ONLINE SUPPLEMENTARY MATERIAL FOR "ADMISSIBLE, SIMILAR TESTS: A CHARACTERIZATION", BY JOSÉ LUIS MONTIEL OLEA. APPENDIX B.

TABLE OF CONTENTS

B.1. *Lemmas for Theorem 1*

B.1.1. *Topological preliminaries*

NOTATION: Let $\mathcal{B}(\mathbb{R}^n)$ denote the Borel σ-algebra on $\mathbb{R}^n$. For any set $\mathcal{S} \in \mathcal{B}(\mathbb{R}^n)$, let $\mathcal{B}(\mathbb{R}^n)_{\mathcal{S}}$ denote the sub-space σ-algebra. *Measurability* of the function $f : \mathcal{S} \to \mathbb{R}$ is always relative to the measurable spaces $(\mathcal{S}, \mathcal{B}(\mathbb{R}^n)_{\mathcal{S}})$-$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The integral of f with respect to the Lebesgue measure in $\mathbb{R}^n$ is denoted by $\int_{\mathcal{S}} f(s)ds$. *Integration* with respect to a different measure μ is denoted $\int_{\mathcal{S}} f(s)d\mu(s)$ or $\int_{\mathcal{S}} f d\mu$ if no ambiguity arises. All vectors are column vectors. For notational convenience, (a, b) will sometimes replace $(a', b')'$. The dimension of the column vector "a" is denoted d_a.

PRELIMINARIES 1 (L^1 and L^∞): Since the sample space $\mathbf{X} \in \mathcal{B}(\mathbb{R}^s)$, the triplet $(\mathbf{X}, \mathcal{B}(\mathbb{R}^s)_{\mathbf{X}}, \lambda^s)$ is a well-defined σ-finite measure space, where λ^S denotes the Lebesgue measure in $\mathbb{R}^s$ restricted to $\mathbf{X}$. Note that $\mathcal{B}(\mathbb{R}^s)_{\mathbf{X}} = \mathcal{B}(\mathbf{X})$ whenever $\mathbf{X}$ is endowed with the sub-space topology relative to $\mathbb{R}^s$. Following Rudin (2006), p. 65, let $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$ denote the space of all real-valued measurable functions f that satisfy $||f||_1 \equiv \int_{\mathbf{X}} |f(x)|dx < \infty$. Let $L^\infty(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$ denote the class of all essentially bounded real-valued measurable functions (Rudin (2006) p. 66).

REMARK 1: Identify the class of all tests $\mathcal{C}$ as a subset of $L^\infty(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$

$$\mathcal{C} \equiv \{\phi \in L^\infty(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s) \mid \phi(x) \in [0,1] \text{ for } \lambda^s\text{-a.e. } x \in \mathbf{X}\}.$$

And note that the elements of any statistical model $\{f(x;\theta)\}_{\theta\in\mathbf{\Theta}}$ are elements of $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$, by the definition of probability density function $\int_{\mathbf{X}} f(x;\theta)dx = 1 < \infty$ for all $\theta \in \mathbf{\Theta}$.

PRELIMINARIES 2 (The dual space of L^1): Let $[L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)]^*$ denote the dual space of $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$, i.e., the space of all continuous (w.r.t. $||f||_1$) linear functionals on $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$; see Rudin (2005), p. 56. Let Λ denote an element of the dual space $[L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)]^*$. By Theorem

6.16 in Rudin (2006), p. 127 and Theorem 1.18 in Rudin (2005), p. 15; the space $[L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)]^*$ is isometrically isomorphic to $L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$. Therefore, one can identify each functional Λ with a unique element (up to equivalence) $g \in L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$, and vice versa: for $f \in L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)^*$, the functional $\Lambda \in [L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)]^*$ is of the form

$$\Lambda(f) \equiv \int_{\mathbf{X}} g(x)f(x)dx \quad \text{for some} \quad g \in L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s).$$

Preliminaries 3 (weak* topology on L^∞): Endow the space $L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$ with the topology induced by the weak*-topology on the space $[L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)]^*$; see Rudin (2005), p. 67, 68. The new topological space is denoted by $(L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s),\mathcal{T}^*)$. Denote convergence in such topology by $\to^*$. Note that, by definition, $\{g_n\}_{n\in\mathbb{N}} \to^* g$ if and only if

$$\int_{\mathbf{X}} f(x)g_n(x)dx \to \int_{\mathbf{X}} f(x)g(x)dx \quad \forall \quad f \in L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s).$$

Let $(L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s),\mathcal{T}^*)$ be the space of essentially bounded functions topologized with the weak* topology. For any $A \subset L^\infty(\mathbf{X},\mathcal{B}(\mathbf{X})$, let $\mathcal{T}^*_A$ denote the subset topology induced by $\mathcal{T}^*$.

B.1.2. *Proof of Lemma 1*

Proof of Lemma 1: The outline of the proof is the following. We show that the set $\mathcal{C}(\alpha\text{-}s)$ is a sequentially closed subset of $\mathcal{C}$ with the relative weak* topology. Then we use the Banach-Alaoglu theorem and the topological separability of $L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$ to establish the compactness of $\mathcal{C}(\alpha\text{-}s)$.

(*Sequential Closedness*) Take any convergent sequence of tests $\phi_n \to^* \phi$ with $\{\phi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{C}(\alpha\text{-}s)$. We want to show that $\phi \in \mathcal{C}(\alpha\text{-}s)$. First, we show that $\phi(x) \in \mathcal{C}$; i.e., $\phi \in [0,1]$ for almost every $x \in \mathbf{X}$. Suppose not. Then there exists a measurable set $A \in \mathcal{B}(\mathbf{X})$ with $\lambda^s(A) > 0$ such that $\phi(x) > 1$ or $\phi(x) < 0$ for all $x \in A$. Without loss of generality assume $\phi(x) > 1$. Since λ^s is σ-finite, there exists a countable collection $\{E_n\}_{n\in\mathbb{N}}$ such that $\cup_{n\in\mathbb{N}}E_n = \mathbf{X}$ and $\lambda^s(E_n) < \infty$ for every n. Consider the sequence of sets $\{A \cap E_n\}_{n\in\mathbb{N}}$. Note that $0 \le \lambda^s(A\cap E_n) < \infty$ for all $n \in \mathbb{N}$. In addition, there exists $N \in \mathbb{N}$ for which $0 < \lambda^s(A\cap E_N)$, otherwise $\lambda^s(A) = \lambda^s(\cup_{n=1}^\infty(A\cap E_n)) \le \sum_{n=1}^\infty \lambda^s(A\cap E_n) = 0$. Consider the indicator function $\mathbb{1}_{A\cap E_N}$. Since $0 < \lambda^s(A\cap E_N) < \infty$, the indicator function $\mathbb{1}_{A\cap E_N} \in L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$. Note that

$$\lambda^s(A\cap E_N) < \int_{\mathbf{X}} \mathbb{1}_{A\cap E_N}(x)\phi(x)dx = \lim_{n\to\infty}\int_{\mathbf{X}} \mathbb{1}_{A\cap E_N}(x)\phi_n(x)dx \le \lambda^s(A\cap E_N),$$

leading to a contradiction. Therefore $\phi(x) \le 1$ λ^s-almost everywhere in $\mathbf{X}$. An analogous argument yields $\phi(x) \ge 0$ λ^s-almost everywhere. Therefore $\phi \in \mathcal{C}$. Now, we need to show that $\phi \in \mathcal{C}(\alpha\text{-}s)$. By assumption, for every $\theta \in \text{Bd}(\Theta_0)$ $f(\cdot;\theta)$ is an element of $L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$. Consequently, $f(\cdot,\theta) \in L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s)$. Since $\phi_n \in \mathcal{C}(\alpha\text{-}s)$ for every $n \in \mathbb{N}$ weak* convergence yields

$$\begin{aligned} 0 = \lim_{n\to\infty}\int_{\mathbf{X}} f(x;\theta)(\phi_n(x)-\alpha)dx &= \Big(\lim_{n\to\infty}\int_{\mathbf{X}} f(x;\theta)\phi_n(x)dx\Big) - \int_{\mathbf{X}} f(x;\theta)\alpha dx \\ &= \int_{\mathbf{X}} f(x;\theta)\phi(x)dx - \int_{\mathbf{X}} f(x;\theta)\alpha dx \\ &= \int_{\mathbf{X}} f(x;\theta)(\phi(x)-\alpha)dx. \end{aligned}$$

So $\phi \in \mathcal{C}(\alpha\text{-}s)$. This implies $\mathcal{C}(\alpha\text{-}s)$ is sequentially closed in $\mathcal{C}$ endowed with the weak* topology.

(*Compactness*) Let

$$V \equiv \Big\{ f \in L^1(\mathbf{X},\mathcal{B}(\mathbf{X}),\lambda^s) \ : \ \int_{\mathbf{X}} |f(x)|dx \le 1 \Big\}$$

Note that V is a neighborhood of the function $\mathbf{0}$ in the space $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$. Let

$$K \equiv \Big\{ g \in L^\infty(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s) \; : \; \Big| \int_{\mathbf{X}} f(x) g(x) \Big| dx \leq 1 \quad \forall \quad f \in V \Big\}. \tag{B.1}$$

Note that $C(\alpha\text{-}s) \subseteq \mathcal{C} \subseteq K$, as for any test $\left| \int_{\mathbf{X}} f(x)\phi(x)dx \right| \leq \int_{\mathbf{X}} |f(x)|\phi(x)dx \leq \int_{\mathbf{X}} |f(x)|dx \leq 1$. By the Banach-Alaouglu Theorem the set K is compact in the weak* topology; see Rudin (2005), p. 68, Theorem 3.15. Furthermore, the space $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$ is topologically separable as $(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$ is a separable measure space; see exercise 10, Chapter 1 of Stein (2011). Therefore, Theorem 3.16 in Rudin (2005) p. 70 implies that the topological space $(K, \mathcal{T}^*_K)$ is compact and metrizable. Since every metrizable space is first-countable—consequently, Frechet-Urysohn— the sequential closure of C(α-s) coincides with its closure. Therefore, the set $C(\alpha - s)$ is a closed subset of the compact topological space $(K, \mathcal{T}^*_K)$. We conclude that $(\mathcal{C}(\alpha\text{-}s), \mathcal{T}^*_{\mathcal{C}(\alpha\text{-}s)})$ is compact and metrizable. That is, the space of α-similar tests is weak* compact.

Q.E.D.

B.1.3. *Proof of Lemma 2*

Proof of Lemma 2: This lemma has three claims. The first claim, denoted as **L2a**, is that $M(w) \neq \emptyset$. The second claim, **L2b**, is that $\phi^* \in M(w) \implies \phi^*$ is admissible in $\mathcal{C}(\alpha\text{-}s)$. The third claim, **L2c**, is that $\phi^* \in M(w) \implies \phi^*$ is admissible in $\mathcal{C}$. Now we prove these claims.

Proof of L2a: Let p denote the p.d.f of w. We have shown that the class of tests $\mathcal{C}(\alpha$-s) is weak* compact. This class is non-empty, as it contains the randomized test $\phi(x) = \alpha$. To establish L2a it will be sufficient to show that the objective function

$$\mathcal{W}^*(\phi) \equiv \int_{\Theta_1} R(\phi, \theta) p(\theta) d\theta$$

is continuous in the weak* topology.

L2a-Step 1 (Fubini's Theorem:) Since the image of any test $\phi \in \mathcal{C}$ is contained in the interval $[0, 1]$ λ^s-a.e. and $f(x; \theta) \in L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$ for all θ, it follows that $\Big(\int_{\mathbf{X}} \phi(x) f(x; \theta) dx \Big) \leq 1$ for every $\theta \in \Theta$. Furthermore,

$$\int_{\Theta_1} \Big(\int_{\mathbf{X}} \phi(x) f(x; \theta) dx \Big) p(\theta) d\theta \leq 1 < \infty.$$

Therefore, an application of Fubini's theorem in Billingsley (1995), p. 234 yields

$$\int_{\Theta_1} R(\phi, \theta) p(\theta) d\theta \equiv \int_{\Theta_1} \Big(\int_{\mathbf{X}} (1 - \phi(x)) f(x; \theta) dx \Big) p(\theta) d\theta = \int_{\mathbf{X}} (1 - \phi(x)) f_1^*(x) dx$$

where f_1^* is the "integrated" likelihood given by

$$f_1^*(x) \equiv \int_{\Theta_1} f(x; \theta) p(x) d\theta, \tag{B.2}$$

Note that f_1^* is an element of $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda^s)$. We can re-write

$$\mathcal{W}^*(\phi) \equiv \int_{\mathbf{X}} (1 - \phi(x)) f_1^*(x) dx \tag{B.3}$$

L2a-Step 2 (Sequential Continuity of $\mathcal{W}^*$:) We now show that $\mathcal{W}^*$ is continuous on the compact metrizable space $(\mathcal{C}(\alpha\text{-}\mathcal{G}), \mathcal{T}^*_{\mathcal{C}(\alpha\text{-s})})$. It suffices to establish sequential continuity. Take any sequence of tests $\phi_n \to^* \phi$. Since f_1^* is an element of $L^1(\mathbf{X}, \mathcal{B}(\mathbf{X}), \lambda_{\mathbf{X}})$, convergence in the weak* topology yields

$$\int_{\mathbf{X}} \phi_n(x) f_1^*(x) dx \to \int_{\mathbf{X}} \phi(x) f_1^*(x) dx.$$

Therefore

$$\begin{aligned} \mathcal{W}^*(\phi_n) &\equiv 1 - \int_{\mathbf{X}} \phi_n(x) f_1^*(x) dx \to 1 - \int_{\mathbf{X}} \phi(x) f_1^*(x) dx, \\ &= \mathcal{W}^*(\phi). \end{aligned}$$

Therefore, $\mathcal{W}^*$ is a continuous functional defined on the compact space $(\mathcal{C}(\alpha\text{-}s), \mathcal{T}^*_{\mathcal{C}(\alpha\text{-}s)})$, and $\mathcal{C}(\alpha\text{-}s) \neq \emptyset$, as it contains the test $\phi(x) = \alpha$. This implies $M(w) \neq \emptyset$.

L2b : Let $\phi^* \in M(w_1)$. We show that if $\phi' \in \mathcal{C}(\alpha\text{-}s)$ satisfies

$$\text{(B.4)} \qquad \mathbb{E}_\theta[\phi'(X)] \geq \mathbb{E}_\theta[\phi^*(X)] \quad \forall \quad \theta \in \Theta_1$$

then

$$\text{(B.5)} \qquad \mathbb{E}_\theta[\phi'(x)] = \mathbb{E}_\theta[\phi^*(x)] \quad \forall \quad \theta \in \Theta_1.$$

Consequently, there is no test $\phi' \in \mathcal{C}(\alpha\text{-}s)$ that "weakly dominates" ϕ^* ; i.e, $R(\phi', \theta) \leq R(\phi^*, \theta)$ with strict inequality for some θ.

Suppose (B.4) hold, but (B.5) does not. Then, the following is true:

C1: There exists $\tilde{\theta} \in \Theta_1$ such that $\Delta_{\phi', \phi^*}(\tilde{\theta}) \equiv \mathbb{E}_{\tilde{\theta}}[\phi'(X)] - \mathbb{E}_{\tilde{\theta}}[\phi^*(X)] > 0$

C1 and the continuity of $\Delta_{\phi', \phi}(\cdot)$ at $\tilde{\theta}$ implies the existence of an open neighborhood $\tau_{\tilde{\theta}}$ for which $\Delta_{\phi', \phi^*}(\theta) > 0$ for all $\theta \in \tau_{\tilde{\theta}}$. Note that $\Theta_1 \neq \emptyset$ is an open set. It follows that the set $\mathcal{S}_{\tilde{\theta}}$ defined by $\mathcal{S}_{\tilde{\theta}} \equiv \tau_{\tilde{\theta}} \cap \Theta_1$ satisfies three properties: it is non-empty, it is open, and it is contained in Θ_1. Since $w_1(\theta)$ has full support on Θ_1, $\int_A dw_1(\theta) > 0$ for any open set A contained in Θ_1. Note that $\Delta_{\phi', \phi^*}(\theta) > 0$ for all $\theta \in S_{\tilde{\theta}}$ and (B.4) implies

$$\int_{\Theta_1} \Big(\int_{\mathbf{X}} (1 - \phi'(x)) f(x; \theta) dx \Big) dw(\theta) < \int_{\Theta_1} \Big(\int_{\mathbf{X}} (1 - \phi^*(x)) f(x; \theta) dx \Big) dw(\theta)$$

This contradicts the fact that $\phi^* \in M(w_1)$. We conclude C1 cannot hold.

Therefore, (B.4) implies (B.5). We conclude that ϕ^* is admissible in $\mathcal{C}(\alpha\text{-}s)$.

L2c : We now show that a test $\phi^* \in M(w)$ is admissible in the class of all tests. This proof is based on the arguments provided in Chernozhukov, Hansen, and Jansson (2009). The proof is divided into two steps.

