-square distribution has a continuous and strictly

increasing distribution function on $(0, \infty)$, implying convergence of its $(1 - \tau)$ -quantiles. When $Q = 0$,

$$0 \leq q_{1-\tau}(Q_k) \leq \|Q_k\|_{\text{op}} q_{1-\tau}(I_d) \longrightarrow 0.$$

Thus, the objective in (27) is continuous on a compact set. Consequently,

$$c_{1-\tau}^*(R_W) = \max_{C \in \mathcal{C}} q_{1-\tau}(LCL'),$$

as claimed.

In particular, a maximizing C^* exists. Applying the lifting construction in Lemma A.1 to

$$\Gamma^* = \mathcal{L}C^*\mathcal{L}'$$

produces an admissible covariance completion $\tilde{\Sigma}^*$ attaining the original worst-case critical value. The least-favorable completion need not be unique. ■

Proof Lemma A.3.

Weyl's inequality gives

$$\max_{1 \leq j \leq k_A} |\hat{\sigma}_{j,n} - \sigma_j(A)| \leq \|\hat{A} - A\|_{\text{op}}.$$

For every $j > r_A$,

$$\frac{\hat{\sigma}_{j,n}}{\tau_{A,n}} \leq \frac{\|\hat{A} - A\|_{\text{op}}}{\tau_{A,n}} \xrightarrow{p} 0,$$

so all population-zero singular values are removed with probability approaching one. Moreover, when $r_A > 0$,

$$\min_{1 \leq j \leq r_A} \hat{\sigma}_{j,n} \geq \sigma_{r_A}(A) - \|\hat{A} - A\|_{\text{op}}.$$

Since $\sigma_{r_A}(A) > 0$ is fixed and $\tau_{A,n} \xrightarrow{p} 0$, all positive population singular values are retained with probability approaching one.

■

Proof of Proposition A.4. We first establish consistency of the estimated ranks and residual covariance factors. We then use invariance to orthogonal rotations and continuity of the quantile maximization problem to obtain critical-value consistency.

Since $W \succ 0$, the assumed convergence rate of $\widehat{W}$ implies the same rate for its symmetric square root. Consequently, for

$$\widehat{X} \equiv \widehat{W}^{1/2} \widehat{G}, \quad X^* \equiv W^{1/2} G^*,$$

we have

$$\|\widehat{X} - X^*\|_{\text{op}} = O_p(n^{-1/2}).$$

The threshold satisfies $\tau_{G,n} \xrightarrow{p} 0$ and $\sqrt{n} \tau_{G,n} \xrightarrow{p} \infty$. Weyl's inequality therefore implies that all fixed positive population singular values are retained and all population zero singular values are removed with probability approaching one. Hence,

$$\Pr\{\widehat{r}_\alpha = r_\alpha\} \longrightarrow 1.$$

If $m - r_\alpha = 0$, correct tangent-rank selection implies $\widehat{c}_{1-\tau} = 0$ and nonrejection. Since $R_W = 0$ in this case, both conclusions follow. Henceforth, suppose that $m - r_\alpha > 0$.

On the event of correct tangent-rank selection, the factor-level perturbation condition gives orthogonal matrices Q_n , common to all database blocks, such that

$$\|\widehat{M}_b - Q_n M_b\|_{\text{op}} = O_p(n^{-1/2}), \quad b = 1, \dots, B.$$

Because

$$\widehat{\Omega}_b = \widehat{M}_b \widehat{M}_b', \quad \Omega_b = M_b M_b',$$

this also implies

$$\|\widehat{\Omega}_b - Q_n \Omega_b Q_n'\|_{\text{op}} = O_p(n^{-1/2}), \quad b = 1, \dots, B.$$

For rank selection, however, we use the factor-level rate rather than this covariance-level bound. Orthogonal rotations preserve singular values. Since $\text{rank}(M_b) = s_b$, Weyl's inequality gives

$$s_j(\widehat{M}_b) = O_p(n^{-1/2}), \quad j > s_b.$$

Here singular values are padded with zeros when necessary. The corresponding covariance eigenvalues therefore satisfy

$$\lambda_j(\widehat{\Omega}_b) = s_j(\widehat{M}_b)^2 = O_p(n^{-1}), \quad j > s_b.$$

Since

$$\tau_{\Omega,b,n} \xrightarrow{p} 0, \quad n\tau_{\Omega,b,n} \xrightarrow{p} \infty,$$

all population-zero eigenvalues are removed with probability approaching one. Every fixed positive population eigenvalue is retained because the covariance estimates converge, up to orthogonal rotation, and the threshold tends to zero. Thus,

$$\Pr\{\widehat{s}_b = s_b \text{ for every } b\} \longrightarrow 1.$$

There are finitely many databases, so the rank decisions are jointly consistent.

