A Connection to linear independence constraint qualification (LICQ)

In this section, we draw connections between linear independence constraint qualification (LICQ) and our Assumption 3, under which the power of the conditional test converges to the power envelope asymptotically. We show that LICQ implies Assumption 3 but is somewhat stronger. We follow the notation provided in Kaido et al. (2019), whose results connect constraint qualifications to other common assumptions made in the moment inequality literature. See also Gafarov (2019); Cho and Russell (2018); Flynn (2019) and Kaido and Santos (2014) for other recent applications of LICQ in moment inequality models.

We begin by defining LICQ. Suppose $\Delta = \{ \delta : A\delta \leq d \}$. Let $T(\Delta, \beta) := \{ \tau_{\text{post}} : m(\tau_{\text{post}}; \beta) \leq 0 \}$, where $m(\tau_{\text{post}}; \beta) = A(\beta - M_{\text{post}}\tau_{\text{post}}) - d$, be the identified set for the full parameter vector $\tau_{\text{post}}$. We define the set of support points in direction $p$ to be $S(p, T) := \{ \tau_{\text{post}} : p'\tau_{\text{post}} = \sup_{\tau_{\text{post}} \in T} p'\tau_{\text{post}} \}$.

**Definition 1.** We say that the linear constraint constraint qualification (LICQ) is satisfied in the direction $p$ if for all support points in the direction $p$, the gradients of the binding constraints are linearly independent. That is, for all $\tau_{\text{post}} \in S(p, T)$, the set $\{ D_{\tau_{\text{post}}}m_j(\tau_{\text{post}}, \beta) : m_j(\tau_{\text{post}}, \beta) = 0 \}$ is linearly independent, where $D_{\tau_{\text{post}}}$ denotes the gradient with respect to $\tau_{\text{post}}$.

We now show that LICQ in the directions $l$ and $-l$ is equivalent to restrictions on the binding moments in the problems $b^{\min}$ and $b^{\max}$.
Lemma A.1. Suppose $\beta_A = \delta_A + M_{\text{post}}\tau_{\text{post}}$ for some $\delta_A \in \Delta = \{\delta : A\delta \leq d\}$ and $\tau_{\text{post}} \in \mathbb{R}^T$. Then the following are equivalent:

(i) LICQ is satisfied in the direction $l$.

(ii) For any solution $\delta^{**}$ to the linear program

\[
\begin{aligned}
b^{\text{min}}(\delta_{\text{pre}}) &= \min_{\delta} l'\delta_{\text{post}} \\
\text{s.t.} & \quad A\delta \leq d, \delta_{\text{pre}} = \delta_{\text{pre}},
\end{aligned}
\]

the matrix $A(B(\delta^{**}),\text{post})$ with rows corresponding with the binding inequality constraints at $\delta^{**}$ has rank $|B(\delta^{**})|$.

The analogous results hold replacing $l$ with $-l$ in (i) and min with max in (ii).

Proof. We first show (i) implies (ii). Let $\delta^{**}$ be a solution to the minimization problem for $b^{\text{min}}$. Let $\tau^{**}_{\text{post}} = \beta_{\text{post}} - \delta^{**}_{\text{post}}$. Observe that $l'\tau^{**}_{\text{post}} = l'\beta_{\text{post}} - b^{\text{min}}(\delta_{\text{post}})$. From (10), we then see that $l'\tau^{**}_{\text{post}} = \theta^{ub}$ and hence $\tau^{**}_{\text{post}} \in S(l, T)$. Now, note that by construction, $m(\beta_A, \tau^{**}_{\text{post}}) = A(\beta_A - M_{\text{post}}\tau^{**}_{\text{post}} - d) = A\delta^{**} - d$, so the binding constraints in $m(\beta_A, \tau^{**}_{\text{post}})$ correspond with the binding constraints in the minimization for $b^{\text{min}}$. Finally, observe that $D_{\tau^{**}_{\text{post}}} m(\beta_A, \tau^{**}_{\text{post}}) = A(\cdot, \text{post})$. It then follows from (i) that the rows of $A(B(\delta^{**}),\text{post})$ are linearly independent, which gives the desired result.

Conversely, suppose $\tau^{**}_{\text{post}} \in S(l, \tau)$. By definition, there exists some $\delta^{**} \in \Delta$ such that $\delta^{**} = \beta_{\text{post}} - M_{\text{post}}\tau^{**}_{\text{post}}$, and $l'\tau^{**}_{\text{post}} = \theta^{ub}$. Thus, $\theta^{ub} = l'\beta_{\text{post}} - l'\delta^{**}$. It then follows from (10) that $l'\delta^{**} = b^{\text{min}}(\delta_{\text{pre}})$, so $\delta^{**}$ is a solution to the optimization $b^{\text{min}}$. (ii) then implies that $A(B(\delta^{**}),\text{post})$ has linearly independent rows. By the same argument as earlier in the proof, $A(B(\delta^{**}),\text{post})$ corresponds with the matrix of gradients for the binding constraints in $m(\beta_A, \tau^{**}_{\text{post}})$, from which we see that LICQ is satisfied.