Step 1: First we show that if $\phi' \in \mathcal{C}$ satisfies

$$\text{(B.6)} \qquad \mathbb{E}_\theta[\phi'(X)] \leq \mathbb{E}_\theta[\phi^*(X)] \quad \forall \quad \theta \in \Theta_0$$

and

$$\text{(B.7)} \qquad \mathbb{E}_\theta[\phi'(X)] \geq \mathbb{E}_\theta[\phi^*(X)] \quad \forall \quad \theta \in \Theta_1$$

with some strict inequality, then ϕ' is α-similar on $\text{Bd}\Theta_0 = \Theta_0$. Consequently, any test ϕ' that "weakly dominates" ϕ^* (i.e, $R(\phi',\theta) \leq R(\phi^*,\theta)$ with strict inequality for some θ) must be α-similar on the boundary of Θ_0.

Let $\mathcal{C}_{ns} \subset \mathcal{C}$ be the class of tests that are not similar on the boundary of Θ_0. This is, $\phi \in \mathcal{C}_{ns}$ if and only if there exists $\theta, \theta' \in \text{Bd}\Theta_0$ such that $\mathbb{E}_\theta[\phi(x)] \neq \mathbb{E}_{\theta'}[\phi(x)]$. Partition $\mathcal{C}$ according to $\mathcal{C}_{ns}$ so that $\mathcal{C} \equiv \mathcal{C}_{ns} \cup (\mathcal{C}\backslash\mathcal{C}_{ns})$. Take any test $\phi' \in \mathcal{C}_{ns}$ that satisfies (B.6). Since ϕ' is an element of $\mathcal{C}_{ns}$ and Θ_0 contains its boundary (as it is closed), there exists $\theta \in \text{Bd}\Theta_0$ such that $\Delta_{\phi',\phi^*}(\theta) \equiv \mathbb{E}_\theta[\phi'(X)] - \mathbb{E}_\theta[\phi^*(X)] < 0$. Because $\Delta_{\phi',\phi^*}(\theta) < 0$ and the function $\Delta_{\phi',\phi^*}(\cdot)$ is continuous at θ, there exists an open neighborhood $\tau_\theta \in \mathcal{T}$ such that $\Delta_{\phi',\phi^*}(\theta) < 0$ for all $\theta \in \tau_\theta$. Since θ is an element of $\text{Bd}\,\Theta_0$, then $\tau_\theta \cap \Theta_1 \neq \emptyset$. The latter implies there exists $\theta_1 \in \Theta_1$ such that $\Delta_{\phi',\phi^*_\alpha}(\theta_1) = \mathbb{E}_{\theta_1}[\phi'(X)] - \mathbb{E}_{\theta_1}[\phi^*_\alpha(X)] < 0$. Therefore, equation $(B.6)$ and $(B.7)$ cannot hold. We conclude there is no test $\phi' \in \mathcal{C}_{ns}$ that satisfies (B.6) and (B.7).
Since $\mathcal{C}_{ns}$ partitions $\mathcal{C}$, a test $\phi' \in \mathcal{C}$ that satisfies (B.6) and (B.4) must be an element of $\mathcal{C}\backslash\mathcal{C}_{ns}$ (as $\phi' \notin \mathcal{C}_{ns}$). Equation (B.6) implies ϕ' is α'-similar on the boundary with $\alpha' \leq \alpha$. Two cases follow: $\alpha' < \alpha$ or $\alpha' = \alpha$. In the first case, the argument in the previous paragraph implies that ϕ' will violate $(B.4)$. We conclude that any test that satisfies (B.6) and (B.4) must be α-similar on $\text{Bd}\Theta_0 = \Theta_0$.

Step 2: We have shown that $\phi^* \in M(w)$ implies ϕ^* is admissible in $\mathcal{C}(\alpha\text{-}s)$. We want to show that there is no nonsimilar α-level test such that $R(\phi',\theta) \leq R(\phi^*,\theta)$ with strict inequality for some $\theta \in \Theta$. By Step 1 any test $\phi' \in C$ that satisfies (B.6) and (B.4) must be α-similar on $\text{Bd}\Theta_0$. Therefore, we conclude ϕ^* is admissible in $\mathcal{C}$ as only α-similar tests can dominate ϕ^* and ϕ^* has been shown to be admissible in $\mathcal{C}(\alpha\text{-}s)$. *Q.E.D.*

B.2. *Lemmas for Result 1*

B.2.1. *Proof of Lemma 3*

Proof of Lemma 3: The Gaussian likelihood for (S,T) given parameters (ρ,ϕ,ω) is

$$f(S,T;\rho,\phi,\omega) = c_1 \exp\Big(-\frac{1}{2}([S',T']' - \rho(\phi\otimes\omega))'([S',T']' - \rho(\phi\otimes\omega))\Big),$$

where c_1 is a positive constant. Algebra shows that

$$\int_{\mathcal{S}^{k-1}} f(S,T;\rho,\phi,\omega) d\lambda_{\mathcal{S}^{k-1}}(\omega)$$

equals

$$a_1(Q)\exp\Big(-\rho^2/2\Big)\int_{\mathcal{S}^{k-1}} \exp\Big(([S,T]\phi)'\rho\omega\Big) d\lambda_{\mathcal{S}^{k-1}}(\omega),$$

where

$$a_1(Q) \equiv c_2 \exp\Big(-\frac{1}{2}[S'S + T'T]\Big)$$

and c_2 is a positive constant.

The density of ρ is given by:

$$m(\rho) \equiv \frac{1}{2^{k/2}\Gamma(k/2)}(\rho^2)^{(k/2)-1} e^{-(\rho^2/2)} 2\rho$$

The integral of interest thus equals

$$
\begin{aligned}
&= a_1(Q)\int_{\mathcal{S}^{k-1}}\Big(\int_{\mathbb{R}_+}\exp\Big(\Big[(S,T)\rho\phi\Big]'\boldsymbol{\omega}\Big)m(\rho)\exp(-\rho^2/2)d\rho\Big)d\lambda_{\mathcal{S}^{k-1}}(\boldsymbol{\omega}),\\
&= a_2(Q)\int_{\mathcal{S}^{k-1}}\Big(\int_{\mathbb{R}_+}\exp\Big(\Big[(S,T)\rho\phi\Big]'\boldsymbol{\omega}\Big)\exp(-\rho^2)\rho^{k-1}dp\Big)d\lambda_{\mathcal{S}^{k-1}}(\boldsymbol{\omega}),\\
&= a_2(Q)\int_{\mathcal{S}^{k-1}}\Big(\int_{\mathbb{R}_+}\exp\Big(\Big[(S,T)\phi\Big]'\rho\omega\Big)\exp(-(\rho\boldsymbol{\omega})'(\rho\boldsymbol{\omega}))\rho^{k-1}dp\Big)d\lambda_{\mathcal{S}^{k-1}}(\boldsymbol{\omega})
\end{aligned}
$$

where the last line follows from $\boldsymbol{\omega}'\boldsymbol{\omega}=1$ and $a_2(Q)=a_1(Q)2/(2^{k/2}\Gamma(k/2))$. Theorem 5.2.2, p. 86 in Stroock (1999) implies the last equation above equals

$$
\begin{aligned}
& a_2(Q)\int_{\mathbb{R}^K}\exp\Big(\Big[(S,T)\phi\Big]'x\Big)\exp(-x'x)dx,\\
& \text{(by applying Theorem 5.2.2 to the function } \exp\Big(\Big[(S,T)\phi\Big]'x\Big)\exp(-x'x))\\
&= a_3(Q)\exp\Big(\frac{1}{4}\phi'Q\phi\Big),\ Q\equiv[S,T]'[S,T],
\end{aligned}
$$

where the last line follows by definition of the moment generating function of a k-dimensional multivariate normal evaluated at $(S,T)\phi$. The constant $a_3(Q)$ equals

$$
\begin{aligned}
a_3(Q) &\equiv (2\pi)^{k/2}a_2(Q),\\
&= (2\pi)^{k/2}2/(2^{k/2}\Gamma(k/2))a_1(Q),\\
&= (2\pi)^{k/2}2/(2^{k/2}\Gamma(k/2))c_2\exp\Big(-\frac{1}{2}[S'S+T'T]\Big).
\end{aligned}
$$

B.2.2. *Proof of Lemma 4*

Proof of Lemma 4: Part ii) of exercise 5.2.4 in p. 87 of Stroock (1999) implies that

$$
\int_{\mathcal{S}^1}\exp\Big(\frac{1}{4}\phi'Q\phi\Big)d\lambda_{\mathcal{S}^1}(\phi)=\frac{1}{2\pi}\int_0^{2\pi}\exp\Big(\frac{1}{4}[\cos(\theta),\sin(\theta)]Q[\cos(\theta),\sin(\theta)]'\Big)d\theta.
$$

Let $L\equiv S'S-\zeta_{min}$, where ζ_{min} is the smallest eigenvalue of Q as defined in the statement of Lemma 4. Note that L is the Likelihood Ratio Statistic as defined in Andrews, Moreira, and Stock (2006) p. 722. The eigenvector associated with the largest eigenvalue of the matrix Q equals:

$$
e_{max}\equiv(L,S'T)'/\sqrt{L^2+(S'T)^2}.
$$

Define $\widehat{\theta}\in[0,2\pi]$ implicitly by the following equation:

$$
[\cos(\widehat{\theta}),\sin(\widehat{\theta})]'=e_{max}
$$

Therefore,

$$
P\equiv\begin{pmatrix}\cos(\widehat{\theta}) & \sin(\widehat{\theta})\\ \sin(\widehat{\theta}) & -\cos(\widehat{\theta})\end{pmatrix}
$$

yields the spectral decomposition of the matrix Q; that is:

$$
P\begin{pmatrix}\zeta_{max} & 0\\ 0 & \zeta_{min}\end{pmatrix}P'=Q.
$$

Note that for any θ:

$$
P'\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}=\begin{pmatrix}\cos(\widehat{\theta})\cos(\theta)+\sin(\widehat{\theta})\sin(\theta)\\ \sin(\widehat{\theta})\cos(\theta)-\cos(\widehat{\theta})\sin(\theta)\end{pmatrix}=\begin{pmatrix}\cos(\widehat{\theta}-\theta)\\ \sin(\widehat{\theta}-\theta)\end{pmatrix}.
$$

Therefore:

$$\begin{aligned}
\int_{\mathcal{S}^1} \exp\Big(\frac{1}{4}\phi' Q\phi\Big) d\lambda_{\mathcal{S}^1}(\phi) &= \frac{1}{2\pi}\int_0^{2\pi} \exp\Big(\frac{1}{4}\Big[\zeta_{max}\cos^2(\widehat{\theta}-\theta)+\zeta_{min}\sin^2(\widehat{\theta}-\theta)\Big]\Big)d\theta,\\
&= \frac{1}{2\pi}\int_{\widehat{\theta}-2\pi}^{\widehat{\theta}} \exp\Big(\frac{1}{4}\Big[\zeta_{max}\cos^2(\theta)+\zeta_{min}\sin^2(\theta)\Big]\Big)d\theta,\\
&\quad \text{(where he have changed the integration variable)}\\
&= \exp\Big(\frac{1}{4}\zeta_{min}\Big)\frac{1)}{2\pi}\int_{\widehat{\theta}-2\pi}^{\widehat{\theta}} \exp\Big(\frac{1}{4}\Big[(\zeta_{max}-\zeta_{min})\cos^2(\theta)\Big]\Big)d\theta,\\
&\quad (\text{as}\sin^2(\theta)+\cos^2(\theta)=1)\\
&= \frac{1}{2\pi}\exp\Big(\frac{1}{4}\zeta_{min}\Big)\exp\Big(\frac{1}{8}(\zeta_{max}-\zeta_{min})\Big)\\
&\quad \int_{\widehat{\theta}-2\pi}^{\widehat{\theta}} \exp\Big(\frac{1}{8}(\zeta_{max}-\zeta_{min})\cos(2\theta)\Big)d\theta,\\
&\quad (\text{as }\cos^2(\theta)=(1+\cos(2\theta))/2)\\
&= \frac{1}{4\pi}\exp\Big(\frac{1}{4}\zeta_{min}\Big)\exp\Big(\frac{1}{8}(\zeta_{max}-\zeta_{min})\Big)\\
x &\quad \int_{2(\widehat{\theta}-2\pi)}^{2\widehat{\theta}} \exp\Big(\kappa(Q)\cos(\theta)\Big)d\theta,\\
&\quad \text{(where we have used the change of variable } \tilde{\theta}=2\theta)\\
&\quad (\kappa(Q)\equiv\frac{1}{8}(\zeta_{max}-\zeta_{min}))\\
&= \exp\Big(\frac{1}{4}\zeta_{min}\Big)\exp\Big(\frac{1}{8}(\zeta_{max}-\zeta_{min})\Big)\\
&\quad \frac{1}{2\pi}\int_0^{2\pi} \exp\Big(\kappa(Q)\cos(u)\Big)du,\\
&\quad \text{(where we have used the change of variable } u=(\widehat{\theta})-(\theta/2).
\end{aligned}$$

Using the definition of the Von-Mises distribution and equation 3.5.18 in p. 36 of Mardia and Jupp (2000) it follows that:

$$\begin{aligned}
\int_{\mathcal{S}^1} \exp\Big(\frac{1}{4}\phi' Q\phi\Big) d\lambda_{\mathcal{S}^1}(\phi) &= \exp\Big(\frac{1}{4}\zeta_{min}\Big)\exp\Big(\frac{1}{8}(\zeta_{max}-\zeta_{min})\Big)I_0\Big(\kappa(Q)\Big),\\
&= \exp\Big(\frac{1}{8}\Big(\zeta_{max}+\zeta_{min}\Big)\Big)I_0\Big(\frac{1}{8}(\zeta_{max}-\zeta_{min})\Big),
\end{aligned}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind, defined in Abramowitz and Stegun (1964), Section 9.6, p. 375.

B.3. *Additional results related to Theorem 1*

B.3.1. *Corollary to Theorem 1: sequences of WAP-similar tests*

COROLLARY 1: *Suppose that Θ is compact. Let ϕ be an admissible, α-similar test. Let $\to^*$ denote convergence in the weak* topology as defined in B.1.1. Under Assumption F0 there exists a sequence of Borel probability measures w_n on Θ, a weight function w^*, and an α-similar test ϕ^* such that:*

$$w_n \xrightarrow{d} w^*, \quad \phi_{WAP}^{w_n,\alpha} \to^* \phi^*.$$

Moreover, if the sequence $\{w_n\}_{n\in\mathbb{N}}$ has a common σ-finite dominating measure, P, and the sequence of corresponding Radon-Nikodym derivaties $\{f_n\}_{n\in\mathbb{N}}$ admits a function g such that:

$$|f_n(\theta)| \leq g(\theta) \quad \textit{and} \quad \int_\Theta |g| dP < \infty,$$

then $WAP(\phi^*, w^*) = WAP(\phi, w^*)$. This means that for any admissible, α-similar test ϕ there is a sequence of weights (with limit w^*) such that the test ϕ is WAP-equivalent to the properly defined limit of w_n-WAP α-similar tests.

PROOF: We will break the proof of the corollary into two steps.

STEP 1 (construction of w^* and ϕ^*): Take any sequence of real numbers $\{\epsilon_k\}_{k\in\mathbb{N}}$ such that $\epsilon_k \to 0$. Since ϕ is admissible and similar, Theorem 1 implies the existence of a sequence of Borel probability measures $\{w_k\}$ satisfying

$$\text{WAP}(\phi_{\text{WAP}}^{w_k,\alpha}, w_k) \geq \text{WAP}(\phi, w_k) \geq \text{WAP}(\phi_{\text{WAP}}^{w_k,\alpha}, w_k) - \epsilon_k. \tag{B.8}$$

The sequence $\{\text{WAP}(\phi_{\text{WAP}}^{w_k,\alpha}, w_k)\}_{k\in\mathbb{N}}$ takes its values on the $[0,1]$ interval. Hence, there exists a subsequence $\{w_{k_l}\}_{l\in\mathbb{N}}$ along which:

$$\text{WAP}(\phi_{\text{WAP}}^{w_{k_l},\alpha}, w_{k_l}) \to \text{WAP}^*,$$

where WAP* is some number in the $[0,1]$ interval. Moreover, Equation (B.8) and $\epsilon_k \to 0$ then imply

$$\text{WAP}(\phi, w_{k_l}) \to \text{WAP}^*.$$

It is well known that if $\Theta \subseteq \mathbb{R}^p$ is endowed with its standard metric, then the space of Borel probability measures on Θ is sequentially compact—relative to the topology induced by the Prokhorov metric, which metrizes weak convergence (see Proposition 4.4 in Chapter F of Ok (2019)). Therefore, it is possible to extract a further subsequence $\{w_{k_{l_m}}\}_{m\in\mathbb{N}}$ such that:

$$w_{k_{l_m}} \xrightarrow{d} w^*.$$

In proving Theorem 1, we have showed that the space of α-similar tests is compact relative to the weak* topology, and also metrizable. The latter implies that the space of α-similar tests is sequentially compact. Consequently, we can extract a further subsequence $\{w_{k_{l_{m_n}}}\}_{n\in\mathbb{N}}$ along which:

$$\phi_{\text{WAP}}^{w_{k_{l_{m_n}}},\alpha} \to^* \phi^*,$$

as n tends to infinity. Consider thus the sequence of weights $\{w_n\}_{n\in\mathbb{N}}$ with n-th element defined as

$$w_n \equiv w_{k_{l_{m_n}}}.$$

By construction, under this sequence

$$w_n \xrightarrow{d} w^*, \quad \phi_{\text{WAP}}^{w_n,\alpha} \to^* \phi^*, \quad \text{WAP}(\phi_{\text{WAP}}^{w_n,\alpha}, w_n) \to \text{WAP}^*, \quad \text{WAP}(\phi, w_n) \to \text{WAP}^*.$$

with ϕ^* α-similar.