Thresholding changes each covariance estimate by at most the largest discarded nonnegative

eigenvalue, which is bounded by $\tau_{\Omega,b,n}$. Thus,

$$\|\widehat{\Omega}_b^{\text{HT}} - Q_n \Omega_b Q_n'\|_{\text{op}} \xrightarrow{p} 0.$$

On the event of correct rank selection, thin square roots can consequently be aligned by orthogonal matrices $O_{b,n} \in \mathbb{R}^{s_b \times s_b}$ so that

$$\|\widehat{L}_b - Q_n L_b O_{b,n}\|_{\text{op}} \xrightarrow{p} 0.$$

Only blocks with $s_b > 0$ enter this expression.

Suppose first that $s = \sum_{b=1}^B s_b > 0$. On the event of correct rank selection, $\widehat{\mathcal{C}} = \mathcal{C}$ under the same block ordering. Let

$$O_n \equiv \text{diag}(O_{b,n} : s_b > 0), \quad \bar{L}_n \equiv Q_n' \widehat{L} O_n'.$$

Then

$$\|\bar{L}_n - L\|_{\text{op}} \xrightarrow{p} 0.$$

The block-correlation set is invariant under these blockwise rotations:

$$C \in \mathcal{C} \iff O_n C O_n' \in \mathcal{C}.$$

Moreover, common rotations in residual coordinates preserve the eigenvalues of the quadratic-form covariance matrix. It follows that

$$\widehat{c}_{1-\tau} = \max_{C \in \mathcal{C}} q_{1-\tau}(\bar{L}_n C \bar{L}_n').$$

The set $\mathcal{C}$ is compact, and the generalized chi-square quantile is continuous on the positive-semidefinite cone, including at the zero matrix, as established in the proof of Theorem 5.1. Consequently,

$$\sup_{C \in \mathcal{C}} \left| q_{1-\tau}(\bar{L}_n C \bar{L}_n') - q_{1-\tau}(L C L') \right| \xrightarrow{p} 0.$$

Taking maxima and applying Theorem 5.1 gives

$$\widehat{c}_{1-\tau} \xrightarrow{p} \max_{C \in \mathcal{C}} q_{1-\tau}(L C L') = c_{1-\tau}^*(R_W).$$

If $s = 0$, every population residual covariance block is zero. Joint rank consistency then implies $\widehat{s} = 0$ with probability approaching one, so both the population and estimated critical values are zero asymptotically.

It remains to establish size control. When $s > 0$, the admissible choice $C = I_s$ gives

$$LCL' = LL' = \sum_{b=1}^B \Omega_b \neq 0.$$

Its $(1 - \tau)$ -quantile is strictly positive, and hence $c_{1-\tau}^*(R_W) > 0$. For any admissible true covariance completion, $Z'R_W Z$ either has a continuous generalized chi-square distribution or is identically zero. In either case, its distribution function is continuous at this positive critical value. The null limit, critical-value consistency, and consistent nondegeneracy decision therefore imply

$$\Pr\{\widehat{\phi}_n(\bar{\beta}) = 1\} \longrightarrow \Pr\{Z'R_W Z > c_{1-\tau}^*(R_W)\} \leq \tau.$$

When $s = 0$, the nonrejection rule and rank consistency instead give $\Pr\{\widehat{\phi}_n(\bar{\beta}) = 1\} \rightarrow 0$.

Finally, replacing the exact plug-in critical value by an approximation with $o_p(1)$ error preserves its probability limit. Retaining the same nonrejection rule, the size argument therefore also applies to the numerical implementation. ■