Step 2 (ϕ and ϕ^* are WAP-equivalent under w^*): We will now show that $\text{WAP}(\phi, w^*) = \text{WAP}^* = \text{WAP}(\phi^*, w^*)$. So that the original admissible, similar test ϕ and the limiting test ϕ^* are WAP equivalent under the limiting weight w^*.

First, Assumption F0 and the definition of weak convergence of probability measures implies

$$\text{WAP}(\phi, w_n) \equiv \int_\Theta R(\phi;\theta) dw_n \to \int_\Theta R(\phi;\theta) dw \equiv \text{WAP}(\phi, w^*).$$

(as $R(\phi,\cdot)$ is bounded and, by Assumption F0, continuous). Since—by construction of w_n—$\text{WAP}(\phi, w_n) \to \text{WAP}^*$, then $\text{WAP}(\phi, w^*) = \text{WAP}^*$.

Second, $\text{WAP}^* = \text{WAP}(\phi^*, w^*)$. To establish such a relation, it is sufficient to show

$$\text{WAP}(\phi^{w_n,\alpha}_{\text{WAP}}, w_n) \to \text{WAP}(\phi^*, w^*),$$

as, by construction, $\text{WAP}(\phi^{w_n,\alpha}_{\text{WAP}}, w_n) \to \text{WAP}^*$. Note that:

$$\begin{aligned}
\text{WAP}(\phi^{w_n,\alpha}_{\text{WAP}}, w_n) - \text{WAP}(\phi^*, w^*) &= \text{WAP}(\phi^{w_n,\alpha}_{\text{WAP}}, w_n) - \text{WAP}(\phi^*, w_n) \\
&+ \text{WAP}(\phi^*, w_n) - \text{WAP}(\phi^*, w^*) \\
&= \text{WAP}(\phi^{w_n,\alpha}_{\text{WAP}}, w_n) - \text{WAP}(\phi^*, w_n) + o(1) \\
&\quad \text{(By Assumption F0 and the definition of weak convergence)} \\
&= \int_\Theta (R(\phi^{w_n,\alpha}_{\text{WAP}};\theta) - R(\phi^*,\theta)) dw_n + o(1) \\
&= \int_\Theta (R(\phi^{w_n,\alpha}_{\text{WAP}};\theta) - R(\phi^*,\theta)) f_n(\theta) dP + o(1) \\
&\quad \text{(by the assumption about the existence of Radon-Nikodym} \\
&\quad \text{derivatives w.r.t. } P \text{ and equation (32.5) in Billingsley (1995))} \\
&= o(1)
\end{aligned}$$

To establish the last equality, define the sequence of functions:

$$h_n(\theta) \equiv (R(\phi^{w_n,\alpha}_{\text{WAP}};\theta) - R(\phi^*,\theta)) f_n(\theta)$$

and note that $|h_n(\theta)| \leq |(R(\phi^{w_n,\alpha}_{\text{WAP}};\theta) - R(\phi^*,\theta))| g(\theta)$. Since $\phi^{w_n,\alpha}_{\text{WAP}} \to^* \phi^*$ it then follows that $R(\phi^{w_n,\alpha}_{\text{WAP}};\theta) \to R(\phi^*,\theta))$ for every θ and consequently:

$$h_n(\theta) \to 0 \quad \forall\, \theta \in \Theta.$$

An application of the dominated convergence theorem yields

$$\text{WAP}(\phi^{w_n,\alpha}_{\text{WAP}}, w_n) - \text{WAP}(\phi^*, w^*) \to 0.$$

Consequently,

$$\text{WAP}(\phi^*, w^*) = \text{WAP}^*.$$

Therefore ϕ and ϕ^* are WAP-equivalent under w^*. That is, $\text{WAP}(\phi, w^*) = \text{WAP}^* = \text{WAP}(\phi^*, w^*)$.

Q.E.D.

Comment: The essentially complete class theorem (ECCT) invoked in the proof of Theorem 1—based on Theorems 2.9.2 and 2.10.3 in Ferguson (1967)—characterizes admissible tests as extended Bayes tests. There are other versions of the ECCT that characterize admissible tests in terms of 'limits of Bayes procedures'. For example, Theorem 4A.10 in Brown (1986) shows that the closure (in weak* topology) of the set of Bayes procedures for priors concentrated on finite subsets of Θ

constitutes—under some assumptions on the action space, the loss function, and the statistical model—an essentially complete class. Note that if we're able to verify such a theorem in our set-up, then for every admissible, α-similar test ϕ there would be a test ϕ^*—on the closure of Bayes procedures—for which $R(\phi^*, \theta) = R(\phi, \theta)$ for every θ. This, by definition of closure, would imply the existence of a sequence of weights w_n (concentrated on finite subsets of Θ) such that:

$$\phi_{WAP}^{w_n,\alpha} \to^* \phi^*,$$

and consequently, by the definition of weak* convergence,

$$R(\phi_{WAP}^{w_n,\alpha}, \theta) \to R(\phi^*, \theta) = R(\phi, \theta), \quad \forall\, \theta \in \Theta.$$

This is a stronger result than the one obtained in Corollary 1. To the best of my knowledge, the stronger version of the complete class theorems seem to require the convexity of the action space as well as strict convexity of the loss function (see for example Theorem 7.15 in Lehmann and Casella (1998)).

B.3.2. WAP-*similar tests with a boundedly complete, null-sufficient statistic.*

Preliminaries: This section generalizes a well-known observation in the IV literature: maximizing constrained average power is straightforward whenever there is a boundedly-complete, null-sufficient statistic. Consider the following assumptions.

Assumption F1 (Null Sufficiency): There is a partition of the data $X = (x_1, x_2)$ such that the conditional density of x_1 given x_2 in the statistical model $f(x_1, x_2; \theta)$ satisfies:

$$f(x_1|x_2; \beta_0) \equiv f(x_1|x_2; \theta) = f(x_1|x_2; \theta') \quad \forall\, \theta, \theta' \in \Theta_0. \tag{B.9}$$

The statistic x_2 arising from such partition of the data will be called a *null-sufficient statistic*.

It is well known that a null-sufficient statistic can be used to control the *null* rejection probability of a test in a two-sided problem with a nuisance parameter [Ferguson (1967), Moreira (2003), Andrews et al. (2006), Lehmann and Romano (2005)].

Let $h(x_2; \theta)$ denote the marginal density of the null-sufficient statistic x_2 based on the statistical model $f(x_1, x_2; \beta, \Pi)$.

Assumption F2: (Bounded Completeness): For any bounded measurable function $m : \mathbf{X}_2 \to \mathbb{R}$, the marginal densities of the null sufficient statistic are such that:

$$\int m(x_2) h(x_2; \theta) dx_2 = 0, \quad \forall\, \theta \in \Theta_0 \implies m(x_2) = 0,$$

except, perhaps, in a set that has zero measure under every element of the collection $\{h(\cdot, \theta)\}_{\theta \in \Theta_0}$.[1]

Theorem 4.3.1 in Lehmann and Romano (2005) provides a sufficient condition to guarantee that a family of distributions is complete, and thus, boundedly complete. In the IV example studied in this paper, it will be sufficient to show that the set Θ_0 contains a rectangle of the same dimension as the null-sufficient statistic.

Bounded completeness will be used to show that all similar tests must be "conditionally" similar. This is a well-known result in the theory of statistical hypothesis testing. See Theorem 4.3.2 in Lehmann and Romano (2005).

[1] See Lehmann and Romano (2005) p. 115 for the definition of bounded completeness.

Description: Lemma 5 will show that under assumptions F1, F2 the test that rejects whenever

$$\phi^*(x_1, x_2) \equiv f^*_{w_1}(x_1, x_2)/f(x_1|x_2; \beta_0) > c(x_2; \alpha), \tag{B.10}$$

is an element of

$$M(w_1) \equiv \arg \min_{\phi \in \mathcal{C}(\alpha\text{-}s)} \int_{\text{Int } \Theta_1} R(\phi, \theta) dw_1(\theta)$$

provided $c(x_2; \alpha)$ is the $1 - \alpha$ quantile of $z(X_1, x_2)$ with $X_1 \sim f(x_1|x_2; \beta_0)$. This is a well-known result and we reproduce it for the sake of completeness.

Relevance of Lemma 5: This implies that the tests in (B.10) are constrained weighted average power maximizers. This property has been discussed in Andrews, Moreira, and Stock (2004), Chernozhukov et al. (2009). Lemma 5 combined with Lemma 2 implies that the tests in (B.10) are admissible in the class of all tests.

Lemma 5: *Let ϕ^* be defined as in (B.10) and let $c(\cdot; \alpha)$ be measurable. Under Assumptions F1-F2, $\phi^* \in M(w_1)$; that is, ϕ^* minimizes average risk inside the class of α-similar tests.*

Proof: Throughout this proof we assume that $\mathbf{X} = \mathbf{X}_1 \times \mathbf{X}_2$. Fubini's theorem (L2a-Step 1) and Theorem 4.3.2 in Lehmann and Casella (1998) implies that $\phi^* \in M(w_1)$ if and only if ϕ^* solves the problem:

$$\min_{\phi \in \mathcal{C}} \int_X (1 - \phi(x)) f^*_1(x) dx$$

$$\int_{\mathbf{X}_1(x_2)} \phi(x_1, x_2) f(x_1|x_2) dx_1 = \alpha$$

except, perhaps, for x_2 that belong to a set of measure zero under the marginal density of $h(x_2, \theta)$ for all $\theta \in \Theta_0$. Re-write the objective function as

$$\max_{\phi \in \mathcal{C}} \int_X \phi(x) f^*_1(x) dx.$$

The product structure of $\mathbf{X}$ and the linearity of the integral allows a further expansion of the previous equation:

$$\max_{\phi \in \mathcal{C}} \int_{X_2} \Big(\int_{\mathbf{X}_1} \phi(x_1, x_2) f^*_1(x_1, x_2) dx_1 \Big) dx_2.$$

Note first that the Neyman Pearson Lemma in Ferguson (1967) p. 204 implies that for a fixed x_2 the WAP test $\phi^*(x_1, x_2)$ solves the problem

$$\max_{\phi \in \mathcal{C}} \int_{\mathbf{X}_1} \phi(x_1, x_2) f^*_1(x_1, x_2)\, dx_1$$

subject to

$$\int_{\mathbf{X}_1} \phi(x_1, x_2) f(x_1|x_2) dx_1 = \alpha.$$

except, perhaps, for x_2 that belong to a set of measure zero under every $h(x_2, \theta)$, $\theta \in \text{Bd}\Theta_0$. Hence, to show that $\phi^*(x_1, x_2) \in M(w_1)$ it only remains to prove that $\phi^*(x_1, x_2)$ is measurable. That is, $\phi^*(x_1, x_2) \in \mathcal{C}(\alpha\text{-s})$. Assumption F0 implies that $\phi^*(x_1, x_2)$ is continuous in x_1, for every x_2.

Furthermore, since $c(\cdot, \alpha)$ is measurable, then $\phi^*(x_1, x_2)$ is measurable in x_2, for every x_1. Therefore, $\phi^*(x_1, x_2)$ is a Carathéodory function, as defined in Aliprantis and Border (2006), p. 153. Since the sample space $\mathbf{X}$ is separable (by assumption) and metrizable (for it is a subset of a euclidean space), Lemma 4.5.1 in Aliprantis and Border (2006) p. 153 implies $\phi^* : \mathbf{X} \rightarrow [0,1]$ is measurable. *Q.E.D.*

B.3.3. *Continuity of the conditional critical value function*

Measurability is required for the proof of Lemma 5. This subsection provides two sufficient conditions that imply the continuity of $c(\cdot;\alpha)$ (and hence, its measurability).

Let $f(x;\theta)$ denote the statistical model. Consider the following auxiliary assumptions:

ASSUMPTION F3: There exists a function $g(\theta)$ such that:

$$f(x;\theta) \leq g(\theta) \quad \forall\, x,$$

and $\int_\theta g(\theta) dw(\theta) < \infty$.

ASSUMPTION F4: $f(x_1|x_2;\beta_0) > 0$ for every (x_1, x_2) and $f(x_1|x_2;\beta_0)$ is continuous in (x_1, x_2).

Assumptions F0, F3, F4 imply that $c(x_2;\alpha)$ is continuous.

PROOF: Note first that Assumption F0 implies that $f^*_w(x)$ is sequentially continuous in x. To see this, consider any sequence $x_n \rightarrow x$. Assumption F0 i) implies that $f(x_n;\theta) \rightarrow f(x;\theta)$ for almost every $\theta \in \Theta$. Since the weight function $w(\theta)$ is assumed to satisfy $f^*_w(x) < \infty$ for every x then:

$$\begin{aligned} \Big| f^*_w(x_n) - f^*_w(x) \Big| &= \Big| \int_\Theta f(x_n;\theta) dw(\theta) - \int_\Theta f(x;\theta) dw(\theta) \Big| \\ &\leq \int_\Theta \Big| f(x_n;\theta) - f(x;\theta) \Big| dw(\theta). \end{aligned}$$

By Assumption F3, the Dominated Convergence Theorem applies and we can conclude that

$$\int_\Theta \Big| f(x_n;\theta) - f(x;\theta) \Big| dw(\theta) \rightarrow 0.$$

Consequently, $f^*_w(x_n) \rightarrow f^*_w(x)$. Furthermore, Assumption F4 implies that the test statistic:

$$z(x_1, x_2) = f_w(x_1, x_2)/f(x_1|x_2;\beta_0)$$

is continuous in (x_1, x_2).

Let $x_{2,n} \rightarrow x_2$ and let $X_{1,n} \sim f(x_1|x_{2,n};\beta_0)$. Consider the sequence of random variables.

$$z(X_{1,n}, x_{2,n}).$$

By Scheffe's theorem and the continuity of $f(x_1|x_2;\beta_0)$ at x_2, $X_{1,n} \xrightarrow{d} X \sim f(x_1|x_2;\beta_0)$. Therefore, the random vector $(X_{1,n}, x_{2,n}) \xrightarrow{d} (X, x_2)$. The continuous mapping theorem implies that

$$z(X_{1,n}, x_{2,n}) \xrightarrow{d} z(X_1, x_2).$$

Lemma 21.2 in Van der Vaart (2000) implies $c(x_{2,n};\alpha) \rightarrow c(x_2;\alpha)$ for any sequence $x_{2,n} \rightarrow x_2$. Hence, the critical value function is continuous and, consequently, measurable.

B.4. *Additional results related to Result 1*

B.4.1. *Asymptotic validity of the test in Result 1*

The test in Result 1 was derived under the assumption that the rotated reduced-form OLS estimators $(S_n', T_n')'$ have the exact distribution:

$$Q^n_{\beta,\Pi,\Sigma} \equiv \mathcal{N}_{2k}\left(\begin{matrix} ([b_0' \otimes \mathbb{I}_k)\Sigma(b_0 \otimes \mathbb{I}_k)]^{-1/2}(\beta-\beta_0)\sqrt{n}\Pi \\ [(a_0' \otimes \mathbb{I}_k)\Sigma^{-1}(a_0 \otimes \mathbb{I}_k)]^{-1/2}(a_0' \otimes \mathbb{I}_k)\Sigma^{-1}(a \otimes \mathbb{I}_k)\sqrt{n}\Pi \end{matrix}, \mathbb{I}_{2k} \right),$$

where Σ is of the form $\Psi \otimes \Phi$. In any finite sample, however, the law of $(S_n', T_n')'$ is a function of (β, Π), the sample size, and the joint distribution between the instrumental variables and reduced-form residuals, denoted F. In fact, one can write:

$$\begin{pmatrix} ([b_0' \otimes \mathbb{I}_k)\widehat{\Sigma}(b_0 \otimes \mathbb{I}_k)]^{-1/2}(b_0' \otimes \mathbb{I}_k) \\ [(a_0' \otimes \mathbb{I}_k)\widehat{\Sigma}^{-1}(a_0 \otimes \mathbb{I}_k)]^{-1/2}(a_0' \otimes \mathbb{I}_k)\widehat{\Sigma}^{-1} \end{pmatrix} \sqrt{n}\widehat{\gamma}_n \sim P^n_{\beta,\Pi,F},$$

where $\widehat{\Sigma}$ is an estimator of the variance of $\sqrt{n}\widehat{\gamma}_n$. This variance depends on F and such dependence is denoted $\Sigma(F)$. The estimator $\widehat{\Sigma}$ need not have the Kronecker form, even when $\Sigma(F)$ does.

If one assumes that for n large enough the distributions P^n and Q^n are 'close' to each other (under the null), then one would expect the rate of Type I error computed under P^n to be close to that obtained under Q^n.

Preliminaries: We introduce some additional notation in order to establish the asymptotic validity of the test in Result 1.

1. *Bounded Lipschitz Distance*: Let $d_{\mathrm{BL}}(P,Q) = \sup_{h \in \mathrm{BL}_1} |\mathbb{E}_P[h(X)] - \mathbb{E}_Q[h(X)]|$ denote the Bounded Lipschitz distance between any pair of probability measures P and Q. For definitions and notation, see p. 73, Section 1.12 of Van der Vaart and Wellner (1996). Note also that the Bounded Lipschitz metric is equivalent to the 'β' metric between Borel probability measures defined in p. 394 of Dudley (2002).

2. *δ-Expansion of a set A:* For any $\delta > 0$ let A^δ denote the δ-expansion of the set $A \subseteq \mathbb{R}^m$. This is $A^\delta = \{y \in \mathbb{R}^m \mid d(x,y) \leq \delta \text{ for some } x \in A\}$.

3. *A bound on the distance between probability measures:* One can show that for any measurable set A and any $\delta > 0$:

$$-Q((A^c)^\delta \backslash A^c) - \frac{1}{\delta} d_{\mathrm{BL}_1}(P,Q) \leq P(A) - Q(A) \leq \frac{1}{\delta} d_{\mathrm{BL}_1}(P,Q) + Q(A^\delta \backslash A), \tag{B.11}$$

where A^c is the complement of $A \subseteq \mathbb{R}^m$. We use the right-hand side of this inequality to establish the main result.

Assumption L0: Suppose that the class of distributions $\mathcal{F}$ is such that:

$$\lim_{n\to\infty} \sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} d_{\mathrm{BL}}\left(P^n_{\beta_0,\Pi,F}, Q^n_{\beta_0,\Pi,\Sigma(F)}\right) \to 0.$$

That is, the Bounded Lipschitz distance between the measures $P^n_{\beta,\Pi,F}$ and $Q^n_{\beta,\Pi,\Sigma(F)}$ coverges to zero as the sample size grows large (uniformly over Π and F).

Asymptotic Validity of the Test in Result 1: If Assumption L0 holds and there are constants $0 < \underline{\lambda} < \overline{\lambda} < \infty$ such that the eigenvalues of $\Sigma(F)$ belong to an interval $[\underline{\lambda}, \overline{\lambda}]$ for any $F \in \mathcal{F}$, then:

$$\limsup_{n\to\infty} \sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} P^n_{\beta_0,\Pi,F}\left(z_{\mathrm{WAP}}(S,T) - c_{\mathrm{WAP}}(T,\alpha) > 0\right) \leq \alpha.$$

This means that the rate of Type I error of the test in Result 1 is uniformly controlled over $(\Pi, \mathcal{F}) \in \mathbb{R}^k \times \mathcal{F}$.
Consider the test statistic

$$z(S,T) \equiv S'S - T'T + 8 \ln \left(I_0 \left[(1/8) \sqrt{(S'S - T'T)^2 + 4(S'T)^2} \right] \right)$$

and let $c(T;\alpha)$ denote its conditional critical value. We would like to show that if Assumption L0 holds over the class $\mathcal{F}$, then:

$$\text{(B.12)} \qquad \limsup_{n\to\infty} \sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} P^n_{\beta_0,\Pi,F}\left(z(S,T) - c(T;\alpha) \geq 0\right) \leq \alpha.$$

We establish the asymptotic validity of the test in Result 1 in six steps:

Step 0: Define

$$A \equiv \{(s',t')' \in \mathbb{R}^{2k} \mid z(s,t) - c(t;\alpha) \geq 0\}.$$

Note immediately that Equation (B.11) implies that for any sample size n and any $(\Pi, F) \in \mathbb{R}^k \times \mathcal{F}$:

$$\begin{aligned} P^n_{\beta_0,\Pi,\mathcal{F}}(A) &\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A) + \frac{1}{\delta} d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F}, Q^n_{\beta_0,\Pi,\Sigma(F)}) + Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta \backslash A), \\ &= \alpha + \frac{1}{\delta} d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F}, Q^n_{\beta_0,\Pi,\Sigma(F)}) + Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta \backslash A). \end{aligned}$$

where the last equality follows by the definition of the conditional critical value $c(T;\alpha)$. Thus, in order to establish (B.12), we need to show that

$$\frac{1}{\delta} d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F}, Q^n_{\beta_0,\Pi,\Sigma(F)}) + Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta \backslash A)$$

can be made arbitrary small, uniformly over the values of (Π, F). By the weak convergence assumption in Part 2 of Result 1, for any fixed δ there is $M_\epsilon(\delta) \in \mathbb{N}$ such that whenever $n \geq M_\epsilon(\delta)$ the term $\delta^{-1} d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F}, Q^n_{\beta_0,\Pi,\Sigma(F)})$ can be made smaller than ϵ. Thus, we only need to establish the following result.

Goal: For every $\epsilon > 0$ there is δ_ϵ and N_ϵ such that for all $n \geq N_\epsilon$

$$\sup_{(\Pi,F)\in\mathbb{R}^k\times F} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon} \backslash A) \leq \epsilon.$$

The proof of this result requires a series of intermediate steps. We exploit the fact that the test statistic $z(s,t)$ satisfies a Lipschitz condition whenever (s,t) is restricted to an appropriate set.

Step 1: (A bound on $Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta \backslash A)$): Define the sets

$$\text{(B.13)} \qquad B(\underline{b}_1, \overline{b}_2) \equiv \{(s',t')' \in \mathbb{R}^{2k} \mid s's \in [\underline{b}_1, \overline{b}_2]\ \},$$

$$\text{(B.14)} \qquad C(\underline{c}_1, \overline{c}_2) \equiv \{(s',t')' \in \mathbb{R}^{2k} \mid t't \in [\underline{c}_1, \overline{c}_2]\ \},$$

$$\text{(B.15)} \qquad D(\underline{d}_1, \overline{d}_2) \equiv \{(s',t')' \in \mathbb{R}^{2k} \mid (s't)^2/t't \in [\underline{d}_1, \overline{d}_2]\ \},$$

where $\underline{b}_1, \overline{b}_2, \underline{c}_1, \overline{c}_2, \underline{d}_1, \overline{d}_2$ are positive, finite constants. We want to study the behavior of $A^\delta \backslash A$ inside and outside the sets defined above. Note that for any n, Π, F and δ:

$$
\begin{aligned}
Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A) &= Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)])\\
&+ Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)]^c)\\
&\quad \text{(by the additivity property of probability measures)}\\
&\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)])\\
&+ Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap B^c(\underline{b}_1,\overline{b}_1)) + Q^n_{\beta_0,\Pi\Sigma(F)}(A^\delta\backslash A \cap C^c(\underline{c}_1,\overline{c}_2))\\
&+ Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap D(\underline{d}_1,\overline{d}_2))\\
&\quad \text{(where we have used Boole's inequality)}\\
&\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)])\\
&+ Q^n_{\beta_0,\Pi,\Sigma(F)}(B^c(\underline{b}_1,\overline{b}_1)) + Q^n_{\beta_0,\Pi\Sigma(F)}(D^c(\underline{d}_1,\overline{d}_2))\\
&+ Q^n_{\beta_0,\Pi\Sigma(F)}(A^\delta\backslash A \cap C^c(\underline{c}_1,\overline{c}_2))\\
&\quad \text{(by the monotonicity of probability measures)}\\
&\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)])\\
&+ \mathbb{P}(\chi^2_k \notin (\underline{b}_1,\overline{b}_2)) + \mathbb{P}(\chi^2_1 \notin (\overline{d}_1,\overline{d}_2))\\
&+ Q^n_{\beta_0,\Pi\Sigma(F)}(A^\delta\backslash A \cap C^c(\underline{c}_1,\overline{c}_2)),
\end{aligned}
$$

where the second to last line uses the fact that under any probability measure $Q^n_{\beta_0,\Pi,\Sigma(F)}$:[2]

$$S_n'S_n \overset{Q^n_{\beta_0,\Pi,F}}{\sim} \chi^2_k \text{ and } (S_n'T_n)^2/T_n'T_n \overset{Q^n_{\beta_0,\Pi,F}}{\sim} \chi^2_1.$$

<u>Main conclusion of Step 1:</u> We have shown that for any $\delta>0$ and any positive finite constants $\underline{b}_1,\overline{b}_2,\underline{c}_1,\overline{c}_2,\underline{d}_1,\overline{d}_2$:

$$
\begin{aligned}
\text{(B.16)}\qquad Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A) &\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)])\\
&+ \mathbb{P}(\chi^2_k \notin (\underline{b}_1,\overline{b}_2)) + \mathbb{P}(\chi^2_1 \notin (\overline{d}_1,\overline{d}_2))\\
&+ Q^n_{\beta_0,\Pi,\Sigma(F)}(A^\delta\backslash A \cap C^c(\underline{c}_1,\overline{c}_2)).
\end{aligned}
$$

We now argue that for an appropriate selection of constants, the test statistic $z(s,t)$ and its critical value $c(t;\alpha)$ satisfy a Lipschitz condition when restricted to the set $[B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)]$.

Step 2—Part a): (Lipschitz property of $z(s,t)$): We show that there exists a constant M_1—that only depends on the sets B,C,D—such that for any

$$(s_0',t_0')',(s_1',t_1')' \in \mathcal{K} \equiv [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)],$$

then:

$$|z(s_0,t_0)-z(s_1,t_1)| < M_1||(s_0',t_0')-(s_1',t_1')||.$$

To verify the Lipschitz property on $[B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1(\epsilon),\overline{c}_2(\epsilon))\cap D(\underline{d}_1,\overline{d}_2)]$, it is sufficient to show that, over this set, the derivative of $z(s,t)$ is continuous in its arguments. This observation, together with the fact that $[B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1(\epsilon),\overline{c}_2(\epsilon))\cap D(\underline{d}_1,\overline{d}_2)]$ is compact gives the desired result. Note that the partial derivative of $z(s,t)$ with respect to s is given by:

$$\text{(B.17)}\qquad z_s(s,t) = 2S + \frac{8I_1\left((1/8)\sqrt{(s's-t't)^2+4t't(s't)^2/t't}\right)}{I_0\left((1/8)\sqrt{(s's-t't)^2+4t't(s't)^2/t't}\right)}\frac{1}{8}\frac{4(s's-t't)s+8t't((s't)/t't)s}{2\sqrt{(s's-t't)^2+4t't((s't)^2/t't)}}$$

[2] We also use the fact that $\Sigma(F)$ is invertible for any element $F\in\mathcal{F}$.

$$\text{(B.18)} \qquad z_t(s,t) = 2t + \frac{8I_1\Big((1/8)\sqrt{(s's-t't)^2+4t't(s't)^2/t't}\Big)}{I_0\Big((1/8)\sqrt{(s's-t't)^2+4t't(s't)^2/t't}\Big)}\frac{1}{8}\frac{4(s's-t't)t+8t't((s't)/t't)t}{2\sqrt{(s's-t't)^2+4t't((s't)^2/t't)}}$$

where $I_v(\cdot)$ is the modified Bessel function of the first kind defined in Section 9.6, p. 374 of Abramowitz and Stegun (1964). The formulae above use the fact that the derivative of the modified Bessel function of the first kind of order 0, I_0, is the modified Bessel function of order 1, I_1; see formula 9.6.27 in p. 376 of Abramowitz and Stegun (1964). The continuity of the derivatives and the fact that:

$$\sqrt{(s's-t't)^2+4t't(s't)^2/t't}$$

is bounded away from zero over the set

$$[B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)]$$

implies that the Lipschitz condition holds.

Step 2—Part b): (Lipschitz property of $c(t;\alpha)$): Part a) showed that for any selection of constants (b,c,d) the test statistic $z(s,t)$ satisfies the Lipschitz condition when restricted to $\mathcal{K}$. We now introduce a parameter γ and show that for any given $\gamma>0$, $\epsilon>0$ and any pair of constants $\underline{c}_1$, $\overline{c}_2$, one can find $\underline{b}_1^*$, $\overline{b}_2^*$, $\underline{d}_1^*$, $\overline{d}_2^*$, and M_2—that depend on $\underline{c}_1$, $\overline{c}_2$, γ and ϵ—such that for any t_0, t_1 satisfying:

$$(s_0,t_0),(s_1,t_1)\in\mathcal{K}^*\equiv[B(\underline{b}_1^*,\overline{b}_1^*)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1^*,\overline{d}_2^*)], \quad \text{for some } s_0,s_1\in\mathbb{R}^k$$

the critical value function satisfies a Lipschitz-type condition:

$$|c(t_0;\alpha)-c(t_1;\alpha)|<M_2||t_0-t_1||+\gamma/2,$$

and:

$$\mathbb{P}(\chi_k^2\notin(\underline{b}_1^*,\overline{b}_1^*)+\mathbb{P}(\chi_1^2\notin(\underline{d}_1^*,\overline{d}_2^*))<\epsilon/3.$$

To show this, note that for any constant $z\in\mathbb{R}$ and $(\Pi,\mathcal{F})\in\mathbb{R}^k\times\mathcal{F}$:

$$\begin{aligned}
Q^n_{\beta_0,\Pi,\Sigma(F)}\left(z(s,t_0)\le z\mid t=t_0\right) &= \mathbb{P}\left(z(S,t_0)\le z\right),\ S\sim\mathcal{N}(0,\mathbb{I}_k)\\
&= \mathbb{P}\left(z(S,t_0)-z(S,t_1)+z(S,t_1)\le z\right)\\
&= \mathbb{P}\left(z(S,t_1)\le z-(z(S,t_0)-z(S,t_1)) \text{ and } (S,t_0),(S,t_1)\in\mathcal{K}\right)\\
&+ \mathbb{P}\left(z(S,t_1)\le z-(z(S,t_0)-z(S,t_1)) \text{ and } (S,t_0),(S,t_1)\in\mathcal{K}^c\right).
\end{aligned}$$

where $\mathcal{K}\equiv[B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)]$. Since $z(s,t)$ satisfies the Lipschitz condition in $\mathcal{K}$ with constant $M_1(\mathcal{K})$ it follows that:

$$\begin{aligned}
\mathbb{P}\Big(z(S,t_1)\le z-M_1(\mathcal{K})||t_0-t_1||\Big) &\le \mathbb{P}\Big(z(S,t_1)\le z-M_1(\mathcal{K})||t_0-t_1|| \text{ and } (S,t_0),(S,t_1)\in\mathcal{K}\Big)\\
&+ \mathbb{P}\Big(z(S,t_1)\le z-M_1(\mathcal{K})||t_0-t_1|| \text{ and } (S,t_0),(S,t_1)\in\mathcal{K}^c\Big)\\
&\le \mathbb{P}\Big(z(S,t_1)\le z-(z(S,t_0)-z(S,t_1)) \text{ and } (S,t_0),(S,t_1)\in\mathcal{K}\Big)\\
&+ \mathbb{P}((S,t_0),(S,t_1)\in\mathcal{K}^c),
\end{aligned}$$

which implies that:

$$\mathbb{P}\Big(z(S,t_1)\le z-M_1(\mathcal{K})||t_0-t_1||\Big)-\mathbb{P}((S,t_0),(S,t_1)\in\mathcal{K}^c)$$

is less than or equal to

$$\mathbb{P}\Big(z(S,t_1)\le z-(z(S,t_0)-z(S,t_1)) \text{ and } (S,t_0),(S,t_1)\in\mathcal{K}\Big).$$

Note also that:

$$\mathbb{P}\Big(z(S,t_1) \leq z - (z(S,t_0) - z(S,t_1)) \text{ and } (S,t_0),(S,t_1) \in \mathcal{K}\Big) \leq \mathbb{P}\Big(z(S,t_1) \leq z + M_1(\mathcal{K})||t_0 - t_1||\Big).$$

Note now that for any $t \in \mathbb{R}^k$, the critical value function is continuous in α. Therefore, there exists a positive constant, $\eta_\gamma(\underline{c}_1, \overline{c}_2) > 0$, such that for any t such that $t't \in [\underline{c}_1, \overline{c}_2]$:

$$|c(t;\alpha + \eta_\gamma(\underline{c}_1,\overline{c}_2)) - c(t;\alpha)| \leq \gamma/2, \quad |c(t;\alpha - \eta_\gamma(\underline{c}_1,\overline{c}_2)) - c(t;\alpha)| \leq \gamma/2.$$

Since for any vectors $t_0, t_1 \neq \mathbf{0}_{k\times 1}$:

$$\mathbb{P}((S,t_0),(S,t_1) \in \mathcal{K}^c) \leq \mathbb{P}(\chi^2_k \notin (\underline{b}_1, \overline{b}_1) + \mathbb{P}(\chi^2_1 \notin (\underline{d}_1, \overline{d}_2)),$$

one can then choose $0 < \underline{b}_1^* \equiv \underline{b}_1(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon) < \overline{b}_1^* \equiv \overline{b}_1(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon) < \infty$, and $0 < \underline{d}_1^* \equiv \underline{d}_1(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon) < \overline{d}_2^* \equiv \overline{d}_2(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon) < \infty$ such that

$$\mathbb{P}(\chi^2_k \notin (\underline{b}_1^*, \overline{b}_1^*) + \mathbb{P}(\chi^2_1 \notin (\underline{d}_1^*, \overline{d}_2^*)) < \min\{\eta_\gamma(\underline{c}_1,\overline{c}_2), \epsilon/3\}$$

for any t_0, t_1. This implies that:

$$\text{(B.19)} \qquad \mathbb{P}\Big(z(S,t_1) \leq z - M_2(\underline{c}_1,\overline{c}_2,\gamma,\epsilon)||t_0 - t_1||\Big) - \eta_\gamma(\underline{c}_1,\overline{c}_2) \;\leq\; \mathbb{P}\left(z(S,t_0) \leq z\right)$$

$$\text{(B.20)} \qquad \mathbb{P}\left(z(S,t_0) \leq z\right) \leq \mathbb{P}\Big(z(S,t_1) \leq z + M_2(\underline{c}_1,\overline{c}_2,\gamma,\epsilon)||t_0 - t_1||\Big) + \eta_\gamma(\underline{c}_1,\overline{c}_2),$$

where

$$M_2(\underline{c}_1,\overline{c}_2,\gamma,\epsilon) \equiv M_1\Big(\underline{b}_1(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon), \overline{b}_1(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon), \underline{c}_1, \overline{c}_2, \underline{d}_1(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon), \overline{d}_2(\eta_\gamma(\underline{c}_1,\overline{c}_2),\epsilon)\Big).$$

For simplicity we write M_2 instead of $M_2(\underline{c}_1,\overline{c}_2,\gamma,\epsilon)$ whenever it is convenient.