References

- AN, SUNGBAE AND FRANK SCHORFHEIDE (2007): “Bayesian analysis of DSGE models,” *Econometric reviews*, 26 (2-4), 113–172.
- ANDREWS, DONALD WK AND PATRIK GUGGENBERGER (2019): “Identification-and singularity-robust inference for moment condition models,” *Quantitative Economics*, 10 (4), 1703–1746.
- ANDREWS, DONALD WK AND WERNER PLOBERGER (1994): “Optimal tests when a nuisance parameter is present only under the alternative,” *Econometrica: Journal of the Econometric Society*, 1383–1414.
- ANDREWS, ISAIAH AND ANNA MIKUSHEVA (2016): “A geometric approach to nonlinear econometric models,” *Econometrica*, 84 (3), 1249–1264.
- ANTOINE, BERTILLE AND ERIC RENAULT (2009): “Efficient GMM with nearly-weak instruments,” *The Econometrics Journal*, 12, S135–S171.
- CHA, JOOYOUNG, YUYA SASAKI, AND NELSON MATTHEW P TAN (2026): “Bounds for Standard Errors in Combined Data,” *arXiv preprint arXiv:2606.24867*.
- COCCI, MATTHEW D AND MIKKEL PLAGBORG-MØLLER (2025): “Standard errors for calibrated parameters,” *Review of Economic Studies*, 92 (5), 2952–2978.
- DAVIES, ROBERT B (1977): “Hypothesis testing when a nuisance parameter is present only under the alternative,” *Biometrika*, 64 (2), 247–254.
- FORNERON, JEAN-JACQUES (2024): “Detecting identification failure in moment condition models,” *Journal of Econometrics*, 238 (1), 105552.
- GREGORY, ALLAN W AND GREGOR W SMITH (1990): “Calibration as estimation,” *Econometric Reviews*, 9 (1), 57–89.
- GUERRON-QUINTANA, PABLO, ATSUSHI INOUE, AND LUTZ KILIAN (2017): “Impulse response matching estimators for DSGE models,” *Journal of Econometrics*, 196 (1), 144–155.
- HANSEN, BRUCE E (1996): “Inference when a nuisance parameter is not identified under the null hypothesis,” *Econometrica: Journal of the econometric society*, 413–430.
- HANSEN, LARS PETER (1982): “Large sample properties of generalized method of moments estimators,” *Econometrica: Journal of the econometric society*, 1029–1054.

- HANSEN, LARS PETER AND JAMES J HECKMAN (1996): “The empirical foundations of calibration,” *Journal of economic perspectives*, 10 (1), 87–104.
- INOUE, ATSUSHI AND LUTZ KILIAN (2016): “Joint confidence sets for structural impulse responses,” *Journal of Econometrics*, 192 (2), 421–432.
- ISKREV, NIKOLAY (2010): “Local identification in DSGE models,” *Journal of Monetary Economics*, 57 (2), 189–202.
- KLEIBERGEN, FRANK (2005): “Testing parameters in GMM without assuming that they are identified,” *Econometrica*, 73 (4), 1103–1123.
- KOCIECKI, ANDRZEJ AND MARCIN KOLASA (2018): “Global identification of linearized DSGE models,” *Quantitative Economics*, 9 (3), 1243–1263.
- (2023): “A solution to the global identification problem in DSGE models,” *Journal of Econometrics*, 236 (2), 105477.
- KOMUNJER, IVANA AND SERENA NG (2011): “Dynamic identification of dynamic stochastic general equilibrium models,” *Econometrica*, 79 (6), 1995–2032.
- NAKAMURA, EMI AND JÓN STEINSSON (2018): “High-frequency identification of monetary non-neutrality: the information effect,” *The Quarterly Journal of Economics*, 133 (3), 1283–1330.
- QU, ZHONGJUN AND DENIS TKACHENKO (2012): “Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models,” *Quantitative Economics*, 3 (1), 95–132.
- (2017): “Global identification in DSGE models allowing for indeterminacy,” *The Review of Economic Studies*, 84 (3), 1306–1345.
- (2023): “Using arbitrary precision arithmetic to sharpen identification analysis for DSGE models,” *Journal of Applied Econometrics*, 38 (4), 644–667.
- SARGAN, JOHN D (1958): “The estimation of economic relationships using instrumental variables,” *Econometrica: Journal of the econometric society*, 393–415.
- SHAPIRO, ALEXANDER (1986): “Asymptotic theory of overparameterized structural models,” *Journal of the American Statistical Association*, 81 (393), 142–149.
- STAIGER, DOUGLAS O AND JAMES H STOCK (1994): “Instrumental variables regression with weak instruments,” .

STOCK, JAMES H AND JONATHAN H WRIGHT (2000): “GMM with weak identification,” *Econometrica*, 68 (5), 1055–1096.

VOHRA, VEDANT (2026): “Inference for welfare metrics,” *Journal of Public Economics*, 256, 105603.