Since (B.19) holds for any z, in particular it holds for $z = c(t_1;\alpha) + M_2||t_0 - t_1||$. Consequently:

$$\begin{aligned} \mathbb{P}\left(z(S,t_0) \leq c(t_1;\alpha) + M_2||t_0 - t_1||\right) \quad &\geq \quad \mathbb{P}\left(z(S,t_1) \leq c(t_1;\alpha)\right) - \eta_\gamma(\underline{c}_1,\overline{c}_2) \\ &= \quad 1 - \alpha - \eta_\gamma(\underline{c}_1,\overline{c}_2). \end{aligned}$$

This implies that

$$\text{(B.21)} \qquad c(t_0;\alpha + \eta_\gamma(\underline{c}_1,\overline{c}_2)) \leq c(t_1;\alpha) + M_2||t_0 - t_1||.$$

Likewise, equation (B.20) holds for any z, in particular it holds for $z = c(t_1;\alpha) - M_2||t_0 - t_1||$. This implies that:

$$\begin{aligned} \mathbb{P}\left(z(S,t_0) \leq c(t_1;\alpha) - M_2||t_0 - t_1||\right) \quad &\leq \quad \mathbb{P}\left(z(S,t_1) \leq c(t_1;\alpha)\right) + \eta_\gamma(\underline{c}_1,\overline{c}_2) \\ &= \quad (1 - \alpha) + \eta_\gamma(\underline{c}_1,\overline{c}_2). \end{aligned}$$

This implies that:

$$\text{(B.22)} \qquad c(t_0;\alpha - \eta_\gamma(\underline{c}_1,\overline{c}_2)) \geq c(t_1;\alpha) - M_2||t_0 - t_1||.$$

<u>Main conclusion of Step 2</u>: Finally, (B.21)-(B.22) and the definition of $\eta_\gamma(\underline{c}_1,\overline{c}_2)$ imply that for any $\gamma > 0$, $\epsilon > 0$ and any pair of constants $\underline{c}_1$, $\overline{c}_2$, one can find $\underline{b}_1^*, \overline{b}_2^*$, $\underline{d}_1^*$, $\overline{d}_2^*$, and M_2—that depend on $\underline{c}_1$, $\overline{c}_2$, γ and ϵ—such that for any:

$$(s_0,t_0),(s_1,t_1) \in \mathcal{K}^* \equiv [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1,\overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)],$$

the critical value function satisfies the Lipschitz-type condition:

$$|c(t_0;\alpha) - c(t_1;\alpha)| < M_2||t_0 - t_1|| + \gamma/2.$$

and:

$$\mathbb{P}(\chi_k^2 \notin (\underline{b}_1^*, \overline{b}_1^*) + \mathbb{P}(\chi_1^2 \notin (\underline{d}_1^*, \overline{d}_2^*)) < \epsilon/3.$$

Step 3: (Exploiting the Lipschitz property to manipulate $A^\delta \backslash A$) The constants in Step 2 depend on $\gamma > 0, \epsilon > 0$ and $\underline{c}_1, \overline{c}_2$. This step fixes $\gamma > 0$ and shows how to choose an appropriate enlargement of the set A as a function of γ. The Lipschitz condition established at the end of Step 2 allows for a convenient upper 'bound' on the set:

$$A^{\delta(\gamma)} \backslash A \ \cap [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1, \overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)]. \tag{B.23}$$

In particular, we show that for any $\gamma > 0$, there exists $\delta(\gamma)$ such that:

$$(s', t')' \in A^{\delta(\gamma)} \backslash A \ \cap [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1, \overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)]$$

implies that:

$$-\gamma \leq z(s, t) - c(t; \alpha) \leq 0.$$

This inclusion relation is convenient as it allows for the selection of the auxiliary parameter γ to make the probability of the set (B.23) uniformly small over (Π, F).

To establish the desired result, note that $x \equiv (s', t')' \in A^\delta \backslash A$ implies that:

$$z(s, t) - c(t; \alpha) < 0, \ \text{(as } x \equiv (s', t')' \notin A),$$

and also that, for any δ, there exists $x_0(\delta) \equiv (s'_{0,\delta}, t'_{0,\delta})' \in A$ such that

$$d(x, x_0(\delta)) \leq \delta.$$

Since the functions $s's$, $t't$, $(s't)^2/(t't)$ defining the set:

$$\mathcal{K}^* \equiv [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1, \overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)],$$

are Lipschitz continuous when restricted to $\mathcal{K}^*$, there exists δ^* small enough for which the corresponding $x_0(\delta^*)$ belongs to the set $\mathcal{K}^*$. In this case we have that:

$$\begin{aligned} ||(z(s,t) - c(t;\alpha)) - (z(s_{0,\delta^*}, t_{0,\delta^*}) - c(t_0;\alpha))|| &\leq ||z(s,t) - z(s_{0,\delta^*}, t_{0,\delta^*})|| + ||c(t;\alpha) - c(t_{0,\delta^*};\alpha)||, \\ &\leq (M_1(\mathcal{K}^*) + M_2(\underline{c}_1, \overline{c}_2, \gamma)) d(x, x_0(\delta^*)) + \gamma/2, \\ &\quad \text{(where we have used Step 2 part a) and b))}, \\ &\leq (M_1 + M_2)\delta^* + \gamma/2 \end{aligned}$$

Since $x \notin A$ and $x_0(\delta^*) \in A$ implies that

$$0 \geq (z(s,t) - c(t;\alpha)) \geq (z(s,t) - c(t;\alpha)) - (z(s_{0,\delta^*}, t_{0,\delta^*}) - c(t_0;\alpha)) \geq -(M_1 + M_2)\delta^* - \gamma/2$$

Main conclusion of Step 3: Taking $\delta(\gamma) \equiv \min\{\delta^*, \frac{\gamma}{2(M_1+M_2)}\}$ it follows that

$$(s', t')' \in A^{\delta(\gamma)} \backslash A \ \cap [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1, \overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)]$$

implies that:

$$-\gamma \leq z(s, t) - c(t; \alpha) \leq 0.$$

We now exploit this relation to show that one can choose γ to guarantee that

$$\sup_{\Pi, F \in \mathbb{R}^k \times \mathcal{F}} Q^n_{\beta_0, \Pi, \mathcal{F}}(A^{\delta(\gamma)} \backslash A \ \cap [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1, \overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)])$$

can be made arbitrarily small.

Step 4: (Choosing γ as a function of ϵ) Remember that equation (B.16) in Step 1 established that for any $\delta > 0$ and any constants $\underline{b}_1, \overline{b}_1, \underline{c}_1, \overline{c}_1, \underline{d}_1, \overline{d}_1$:

$$\begin{aligned} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta}\backslash A) &\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta}\backslash A \cap [B(\underline{b}_1,\overline{b}_1)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1,\overline{d}_2)]) \\ &+ \mathbb{P}(\chi^2_k \notin (\underline{b}_1,\overline{b}_2)) + \mathbb{P}(\chi^2_1 \notin (\overline{d}_1,\overline{d}_2)) \\ &+ Q^n_{\beta_0,\Pi\Sigma(F)}(A^{\delta}\backslash A \cap C^c(\underline{c}_1,\overline{c}_2)). \end{aligned}$$

Step 3 showed that for any $\gamma > 0$ there is a way of selecting the enlargement parameter $\delta(\gamma) > 0$ and constants $\underline{b}_1^*, \overline{b}_2^*, \underline{d}_1^*, \overline{d}_2^*$—that depend on $\underline{c}_1$, $\overline{c}_2$ and γ—such that the probability

$$Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta(\gamma)}\backslash A \cap [B(\underline{b}_1^*,\overline{b}_1^*)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1^*,\overline{d}_2^*)])$$

is less than or equal to

$$\text{(B.24)}\qquad Q^n_{\beta_0,\Pi,\Sigma(F)}(-\gamma \leq z(s,t) - c(t;\alpha) \leq 0 \cap [B(\underline{b}_1^*,\overline{b}_1^*)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1^*,\overline{d}_2^*)])$$

We now show that there exists $\gamma_\epsilon > 0$ small enough such that:

$$Q^n_{\beta_0,\Pi,\Sigma(F)}(-\gamma_\epsilon \leq z(s,t) - c(t;\alpha) \leq 0 \cap [B(\underline{b}_1^*,\overline{b}_1^*)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1^*,\overline{d}_2^*)]) < \epsilon/3$$

for any n, Π, F.

To show this, define—for any t such that $t't \in [\underline{c}_1, \overline{c}_2]$—the function $\gamma_\epsilon(t)$ to satisfy:

$$\mathbb{P}_S(-\gamma_\epsilon(t) \leq g(S,t) \leq 0 \text{ and } S'S \in (\underline{b}_1^*,\overline{b}_1^*) \text{ and } \text{ and } (S't)^2/(t't) \in (\underline{d}_1^*,\overline{d}_1^*)), \quad S \sim \mathcal{N}(0,\mathbb{I}_k),$$

where $g(s,t;\alpha) \equiv z(s,t) - c(t;\alpha)$ and t is treated as a fixed vector. Let

$$\gamma_\epsilon \equiv \inf_{\{t \mid t't \in [\underline{c}_1,\overline{c}_2]\}} \gamma_\epsilon(t)$$

and note that $\gamma_\epsilon > 0$ (otherwise, there will be a value t^* for which the distribution of $\mathbb{P}_S(g(S,t^*) = 0) > \epsilon/3$). Note that for any n, Π, F:

$$Q^n_{\beta_0,\Pi,\Sigma(F)}(-\gamma_\epsilon \leq g(s,t;\alpha) \leq 0 \text{ and } s's \in (\underline{b}_1^*,\overline{b}_1^*) \text{ and } t't \in (\underline{c}_1,\overline{c}_2), \text{ and } (s't)^2/t't \in (\underline{d}_1^*,\overline{d}_1^*)),$$

is the same as:

$$\int_{t\in C(\underline{c}_1,\overline{c}_2)} \Big(Q^n_{\beta_0,\Pi,\Sigma(F)}(-\gamma_\epsilon \leq g(s,t;\alpha) \leq 0 \text{ and } s's \in (\underline{b}_1^*,\overline{b}_1^*) \text{ and } \text{ and } (s't)^2/t't \in (\underline{d}_1^*,\overline{d}_1^*) \mid t)\Big) d\mathbb{P}^n_{\beta_0,\Pi,F}(t),$$

where $\mathbb{P}^n_{\beta_0,\Pi,F}$ is the marginal distribution that $Q^n_{\beta_0,\Pi,F}$ induces over (t). Note that (B.25) equals:

$$\int_{t\in C(\underline{c}_1,\overline{c}_2)} \Big(\mathbb{P}_S(-\gamma_\epsilon \leq g(S,t) \leq 0 \text{ and } S'S \in (\underline{b}_1^*,\overline{b}_1^*) \text{ and } \text{ and } (S't)^2/(t't) \in (\underline{d}_1^*,\overline{d}_1^*))\Big) d\mathbb{P}^n_{\beta_0,\Pi,F}(t),$$

$s|t$ has distribution $\mathcal{N}_k(\mathbf{0},\mathbb{I}_k)$ for any n, Π, F. And this is smaller than or equal:

$$\begin{aligned} &\int_{t\in C(\underline{c}_1,\overline{c}_2)} \Big(\mathbb{P}_S(-\gamma_\epsilon(t) \leq g(S,t) \leq 0 \text{ and } S'S \in (\underline{b}_1^*,\overline{b}_1^*) \text{ and } \text{ and } (S't)^2/(t't) \in (\underline{d}_1^*,\overline{d}_1^*))\Big) d\mathbb{P}^n_{\beta_0,\Pi,F}(t), \\ = \;& \mathbb{P}^n_{\beta_0,\Pi,F}(t \in C(\underline{c}_1,\overline{c}_2))\frac{\epsilon}{3} < \frac{\epsilon}{3}, \text{ (by definition of } \gamma_\epsilon(t)). \end{aligned}$$

<u>Main conclusion of Step 4</u>: This means that for any $\epsilon > 0$ there exists $\gamma_\epsilon > 0$ small enough such that:

$$Q^n_{\beta_0,\Pi,\Sigma(F)}(-\gamma_\epsilon \leq z(s,t) - c(t;\alpha) \leq 0 \cap [B(\underline{b}_1^*,\overline{b}_1^*)\cap C(\underline{c}_1,\overline{c}_2)\cap D(\underline{d}_1^*,\overline{d}_2^*)]) < \epsilon/3$$

for any n, Π, F.

Step 5 (choosing $\underline{c}_1$ and $\underline{c}_2$): Step 1 through Step 4 have shown that for any $\epsilon > 0$ there is a constant $\delta_\epsilon \equiv \delta(\gamma_\epsilon)$ and constants $\underline{b}_1^*, \overline{b}_2^*, \underline{d}_1^*, \overline{d}_2^*$—that depend on $\underline{c}_1$, $\overline{c}_2$ and ϵ such that for any n, Π, F:

$$\begin{aligned}
Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A) &\leq Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap [B(\underline{b}_1^*, \overline{b}_1^*) \cap C(\underline{c}_1, \overline{c}_2) \cap D(\underline{d}_1^*, \overline{d}_2^*)]) \\
&+ \mathbb{P}(\chi^2_k \notin (\underline{b}_1^*, \overline{b}_2^*)) + \mathbb{P}(\chi^2_1 \notin (\overline{d}_1^*, \overline{d}_2^*)) \\
&+ Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2)). \\
&\leq \frac{2\epsilon}{3} + Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2)).
\end{aligned}$$

This means that:

$$\sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A) \leq \frac{2\epsilon}{3} + \sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2)).$$

Thus, we only need to show that $\underline{c}_1$ and $\overline{c}_2$ can be chosen to make the second term on the right of the inequality above smaller than $\epsilon/3$. Let λ^* be defined as:

$$\lambda^* \equiv \max_{F\in\mathcal{F}} (a_0' \otimes \mathbb{I}_k)\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)$$

By assumption, there are constants $0 < \underline{\lambda} < \overline{\lambda} < \infty$ such that $\underline{\lambda} < \lambda^* < \overline{\lambda}$. Fix $c^* \in \mathbb{R}^k$ and partition $\mathbb{R}^k$ as follows:

$$\{\Pi \in \mathbb{R}^k \mid : n||\Pi||^2\lambda^* \leq c^*\} \cup \{\Pi \in \mathbb{R}^k \mid : n||\Pi||^2\lambda^* > c^*\} \equiv \mathbf{\Pi}^n_1(c^*) \cup \mathbf{\Pi}^n_2(c^*).$$

Note that

$$\sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} Q^n_{\beta_0,\Pi\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2))$$

is smaller than or equal to the sum of:

$$\sup_{(\Pi,F)\in\mathbf{\Pi}^n_1(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2)), \tag{B.25}$$

and

$$\sup_{(\Pi,F)\in\mathbf{\Pi}^n_2(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2)). \tag{B.26}$$

Step 5—Part a): First, we bound the term (B.25). Let $\chi^2_k(c)$ denote a non-central chi-square with k degrees of freedom and centrality parameter c. Note that:

$$\begin{aligned}
\sup_{(\Pi,F)\in\mathbf{\Pi}^n_1(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1, \overline{c}_2)) &\leq \sup_{(\Pi,F)\in\mathbf{\Pi}^n_1(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(C^c(\underline{c}_1, \overline{c}_2)), \\
&= \sup_{(\Pi,F)\in\mathbf{\Pi}^n_1(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(t't \notin (\underline{c}_1, \overline{c}_2)), \\
&= \sup_{(\Pi,F)\in\mathbf{\Pi}^n_1(c^*)\times\mathcal{F}} \mathbb{P}\Big(\chi^2_k(n||\Pi||^2\lambda(F)) \notin (\underline{c}_1, \overline{c}_2)\Big), \\
&\quad \text{(where } \lambda(F) \equiv (a_0' \otimes \mathbb{I}_k)\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)).
\end{aligned}$$

Therefore, one can choose constants $\underline{c}_1^*, \overline{c}_2^*$ that depend on c^* and ϵ (but do not depend on the sample size) such that for any $\underline{c}_1 < \underline{c}_1^*$ and $\overline{c}_2 > \overline{c}_2^*$:

$$\sup_{(\Pi,F)\in\mathbf{\Pi}_1^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \cap C^c(\underline{c}_1,\overline{c}_2)) < \frac{\epsilon}{6}.$$

Step 5—Part b): Now, we bound the term (B.26). To do this, choose $\overline{e}$ to satisfy

$$\mathbb{P}(\chi^2_k > \overline{e}) < \frac{\epsilon}{12}.$$

Since this selection of $\overline{e}$ implies that

$$\sup_{(\Pi,F)\in\mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \,\cap\, s's > \overline{e} \,\cap\, C^c(\underline{c}_1,\overline{c}_2)) < \frac{\epsilon}{12},$$

it is sufficient to show that there is $\overline{c}_2$ large enough such that:

$$\sup_{(\Pi,F)\in\mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A \,\cap\, s's \leq \overline{e} \,\cap\, C^c(\underline{c}_1,\overline{c}_2)) < \frac{\epsilon}{12}.$$

The key to establish this result is to show that for $t't$ large enough and $s's$ in a compact set, the test statistic $z(s,t)$ defined in Result 1 is close to the statistic $(s't)^2/(t't)$.

Step 5—Part c): Let $\mathbf{o}(t't)$ denote the function:

$$\mathbf{o}(t't) \equiv \Big(((s's/t't)-1)^2 + 4(s't)^2/(t't)^2\Big)^{1/2} - 1.$$

Note that:

$$\begin{aligned}
& 8\ln\Big[I_0\Big(((t't/8)(1+\mathbf{o}(t't))\Big)\Big], \\
= \;& 8\ln\Big[\frac{e^{(t't/8)(1+\mathbf{o}(t't))}}{\sqrt{2\pi((t't/8)(1+\mathbf{o}(t't)))}}\Big(1+O\Big(\frac{1}{(t't/8)(1+\mathbf{o}(t't))}\Big)\Big)\Big], \\
& \text{(where we have used the asymptotic approximation} \\
& \text{for } I_0(z) \text{ in p. 435 of Olver (1997) and the definition} \\
& \text{of } \sim \text{ in p. 4 of the same book)} \\
= \;& 8\ln\Big[\frac{e^{(t't/8)(1+\mathbf{o}(t't))}}{\sqrt{2\pi((t't/8)(1+\mathbf{o}(t't)))}}\Big] \\
+ \;& 8\ln\Big[\Big(1+O\Big(\frac{1}{(t't/8)(1+\mathbf{o}(t't))}\Big)\Big)\Big], \\
= \;& t't(1+\mathbf{o}(t't)) - 4\ln(2\pi) - 4\ln(t't/8) - 4\ln(1+\mathbf{o}(t't)) \\
+ \;& 8\ln\Big[\Big(1+O\Big(\frac{1}{(t't/8)(1+\mathbf{o}(t't))}\Big)\Big)\Big].
\end{aligned}$$

Therefore, $z_{\text{WAP}}(s,t)$ in Result 1:

$$(s's-t't) + 8\ln\Big[\;I_0\Big(\frac{1}{8}\Big[(s's-t't)^2+4(s't)^2\Big]^{1/2}\Big)\;\Big] + 4\ln(2pi) + 4\ln((1/8)t't)$$

can be written in terms of the conditional likelihood ratio statistic (CLR) as follows:

$$
\begin{aligned}
z_{\text{WAP}}(s,t) \quad &\equiv \quad (s's - t't) + t't(1+\mathbf{o}(t't)) - 4\ln(1+\mathbf{o}(t't)) \\
&+ \quad 8\ln\Big[\Big(1 + O\Big(\frac{1}{(t't/8)(1+\mathbf{o}(t't))}\Big)\Big)\Big], \\
&= \quad 2\text{CLR}(s,t) - 4\ln(1+\mathbf{o}(t't)) \\
&+ \quad 8\ln\Big[\Big(1 + O\Big(\frac{1}{(t't/8)(1+\mathbf{o}(t't))}\Big)\Big)\Big], \\
& \qquad \text{(where we have used the fact that } t't(1+\boldsymbol{o}(t't)) \\
& \qquad \text{equals } [(s's-t't)^2 + (s't)]^{1/2}\ \text{)}.
\end{aligned}
$$

It is well-known for large values of $t't$ and for values of $s's$ in a compact set the CLR can be approximated by the LM statistic ($\equiv s't/t't$) uniformly over the values of s. Choose $\zeta^* > 0$ to satisfy:

$$
\mathbb{P}(-\eta^* \leq N(0,1) \leq \eta^*) = \frac{\epsilon}{24}.
$$

Therefore, using the same argument as in part b) of step 2 one can show for $\zeta^* > 0$ there is $\overline{c}_2^*$—that depends on ζ^*—such that uniformly over $s's \leq \overline{e}$

$$
|z_{\text{WAP}}(s,t) - c_{\text{WAP}}(t;\alpha) - 2LM(s,t) + 2\chi^2_{1,1-\alpha}| < \zeta^*,
$$

where $\chi^2_{1,1-\alpha}$ is the 1-α quantile of a chi-squared random variable with one degree of freedom and $c_{\text{WAP}}(t;\alpha)$ is the conditonal critical value of $z_{\text{WAP}}(s,t)$.

Step 5—Part d): Note that

$$x \equiv (s,t) \in (A^{\delta_\epsilon} \backslash A \, \cap \, s's \leq \overline{e} \, \cap \, t't > \overline{c}_2^*),$$

implies that:

$$z_(s,t) - c(t;\alpha) < 0,$$

and also that there is $x_0(\delta_\epsilon) \equiv (s_0, t_0) \in A$ such that $z(s_0,t_0) - c(t;\alpha) > 0$ and $d(x, x_0(\delta_\epsilon)) < \delta_\epsilon$. Since the test based on the test statistic $z(s,t)$ with conditional critical value $c(t;\alpha)$ is equivalent to the test based on $z_{\text{WAP}}(s,t)$ and $c_{\text{WAP}}(t;\alpha)$, it follows that:

$$z_{\text{WAP}}(s,t) - c_{\text{WAP}}(t;\alpha) < 0, \text{ and } z_{\text{WAP}}(s_0,t_0) - c_{\text{WAP}}(t;\alpha) > 0.$$

Consequently:

$$\text{LM}(s,t) - \chi^2_{1,1-\alpha} < \zeta^*/2,$$

and

$$\text{LM}(s_0,t_0) - \chi^2_{1,1-\alpha} > -\zeta^*/2.$$

Note that the LM statistic can be written as a function of (S, ω_t) where $\omega_t \equiv t/||t||$. Since the partial derivatives of $(s'\omega_t)^2$ are bounded whenever $s's \leq \overline{e}$, the LM statistic satisfies the Lipschitz condition when s's belongs to the desired domain. Let $M(\overline{e})$ denote the Lipschitz constant of the LM statistic. Since:

$$-d(x, x_0(\delta_\epsilon))M(\overline{e}) \leq \text{LM}(s,t) - \text{LM}(s_0,t_0) \leq M(\overline{e})d(x, x_0(\delta_\epsilon)),$$

then:

$$\begin{aligned}
-\delta_\epsilon M(\overline{e}) &\leq -d(x, x_0(\delta_\epsilon))M(\overline{e}), \\
&\leq \text{LM}(s,t) - \chi^2_{1,1-\alpha} + \chi^2_{1,1-\alpha} - \text{LM}(s_0,t_0), \\
&\leq \text{LM}(s,t) - \chi^2_{1,1-\alpha} + \zeta^*/2, \\
&\quad \text{(where we have used the fact that } \text{LM}(s_0,t_0) - \chi^2_{1,1-\alpha} > -\zeta^*/2), \\
&\leq \zeta^*, \\
&\quad \text{(where we have used the fact that } \text{LM}(s,t) - \chi^2_{1,1-\alpha} < \zeta^*/2).
\end{aligned}$$

One can further shrink δ_ϵ to satisfy $\delta_\epsilon M(\overline{e}) < -\zeta^*$. This means that:

$$\begin{aligned}
\sup_{(\Pi,F)\in \mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon} \backslash A \, \cap \, s's \leq \overline{e} \, \cap \, t't > \overline{c}_2^*) &\leq \sup_{(\Pi,F)\in \mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(-\eta^* \leq \text{LM}(s,t) \leq \eta^*) \\
&= \mathbb{P}(-\eta^* \leq N(0,1) \leq \eta^*) \\
&= \frac{\epsilon}{24}
\end{aligned}$$

Since

$$\sup_{(\Pi,F)\in \mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon} \backslash A \, \cap \, s's \leq \overline{e} \, \cap \, C^c(\underline{c}_1, \overline{c}_2^*)),$$

is smaller than or equal to

$$\sup_{(\Pi,F)\in \mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon} \backslash A \, \cap \, s's \leq \overline{e} \, \cap \, t't \leq \underline{c}_1),$$

plus

$$\sup_{(\Pi,F)\in \mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon} \backslash A \, \cap \, s's \leq \overline{e} \, \cap \, t't > \underline{c}_2^*),$$

it follows that:

$$\sup_{(\Pi,F)\in \mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon} \backslash A \, \cap \, s's \leq \overline{e} \, \cap \, C^c(\underline{c}_1, \overline{c}_2^*)) < \frac{\epsilon}{24} + \frac{\epsilon}{24} = \frac{\epsilon}{12}.$$

This implies that:

$$\sup_{(\Pi,F)\in\mathbf{\Pi}_2^n(c^*)\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A\cap C^c(\underline{c}_1,\overline{c}_2)) < \frac{\epsilon}{12}+\frac{\epsilon}{12}=\frac{\epsilon}{6}.$$

<u>Main conclusion of Step 5</u>: The conclusion of Step 5 is that there are constants $\underline{c}_1^*, \overline{c}_2^*$ such that:

$$\sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A) \le \frac{2\epsilon}{3} + \sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A\cap C^c(\underline{c}_1^*,\overline{c}_2^*)) \le \frac{2\epsilon}{3}+\frac{\epsilon}{3}=\epsilon.$$

Step 1 to Step 5: We have shown that for every $\epsilon>0$ one can choose constants $\underline{b}_1^*, \overline{b}_2^*$, $\underline{c}_1^*, \overline{c}_2^*$, $\underline{d}_1^*$, $\overline{d}_2^*$ such that for any sample size and $(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}$:

$$\begin{aligned} Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta_\epsilon}\backslash A) &\le Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta}\backslash A\cap[B(\underline{b}_1^*,\overline{b}_1^*)\cap C(\underline{c}_1^*,\overline{c}_2^*)\cap D(\underline{d}_1^*,\overline{d}_2^*)]) \\ &+ \mathbb{P}(\chi^2_k\notin(\underline{b}_1^*,\overline{b}_2^*))+\mathbb{P}(\chi^2_1\notin(\overline{d}_1^*,\overline{d}_2^*)) \\ &+ Q^n_{\beta_0,\Pi\Sigma(F)}(A^{\delta}\backslash A\cap C^c(\underline{c}_1^*,\overline{c}_2^*)) \\ &\le \epsilon \end{aligned}$$

Since for any $\delta>0$:

$$\begin{aligned} P^n_{\beta_0,\Pi,\mathcal{F}}(A) &\le Q^n_{\beta_0,\Pi,\Sigma(F)}(A)+\frac{1}{\delta}d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F},Q^n_{\beta_0,\Pi,\Sigma(F)})+Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta}\backslash A), \\ &= \alpha + \frac{1}{\delta}d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F},Q^n_{\beta_0,\Pi,\Sigma(F)})+Q^n_{\beta_0,\Pi,\Sigma(F)}(A^{\delta}\backslash A), \end{aligned}$$

and, by assumption, there is $M_\epsilon\in\mathbb{N}$ such that for any $n\ge M_\epsilon$:

$$d_{\mathrm{BL}_1}(P^n_{\beta_0,\Pi,F},Q^n_{\beta_0,\Pi,\Sigma(F)})\le\epsilon\delta_\epsilon,$$

we conclude that for $n\ge M_\epsilon$:

$$\sup_{(\Pi,F)\in\mathbb{R}^k\times\mathcal{F}} P^n_{\beta_0,\Pi,\mathcal{F}}(A)\le\alpha+2\epsilon.$$

This establishes the asymptotic validity of the test in Result 1.

B.4.2. *Local Asymptotic Power of the test in Result 1*

Now we derive the local asymptotic power of the test in Result 1 under the following assumption:

Assumption L1: The class of distributions $\mathcal{F}$ is such that:

$$\lim_{n\to\infty} d_{\mathrm{BL}}\left(P^n_{\beta_0+\frac{c}{\sqrt{n}},\Pi,F},Q^n_{\beta_0+\frac{c}{\sqrt{n}}}\right)\to 0.$$

This is, the Bounded Lipschitz distance between the measures $P^n_{\beta,\Pi,F}$ and $Q^n_{\beta,\Pi,\Sigma(F)}$ coverges to zero as the sample size grows large for any local alternative of the form $\beta_0+c/\sqrt{n}$.

Asymptotic Efficiency of the test in Result 1: Suppose that Assumption L1 holds and suppose that there are constants $0<\underline{\lambda}<\overline{\lambda}<\infty$ such that the eigenvalues of $\Sigma(F)$ belong to an interval $[\underline{\lambda},\overline{\lambda}]$ for any $F\in\mathcal{F}$. If $\Sigma(F)=\Psi(F)\otimes\Phi(F)$ and $\Pi\neq\mathbf{0}_{k\times 1}$, then:

$$\liminf_{n\to\infty} P^n_{\beta_0+\frac{c}{\sqrt{n}},\Pi,F}(z_{\text{WAP}}(s,t)-c_{\text{WAP}}(t;\alpha)\ge 0)\ge\mathbb{P}\Big(\chi^2_1(\mu^2(\beta_0,c,\Pi,F))>\chi^2_{1,1-\alpha}\Big),$$

where $\chi^2_1(\mu^2(\beta_0,c,\Pi,F))$ is a non-central chi-square distribution with centrality parameter:

$$\mu^2(\beta_0,c,\Pi,F)\equiv c^2(\Pi'\Phi(F)^{-1}\Pi)(b_0'\Psi(F)b_0)^{-1}.$$

We establish the local efficiency of the test in Result 1 in 6 steps:

Step 0: Fix $(\Pi, F) \in \mathbb{R}^k \times \mathcal{F}$ and suppose that $||\Pi|| \neq 0$. Let $\overline{e}$ denote a positive scalar. In Part c) of Step 5 we have shown that for any $\zeta > 0$ there is $\underline{e}(\eta)$ such that for any $t't > \underline{e}(\eta)$—and uniformly over the values of $s's < \overline{e}$—it follows that:

$$|z_{\text{WAP}}(s,t) - c_{\text{WAP}}(t;\alpha) - 2\text{LM}(s,t) + 2\chi^2_{1,1-\alpha}| < \zeta/2,$$

where $\text{LM}(s,t) \equiv (s't)^2/t'^t$. This means that if $s's < \overline{e}$ and $t't > \underline{e}(\zeta)$:

$$LM(s,t) - \chi^2_{1,1-\alpha} > \zeta/4 \implies z_{\text{WAP}}(s,t) - c_{\text{WAP}}(t;\alpha) \geq 0.$$

Therefore, for any local alternative $\beta(c) \equiv \beta_0 + \frac{c}{\sqrt{n}}$:

$$\begin{aligned}
P^n_{\beta(c),\Pi,F}(z_{\text{WAP}}(s,t) - c_{\text{WAP}}(t;\alpha) \geq 0) \quad &\geq \quad P^n_{\beta(c),\Pi,F}(LM(s,t) \geq \chi^2_{1-\alpha} + \zeta \text{ and } s's < \overline{e} \text{ and } t't > \underline{e}(\zeta)),\\
&\geq \quad P^n_{\beta(c),\Pi,F}(LM(s,t) \geq \chi^2_{1-\alpha} + \zeta/4)\\
&+ \quad P^n_{\beta(c),\Pi,F}(s's < \overline{e})\\
&+ \quad P^n_{\beta(c),\Pi,F}(t't > \underline{e}(\zeta)) - 2,\\
&\quad\quad \text{(where we have used } P(A \cap B) = P(A) + P(B) - P(A \cup B)\\
&\quad\quad \text{twice, and also the fact that } P(A \cup B) \leq 1).
\end{aligned}$$

We now characterize the asymptotic behavior of the terms:

$$P^n_{\beta(c),\Pi,F}(LM(s,t) \geq \chi^2_{1-\alpha} + \zeta/4), \quad P^n_{\beta(c),\Pi,F}(s's < \overline{e}), \quad P^n_{\beta(c),\Pi,F}(t't > \underline{e}(\zeta)).$$

Step 1: Consider first the term:

$$P^n_{\beta(c),\Pi,F}(LM(s,t) \geq \chi^2_{1-\alpha} + \zeta/4).$$

Note that the event:

$$E_1 = \{(s',t')' \in \mathbb{R}^{2k} \mid LM(s,t) > \chi^2_{1,1-\alpha} + \zeta/4\},$$

is the same as the event:

$$E_1 = \{(s',t')' \in \mathbb{R}^{2k} \mid LM(s,t/\sqrt{n}) > \chi^2_{1,1-\alpha} + \zeta/4\},$$

as $\text{LM}(s,t) = (s't)^2/t't = s't/\sqrt{n})^2/(t/\sqrt{n})'(t/\sqrt{n})$.

Let $\widetilde{P}^n_{\beta(c),\Pi,F}$ denote the distribution of the random vector $(S', (t/\sqrt{n})')'$. Since the transformation $(x,y) \to (x, y/\sqrt{n})$ is Lipschitz for any $n \in \mathbb{N}$ (with constant 1) it follows—by assumption—that as $n \to \infty$:[3]

$$d_{\text{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, \widetilde{Q}^n_{\beta(c),\Pi,\Sigma(F)}) \to 0,$$

where $\widetilde{Q}$ is the distribution of:

$$\widetilde{Q}^n_{\beta,\Pi,\Sigma(F)} \equiv \mathcal{N}_{2k}\left(\begin{matrix} ([b_0' \otimes \mathbb{I}_k)\Sigma(b_0 \otimes \mathbb{I}_k)]^{-1/2}c\Pi \\ [(a_0' \otimes \mathbb{I}_k)\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)]^{-1/2}(a_0' \otimes \mathbb{I}_k)\Sigma(F)^{-1}(a_n(c) \otimes \mathbb{I}_k)\Pi \end{matrix}, \begin{pmatrix} \mathbb{I}_k & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \mathbb{I}_k/\sqrt{n} \end{pmatrix} \right),$$

[3] To see this, note that for any function $h \in \text{BL}_1$ we have that:

$$\left|\mathbb{E}_{\widetilde{P}^n_{\beta(c),\Pi,F}}[h(X)] - \mathbb{E}_{\widetilde{Q}^n_{\beta(c),\Pi,F}}[h(X)]\right| = \left|\mathbb{E}_{P^n_{\beta(c),\Pi,F}}[h \circ g_n(S,T)] - \mathbb{E}_{Q^n_{\beta(c),\Pi,F}}[h \circ g_n(S,T)]\right|,$$

where $g_n(s,t) = (s', (t/\sqrt{n})')'$ is an element of BL_1. Consequently,

$$0 \leq d_{\text{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, \widetilde{Q}^n_{\beta(c),\Pi,\Sigma(F)}) \leq d_{\text{BL}_1}(P^n_{\beta(c),\Pi,F}, Q^n_{\beta(c),\Pi,\Sigma(F)}).$$

with $a_n(c) = (\beta_0 + c/\sqrt{n}, 1)'$. Moreover, the convergence of the mean vector of this distribution and its covariance matrix this implies that:

$$d_{\mathrm{BL}_1}(\widetilde{Q}^n_{\beta(c),\Pi,\Sigma(F)}, Q_{\beta_0,\Pi,\Sigma(F)}) \to 0$$

where

$$Q_{\beta_0,\Pi,\Sigma(F)} \equiv \mathcal{N}_{2k}\left(\begin{matrix} ([b_0' \otimes \mathbb{I}_k)\Sigma(b_0 \otimes \mathbb{I}_k)]^{-1/2} c\Pi \\ [(a_0' \otimes \mathbb{I}_k)\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)]^{1/2}\Pi \end{matrix}, \begin{pmatrix} \mathbb{I}_k & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \mathbf{0}_{k\times k} \end{pmatrix} \right).$$

Thus, one can conclude that

$$d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,\Sigma(F)}, Q_{\beta_0,\Pi,\Sigma(F)}) \to 0.$$

For any event, and in particular for E_1,

$$\widetilde{P}^n_{\beta(c),\Pi,F}(E_1) \geq Q_{\beta_0,\Pi,\Sigma(F)}(E_1) - \frac{1}{\delta} d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, Q_{\beta_0,\Pi,\Sigma(F)}) - Q_{\beta_0,\Pi,\Sigma(F)}((E_1^c)^\delta \backslash E_1^c).$$

Under the probability measure $Q_{\beta_0,\Pi,\Sigma(F)}$ (which does not depend on n), the topological boundary of E_1^c

$$\mathrm{Bd}(E_1^c) \equiv \{(s', t')' \in \mathbb{R}^{2k} \mid LM(s,t) = \chi^2_{1,1-\alpha} + \zeta/4\},$$

has probability zero. Therefore, there exists $\delta_{\epsilon,\zeta}$ (independent of n) such that:

$$Q_{\beta_0,\Pi,F}((E^c)_!^{\delta_{\epsilon,\zeta}} \backslash E_1^c) < \frac{\epsilon}{6}.$$

Moreover, by choosing $N(\epsilon,\zeta) \in \mathbb{N}$ to be such that for $n \geq N(\epsilon,\zeta)$:

$$d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,\Sigma(F)}, Q_{\beta_0,\Pi,\Sigma(F)}) \leq \delta_{\epsilon,\zeta}\frac{\epsilon}{6},$$

we have that:

$$\begin{aligned} P^n_{\beta(c),\Pi,F}(E_1) &= \widetilde{P}^n_{\beta(c),\Pi,F}(E_1), \\ &\geq Q_{\beta_0,\Pi,\Sigma(F)}(E_1) - \frac{1}{\delta_{\epsilon,\zeta}} d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, Q_{\beta_0,\Pi,\Sigma(F)}) - Q_{\beta_0,\Pi,\Sigma(F)}((E_1^c)^{\delta_{\epsilon,\zeta}} \backslash E_1^c), \\ &\geq Q_{\beta_0,\Pi,\Sigma(F)}(E_1) - \frac{\epsilon}{6} - \frac{\epsilon}{6}, \\ &= Q_{\beta_0,\Pi,\Sigma(F)}(E_1) - \frac{\epsilon}{3}. \end{aligned}$$

Moreover, under the probability measure $Q_{\beta_0,\Pi,\Sigma(F)}(E_1)$:

$$\frac{t'S}{\sqrt{t't}} \sim \mathcal{N}\left(\mu(\beta_0, c, \Pi, F)\ ,\ 1\right),$$

where

$$\mu(\beta_0, c, \Pi, F) \equiv c\frac{\Pi'\Big([(a_0' \otimes \mathbb{I}_k)]\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)]^{1/2}[(b_0' \otimes \mathbb{I}_k)\Sigma(F)(b_0 \otimes \mathbb{I}_k)]^{-1/2}\Big)\Pi}{\Big(\Pi'[(a_0' \otimes \mathbb{I}_k)]\Sigma^{-1}(F)(a_0 \otimes \mathbb{I}_k)]\Pi\Big)^{1/2}}.$$

This means that $Q_{\beta_0,\Pi,\Sigma(F)}(E_1)$ is the same as:

$$\mathbb{P}\Big(\chi_1^2(\mu^2(\beta_0, c, \Pi, F)) > \chi^2_{1,1-\alpha} + \eta\Big).$$

This means that for any $\zeta > 0, \epsilon > 0$ there is $N(\epsilon,\zeta) \in \mathbb{N}$ such that for $n \geq N(\epsilon,\zeta)$:

$$P^n_{\beta(c),\Pi,F}(E_1) \geq \mathbb{P}\Big(\chi_1^2(\mu^2(\beta_0, c, \Pi, F)) > \chi^2_{1,1-\alpha} + \zeta\Big) - \frac{\epsilon}{3},$$

where $\chi_1^2(\mu^2(c, \beta_0, c, \Pi, F))$ is a non-central chi-square distribution with centrality parameter:

$$\mu(\beta_0, c, \Pi, F) \equiv c\frac{\Pi'\Big([(a_0' \otimes \mathbb{I}_k)]\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)]^{1/2}[(b_0' \otimes \mathbb{I}_k)\Sigma(F)(b_0 \otimes \mathbb{I}_k)]^{-1/2}\Big)\Pi}{\Big(\Pi'[(a_0' \otimes \mathbb{I}_k)]\Sigma^{-1}(F)(a_0 \otimes \mathbb{I}_k)]\Pi\Big)^{1/2}}.$$

Step 2: Now we take care of the term $P^n_{\beta(c),\Pi,F}(s's < \overline{e})$. In particular, we show how to choose $\overline{\epsilon}$ to make the term of interest to be at least $1 - \epsilon/3$ for n large enough. Let E_2 denote the event:

$$E_2 \equiv \{(s', t')' \mid s's < \overline{e}\}$$

Note that for any $\delta > 0$:

$$\begin{aligned} P^n_{\beta(c),\Pi,F}(E_2) &\geq Q^n_{\beta(c),\Pi,\Sigma(F)}(E_2) - \frac{1}{\delta}d_{\mathrm{BL}_1}(P^n_{\beta(c),\Pi,F}, Q^n_{\beta(c),\Pi,\Sigma(F)}) - Q^n_{\beta(c),\Pi,\Sigma(F)}((E_2^c)^\delta \backslash E^c), \\ &= \mathbb{P}_S(E_2) - \frac{1}{\delta}d_{\mathrm{BL}_1}(P^n_{\beta(c),\Pi,F}, Q^n_{\beta(c),\Pi,\Sigma(F)}) - \mathbb{P}_S((E_2^c)^\delta \backslash E_2^c), \end{aligned}$$

where $S \sim \mathcal{N}_k([b_0' \otimes \mathbb{I}_k \Sigma(F)(b_0 \otimes \mathbb{I}_k]^{-1/2}\Pi c\,, \mathbb{I}_{2k})$. Once again, since the topological boundary of E_2 has zero measure under $\mathbb{P}_S$ there exists $\delta_{\epsilon,\overline{e}}$ such that:

$$\mathbb{P}_S((E_2^c)^{\delta_{\epsilon,\overline{e}}} \backslash E_2^c) < \frac{\epsilon}{9}.$$

This means that one can choose $N(\epsilon, \overline{e}) \in \mathbb{N}$ such that for $n \geq N(\epsilon, \overline{e})$:

$$P^n_{\beta(c),\Pi,F}(E_2) \geq \mathbb{P}_S(E_2) - \frac{2\epsilon}{9}.$$

Moreover, there is $\overline{e}^*$ large enough such that $\mathbb{P}_S(E_2) \geq 1 - \frac{\epsilon}{9}$. Therefore, for $n \geq N(\epsilon, \overline{e}^*)$,

$$P^n_{\beta(c),\Pi,F}(E_2) \geq 1 - \frac{\epsilon}{3}.$$

Step 3: Finally, consider the term $P^n_{\beta(c),\Pi,F}(t't > \underline{e})$. We show that for any fixed $\overline{\epsilon}$ there is n large enough such that this term is at least $1 - \epsilon/3$. Define:

$$E_{3,n} \equiv \{(s', t')' \mid t't > \underline{e}/\sqrt{n}\}.$$

Note that:

$$\begin{aligned} P^n_{\beta(c),\Pi,F}(t't > \underline{e}) &= \widetilde{P}^n_{\beta(c),\Pi,F}(E_{3,n}), \\ &= Q_{\beta_0,\Pi,\Sigma(F)}(E_{3,n}) - \frac{1}{\delta}d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, Q_{\beta_0,\Pi,\Sigma(F)}) - Q_{\beta_0,\Pi,\Sigma(F)}((E_{3,n}^c)^\delta \backslash E_{3,n}^c), \\ &\geq Q_{\beta_0,\Pi,\Sigma(F)}(E_{3,n}) - \frac{1}{\delta}d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, Q_{\beta_0,\Pi,\Sigma(F)}) - Q_{\beta_0,\Pi,\Sigma(F)}((E_{3,n}^c)^\delta). \end{aligned}$$

Under $Q_{\beta_0,\Pi,\Sigma(F)}$, the statistic t has a degenerate distribution with all of its mass at $t^* \equiv [(a_0' \otimes \mathbb{I}_k)]\Sigma^{-1}(F)(a_0 \otimes \mathbb{I}_k)]^{1/2}\Pi \equiv 0$. This means that for n large enough:

$$Q_{\beta_0,\Pi,\Sigma(F)}(E_{3,n}) = \mathbf{1}\{t^* \in E_{3,n}\} = 1.$$

Take any δ^* such that the ball of radius δ^* around t^*, $B_{\delta^*}(t^*)$, excludes the origin. Note that for any n such that $e/\sqrt{n} < \delta^*$:

$$B_{\underline{e}^{1/2}/n^{1/4}}(\mathbf{0}_{k\times 1}) \cap B_{\delta^*}(t^*) = \emptyset.$$

This means that for n large enough there is no $t_0 \in \mathbb{R}^k$ such that: $t_0't_0 < \underline{e}/\sqrt{n}$ and $d(t_0, t^*) \leq \delta^*$. This means that for n large enough $t^* \notin (E_{3,n}^c)^\delta$. Therefore for n large enough:

$$
\begin{aligned}
P^n_{\beta(c),\Pi,F}(t't > \underline{e}) &\geq 1 - \frac{1}{\delta^*} d_{\mathrm{BL}_1}(\widetilde{P}^n_{\beta(c),\Pi,F}, Q_{\beta_0,\Pi,\Sigma(F)}) - 0 \\
&\geq 1 - \frac{\epsilon}{3}.
\end{aligned}
$$

Step 4: We have shown that:

$$
\begin{aligned}
P^n_{\beta(c),\Pi,F}(z(s,t) - c(t;\alpha) \geq 0) &\geq P^n_{\beta(c),\Pi,F}(LM(s,t) \geq \chi^2_{1-\alpha} + \zeta/4) \\
&+ P^n_{\beta(c),\Pi,F}(s's < \overline{e}^*) \\
&+ P^n_{\beta(c),\Pi,F}(t't > \underline{e}(\zeta)) - 2.
\end{aligned}
$$

Steps 1, 2, 3 of this proof have established the existence of $N(\epsilon,\zeta,\overline{e}^*)$ such that for any $n \geq N(\epsilon,\zeta,\overline{e}^*)$:

$$
\begin{aligned}
P^n_{\beta(c),\Pi,F}(LM(s,t) \geq \chi^2_{1-\alpha} + \zeta/4) &\geq \mathbb{P}\Big(\chi^2_1(\mu^2(\beta_0,c,\Pi,F)) > \chi^2_{1,1-\alpha} + \zeta\Big) - \frac{\epsilon}{3}, \\
P^n_{\beta(c),\Pi,F}(s's < \overline{e}^*) &\geq 1 - \frac{\epsilon}{3}, \\
P^n_{\beta(c),\Pi,F}(t't > \underline{e}(\zeta)) &\geq 1 - \frac{\epsilon}{3},
\end{aligned}
$$

where the centrality parameter $\mu^2(\beta_0,c,\Pi,F)$ is given by:

$$
\mu^2(\beta_0,c,\Pi,F) \equiv \left(c \frac{\Pi'\Big([(a_0' \otimes \mathbb{I}_k)]\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)]^{1/2}[(b_0' \otimes \mathbb{I}_k)\Sigma(F)(b_0 \otimes \mathbb{I}_k)]^{-1/2}\Big)\Pi}{\Big(\Pi'[(a_0' \otimes \mathbb{I}_k)]\Sigma^{-1}(F)(a_0 \otimes \mathbb{I}_k)]\Pi\Big)^{1/2}} \right)^2 .
$$

This implies that:

$$
\liminf_{n\to\infty} P^n_{\beta_0+\frac{c}{\sqrt{n}},\Pi,F}(z_{\mathrm{WAP}}(s,t) - c_{\mathrm{WAP}}(t;\alpha) \geq 0) \geq \mathbb{P}\Big(\chi^2_1(\mu^2(\beta_0,c,\Pi,F)) > \chi^2_{1,1-\alpha} + \zeta\Big),
$$

for any $\zeta > 0$. Since the distribution on the right-hand side is continuous in ζ, then:

$$
\liminf_{n\to\infty} P^n_{\beta_0+\frac{c}{\sqrt{n}},\Pi,F}(z_{\mathrm{WAP}}(s,t) - c_{\mathrm{WAP}}(t;\alpha) \geq 0) \geq \mathbb{P}\Big(\chi^2_1(\mu^2(\beta_0,c,\Pi,F)) > \chi^2_{1,1-\alpha}\Big).
$$

Step 5: If $\Sigma(F) = \Psi(F) \otimes \Phi(F)$:

$$
[(a_0' \otimes \mathbb{I}_k)]\Sigma(F)^{-1}(a_0 \otimes \mathbb{I}_k)] = (a_0'\Psi^{-1}a_0)\Phi^{-1}
$$

and

$$
[(b_0' \otimes \mathbb{I}_k)\Sigma(F)(b_0 \otimes \mathbb{I}_k)] = (b_0'\Psi b_0)\Phi.
$$

Therefore, the expression for the centrality parameter simplifies to:

$$
\mu^2(\beta_0,c,\Pi,F) \equiv c^2(\Pi'\Phi^{-1}\Pi)(b_0'\Psi b_0)^{-1}.
$$

Final Comments: The lower bound on local power above implies that the test in Result 1 is as powerful (locally) as a GMM-Wald test for β_0 based on the sample moment condition:

$$
\frac{1}{\sqrt{n}} Z'(y - \beta_0 x) = \mathbf{0}.
$$

To see this, note that under Assumption L0 the asymptotic variance of the sample moment condition is simply QW_0Q where Q is the probability limit of $Z'Z/n$ and $W_0 \equiv (b_0' \otimes \mathbb{I}_k)\Sigma(b_0 \otimes \mathbb{I}_k)$. Therefore, the efficient GMM estimator for β (assuming W_0 is known) is:

$$
\begin{aligned}
\beta_{\mathrm{GMM}} &= \Big(X'Z(Z'ZW_0Z'Z)^{-1}Z'X\Big)^{-1} X'Z(Z'ZW_0Z'Z)^{-1}Z'y \\
&= \Big(\widehat{\gamma}_2 W_0^{-1}\widehat{\gamma}_2\Big)^{-1}\widehat{\gamma}_2 W_0^{-1}\widehat{\gamma}_1.
\end{aligned}
$$

The efficient α-level GMM-Wald test for $\beta = \beta_0$ rejects whenever:

$$\left(\left(\widehat{\gamma}_2 W_0^{-1}\widehat{\gamma}_2\right)^{-1/2}\widehat{\gamma}_2 W_0^{-1}\sqrt{n}(\widehat{\gamma}_1 - \beta_0\widehat{\gamma}_2)\right)^2 > \chi^2_{1,1-\alpha},$$

and this test has local power, under alternatives of the form $\beta_0 + c/\sqrt{n}$, given by:

$$\mathbb{P}\Big(\chi^2_1(c^2(\Pi' W_0^{-1}\Pi)) > \chi^2_{1,1-\alpha}\Big).$$

If Σ is of the form $\Psi \otimes \Omega$, then $c^2(\Pi' W_0^{-1}\Pi)$ coincides with the centrality parameter $\mu^2(\beta_0, c, \Pi, F)$.

B.4.3. *Details of the weights for* (β, Π) *in the Kronecker case*

This section analyzes the properties of the weights:

$$\text{(B.27)} \qquad \begin{pmatrix}\beta\Pi\\ \Pi\end{pmatrix} = n^{-1/2}\left(\Psi_\Sigma \widetilde{C}_0' \otimes \Phi_\Sigma^{1/2}\right)\rho(\phi\otimes\omega),\ \widetilde{C}_0 \equiv \begin{pmatrix}(b_0'\Psi_\Sigma b_0)^{-1/2} b_0'\\ (a_0'\Psi_\Sigma^{-1}a_0)^{-1/2}a_0'\Psi_\Sigma^{-1}\end{pmatrix},$$

with

$$\text{(B.28)} \qquad \phi \sim \mathcal{U}(\mathcal{S}^1), \quad \omega \sim \mathcal{U}(\mathcal{S}^{k-1}),$$

and

$$\text{(B.29)} \qquad \rho|\phi,\omega \sim \sqrt{\chi^2_k}\ /\ (\phi'\otimes\omega')\left(\widetilde{C}_0\Psi_\Sigma' \otimes {\Phi_\Sigma'}^{1/2}\right)\Sigma^{-1}\left(\Psi_\Sigma\widetilde{C}_0'\otimes\Phi_\Sigma^{1/2}\right)(\phi\otimes\omega).$$

The main assumption of this section is that $\Sigma = \Psi\otimes\Phi$. Note first that when $\Sigma = \Psi\otimes\Phi$:

$$\Psi_\Sigma = (\text{vec}(\Phi)'\text{vec}(\Phi))^{1/2}\Psi, \quad \Phi_\Sigma = \Phi/(\text{vec}(\Phi)'\text{vec}(\Phi))^{1/2}.$$

Therefore, we can write the weights in (B.27) as:

$$\text{(B.30)} \qquad \begin{pmatrix}\beta\Pi\\ \Pi\end{pmatrix} = n^{-1/2}\left(\Psi C_0'\otimes\Phi\right)\rho(\phi\otimes\omega),\ C_0 \equiv \begin{pmatrix}(b_0'\Psi b_0)^{-1/2}b_0'\\ (a_0'\Psi^{-1}a_0)^{-1/2}a_0'\Psi^{-1}\end{pmatrix},$$

Weight for β: Under (B.30), the parameter β equals:

$$\beta \quad = \quad \frac{[1,0]\Psi C_0'\phi}{[0,1]\Psi C_0'\phi}.$$

This ratio can be simplified using the following equalities. First:

$$\begin{aligned}
[1,0]\Psi C_0'\phi \quad &= \quad [1,0]\Psi\Big[b_0(b_0'\Psi b_0)^{-1/2}\ ,\ \Psi^{-1}a_0(a_0'\Psi^{-1}a_0)^{-1/2}\Big]\phi,\\
&\qquad \text{(by definition of } C_0)\\
&= \quad \Big[[1,0]\Psi b_0(b_0'\Psi b_0)^{-1/2}\ ,\ \beta_0(a_0'\Psi^{-1}a_0)^{-1/2}\Big]\phi,\\
&\qquad \text{(as } [1,0]a_0 = \beta_0).
\end{aligned}$$

Second:

$$a_0'\Psi^{-1}a_0 = \frac{1}{\det(\Psi)}a_0'\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\Psi\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}a_0 = \det\left(\begin{pmatrix}1 & -\beta_0\\ 0 & 1\end{pmatrix}\Psi\begin{pmatrix}1 & -\beta_0\\ 0 & 1\end{pmatrix}'\right)^{-1} b_0'\Psi b_0,$$

and

$$1 - r(\beta_0)^2 = \det\left(\begin{pmatrix}1 & -\beta_0\\ 0 & 1\end{pmatrix}\Psi\begin{pmatrix}1 & -\beta_0\\ 0 & 1\end{pmatrix}'\right)\Big/(b_0'\Psi b_0)([0,1]\Psi[0,1]'),$$

where $r(\beta_0)$ refers to the correlation coefficient of $(b_0'; [0,1])\Psi(b_0'; [0,1])'$. This means that the numerator for β is given by:

$$\begin{aligned}
[1,0]\Psi C_0'\phi &= [1,0]\Psi b_0(b_0'\Psi b_0)^{-1/2}\phi_1 + \beta_0\sqrt{1-r^2(\beta_0)}([0,1]\Psi[0,1]')^{1/2}\phi_2, \\
&= [1,0]\Psi b_0(b_0'\Psi b_0)^{-1/2}\phi_1 - \beta_0([0,1]\Psi b_0)(b_0'\Psi b_0)^{-1/2}\phi_1 \\
&+ \beta_0\left(r(\beta_0)\phi_1 + \sqrt{1-r^2(\beta_0)}([0,1]\phi_2\right)([0,1]\Psi[0,1]')^{1/2}, \\
&\quad \text{(where we have added and subtracted } \beta_0 r(\beta_0)([0,1]\Psi[0,1]')^{1/2}\phi_1).
\end{aligned}$$

Therefore:

$$[1,0]\Psi C_0'\phi = (b_0'\Psi b_0)^{1/2}\phi_1 + \beta_0[0,1]\Psi C_0'\phi,$$

where we have used the fact that:

$$[0,1]\Psi C_0'\phi = \left(r(\beta_0)\phi_1 + \sqrt{1-r^2(\beta_0)}([0,1]\phi_2\right)([0,1]\Psi[0,1]')^{1/2}.$$

This means that β can be written as:

$$\beta = \frac{[1,0]\Psi C_0'\phi}{[0,1]\Psi C_0'\phi} = \frac{(b_0'\Psi b_0)^{1/2}\phi_1}{[0,1]\Psi C_0'\phi} + \beta_0. \tag{B.31}$$

Weight for Π: The distribution of the first-stage coefficient is given by:

$$\sqrt{n}\Pi = ([0,1]\Psi C_0'\phi)\Phi^{1/2}\rho\omega. \tag{B.32}$$

This means that:

$$\sqrt{n}\Pi \mid \phi \sim \mathcal{N}_k\left(\mathbf{0}\,,\ ([0,1]\Psi C_0'\phi)^2\Phi\right).$$

Comparison to the MM2 weights: We claim that if $\Sigma = \Psi \otimes \Phi$ the weights in (3.3) and (3.4) are equivalent to the 'MM2' weights proposed in MM15. To see this, note that (B.31) implies that the vector $(\beta,1)'$ can be written as $\Psi C_0'\phi$ divided by $[0,1]\Psi C_0'\phi$. Since the vector $(c_\beta, d_\beta)'$ in MM15 equals $C_0(\beta,1)'$, then:

$$||(c_\beta, d_\beta)'|| = ||C_0(\beta,1)'|| = ||C_0\Psi C_0'\phi||/|[0,1]\Psi C_0'\phi| = 1/|[0,1]\Psi C_0'\phi|.$$

Therefore,

$$1/||(c_\beta, d_\beta)'||^2 = ([0,1]\Psi C_0'\phi)^2.$$

This implies that

$$\sqrt{n}\Pi \mid \phi \sim \mathcal{N}_k\left(\mathbf{0}\,,\ (||(c_\beta, d_\beta)'||^{-2}\Phi\right),$$

which is the same distribution as in MM15, up to a scaling constant. Also, MM15 assumes that the distribution of the angle of

$$C_0(\beta,1)'/||C_0(\beta,1)'||$$

is uniform on $[-\Pi, \Pi]$. Under (B.31) it follows that

$$C_0(\beta,1)'/||C_0(\beta,1)'|| = \phi,$$

where ϕ is uniformly distributed on the unit circle $\mathcal{S}^1$. Part ii) of exercise 5.2.4 in Stroock (1999) implies that ϕ can be written as $[\cos(\theta)', \sin(\theta)']'$ where θ is uniformly distributed on a connected interval of length $2pi$.

Distribution of $\sqrt{\lambda}(\beta - \beta_0)$ and λ: The Monte-Carlo exercises in Andrews et al. (2006) depend on the parameters:

$$\lambda \equiv n\Pi\Phi^{-1}\Pi, \text{ and } \sqrt{\lambda}(\beta - \beta_0).$$

Equation (B.32) implies

$$\begin{aligned} \lambda &\equiv \Big(([0,1]\Psi C_0'\phi)\Phi^{1/2}\rho\omega\Big)\,\Phi^{-1}\,\Big(([0,1]\Psi C_0'\phi)\Phi^{1/2}\rho\omega\Big) \\ &= ([0,1]\Psi C_0'\phi)^2\rho^2\omega'\Phi^{1/2}\Phi^{-1}\Phi^{1/2}\omega \\ &= ([0,1]\Psi C_0'\phi)^2\rho^2. \end{aligned}$$

Consequently, equation (B.31) implies

$$\begin{aligned} \sqrt{\lambda}(\beta-\beta_0) &= \sqrt{([0,1]\Psi C_0'\phi)^2\rho^2}\left(\frac{(b_0'\Phi b_0)^{1/2}\phi_1}{[0,1]\Psi C_0'\phi}\right) \\ &= (b_0'\Phi b_0)^{1/2}\rho\phi_1. \end{aligned}$$

Therefore:

$$\sqrt{\lambda}(\beta-\beta_0) = \Big(b_0'\Psi b_0\Big)^{1/2}\rho\phi_1, \tag{B.33}$$

$$\lambda = ([0,1]\Psi C_0'\phi)^2\rho^2, \tag{B.34}$$

The probability density function of $(\sqrt{\lambda}(\beta-\beta_0),\lambda)$ is given in Figure 1 in the main text of the paper.

B.4.4. *Asymptotic equivalence between the test in Result 1 and the CLR as $t't\to\infty$*

The CLR statistic can be written as

$$\frac{1}{2}\Big((s's-t't)+8x(s,t)\Big),$$

where

$$\begin{aligned} x(s,t) &\equiv \frac{1}{8}\Big((s's-t't)^2+4(s't)^2\Big)^{1/2}, \\ &= \frac{t't}{8}\left(\left(\frac{s's}{t't}-1\right)^2+4\left(\frac{s't}{\sqrt{t't}}\right)^2\frac{1}{t't}\right)^{1/2}. \end{aligned}$$

The equation above implies $x(s,t)\to\infty$ as $t't\to\infty$ for any s. Moreover, the test statistic $z_{\text{WAP}}(s,t)$ in Result 1 equals

$$(s's-t't)+8\ln\Big(I_0\Big(x(s,t)\Big)\Big)+4\ln(2\pi)+4\ln((1/8)t't).$$

The asymptotic approximation for the modified Bessel function $I_0(x)$ given in Olver (1997), p. 435 implies that

$$I_0(x)\Bigg/\frac{e^x}{(2\pi x)^{1/2}}\to 1,\ \text{ as } x\to\infty,$$

Therefore, as $t't\to\infty$

$$\begin{aligned} z_{\text{WAP}}(s,t) &= (s's-t't)+8\ln\left(\frac{e^{x(s,t)}}{(2\pi x(s,t))^{1/2}}\right)+4\ln(2\pi)+4\ln((1/8)t't+o(1) \\ &= (s's-t't)+8x(s,t)-4\ln(2\pi x(s,t))+4\ln(2\pi)+4\ln((1/8)t't+o(1) \\ &= 2\text{CLR}-4\ln(x(s,t)/(t't/8))+o(1) \\ &= 2\text{CLR}+o(1). \end{aligned}$$

REFERENCES

Abramowitz, M. and I. Stegun (1964): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, vol. 55, Dover publications, New York.

Aliprantis, C. and K. Border (2006): *Infinite Dimensional Analysis: a Hitchhiker's Guide*, Springer Verlag, 3rd ed.

Andrews, D., M. Moreira, and J. Stock (2004): "Optimal Invariant Similar Tests for Instrumental Variables Regression," *Working paper, Harvard University*.

——— (2006): "Optimal Two-sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*, 74, 715–52.

Billingsley, P. (1995): *Probability and Measure*, John Wiley & Sons, New York, 3rd ed.

Brown, L. D. (1986): "Fundamentals of statistical exponential families with applications in statistical decision theory," *Lecture Notes-monograph series*, 9, i–279.

Chernozhukov, V., C. Hansen, and M. Jansson (2009): "Admissible Invariant Similar Tests for Instrumental Variables Regression," *Econometric Theory*, 25, 806–818.

Dudley, R. (2002): *Real Analysis and Probability*, vol. 74, Cambridge University Press.

Ferguson, T. (1967): *Mathematical Statistics: A Decision Theoretic Approach*, vol. 7, Academic Press New York.

Lehmann, E. and J. Romano (2005): *Testing Statistical Hypotheses*, Springer Texts in Statistics, Springer Verlag.

Lehmann, E. L. and G. Casella (1998): *Theory of point estimation*, vol. 31 of *Springer Texts in Statistics*, Springer, New York, 2nd ed.

Mardia, K. and P. Jupp (2000): *Directional Statistics*, Wiley Series in Probability and Statistics, Wiley.

Moreira, M. (2003): "A conditional likelihood ratio test for structural models," *Econometrica*, 71, 1027–1048.

Ok, E. (2019): *Probability theory with economic applications*, forthcoming.

Olver, F. W. (1997): *Asymptotics and special functions*, AKP Classics, Academic press, 2nd ed.

Rudin, W. (2005): *Functional Analysis.*, International Series in Pure and Applied Mathematics, McGraw-Hill, New York.

——— (2006): *Real and Complex Analysis*, Tata McGraw-Hill Education, 3rd ed.

Stein, E. (2011): *Functional Analysis: Introduction to Further Topics in Analysis*, Princeton Univ Press.

Stroock, D. (1999): *A Concise Introduction to the Theory of Integration*, Birkhauser, 2nd ed.

Van der Vaart, A. (2000): *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Van der Vaart, A. and J. Wellner (1996): *Weak Convergence and Empirical Processes.*, Springer, New York.