The proof replacing $l$ with $-l$ and min with max is analogous.

Lemma A.1 makes clear that if LICQ holds in the directions $l$ and $-l$, then Assumption 3 is satisfied. Indeed, if LICQ holds, then Lemma 3 implies that the rank condition in Assumption 3 holds for any solutions $\delta^*$ and $\delta^{**}$ to the problems $b^{\text{max}}$ and $b^{\text{min}}$. By contrast, Assumption 3 only requires the rank condition to hold for at least one solution to $b^{\text{max}}$ and $b^{\text{min}}$. It is possible for a linear program to have multiple solutions, and for the rows of the binding constraints to be linearly independent (non-degenerate) for some solutions but not for others; Sierksma (2001) provides an example (Example 1, p. 146). Assumption 3 is thus potentially weaker than LICQ if there are multiple solutions to the problem for $b^{\text{max}}$ or $b^{\text{min}}$. 

S-2
B Constructing conditional-FLCI hybrid Confidence Sets

In this section, we provide further technical details on the construction of conditional-FLCI hybrid confidence sets, which are discussed in Section 7. To construct the conditional-FLCI hybrid confidence sets, we begin by providing a lemma that implies that the affine estimator at which the optimal FLCI is centered can be written as an affine function of $A\hat{\theta}$, where recall that $A$ is the matrix that defines the polyhedral set, $\Delta = \{ \delta : A\delta \leq d \}$.

**Lemma B.1.** Suppose $\Delta = \{ \delta : A\delta \leq d \} \neq \emptyset$, and $(a, v)$ are such that $b(a, v) < \infty$. Then, there exists $\tilde{v}$ such that $v' = \tilde{v}'A$.

**Proof.** Note that

$$b(a, v) = \max_{\delta, \tau_{\text{post}}} |v'(\delta + M_{\text{post}}\tau_{\text{post}}) - l'\tau|$$

s.t. $A\delta - d \leq 0$

We will show that if $b(a, v)$ is finite, then for all $\tilde{\delta} \in \Delta$, $A\tilde{\delta} = 0$ implies $v'\tilde{\delta} = 0$. This implies that $\tilde{\delta}$ is in the rowspace of $A$, from which the result follows. To prove this, suppose towards contradiction that $\tilde{\delta}$ is such that $A\tilde{\delta} = 0$ and $v'\tilde{\delta} \neq 0$. Since $A\tilde{\delta} = 0$, then, $\tilde{\delta} := (\delta_0 + c \cdot \tilde{\delta}) \in \Delta$ for any $\delta_0 \in \Delta$, and $c \in \mathbb{R}$. However, it then follows that for any fixed $\tau_{\text{post}}$ and $\delta_0$, the objective in the previous display can be made arbitrarily large at $(\delta_0, \tau_{\text{post}})$ by $c \to \infty$. 

Consider the level $1 - \kappa$ optimal FLCI, $C_{\kappa,n}^{FLCI} = a_n + v'\hat{\beta}_n \pm \chi_n$. By Lemma B.1, there exists some vector $\tilde{v}_n$ such that the level $1 - \kappa$ optimal FLCI can be written as $a_n + \tilde{v}'_n A\hat{\beta}_n \pm \chi_n$. Recalling further that $\tilde{Y}_n(\theta) = A\hat{\beta}_n - \bar{A}_{(\cdot,1)}\theta - d$, it immediately follows that the parameter value $\tilde{\theta} \in C_{\kappa,n}^{FLCI}$ if and only if

$$\tilde{v}'_n \tilde{Y}_n(\tilde{\theta}) \leq \chi_n - a_n - \tilde{v}'_n d + (1 - \tilde{v}'_n \bar{A}_{(\cdot,1)})\tilde{\theta},$$

$$-\tilde{v}'_n \tilde{Y}_n(\tilde{\theta}) \leq \chi_n + a_n + \tilde{v}'_n d - (1 - \tilde{v}'_n \bar{A}_{(\cdot,1)})\tilde{\theta}.$$ 

One can further show that $(1 - \tilde{v}'_n \bar{A}_{(\cdot,1)})\tilde{\theta} = 0$, which simplifies the upper bounds above.\(^{39}\) Defining the matrix $\tilde{V}_n = (\tilde{v}'_n, -\tilde{v}'_n)'$ and the vector $d_n(\tilde{\theta})$ which stacks the upper-bounds of the inequalities in the previous display, we see that the optimal level $1 - \kappa$ FLCI contains the parameter value $\tilde{\theta}$ if and only if $\tilde{V}_n \tilde{Y}_n(\tilde{\theta}) \leq d_n(\tilde{\theta})$.

With this equivalent representation of the optimal FLCI in hand, we are now able to characterize the distribution of the test statistic $\hat{\eta}$ defined earlier in (15) conditional on the parameter $\theta$ falling within the optimal FLCI.

---

\(^{39}\)Applying the definitions of $\hat{A}$ and $\tilde{v}_n$, we obtain that $\tilde{v}_n \bar{A}_{(\cdot,1)} = v_{n,\text{post}}\Gamma^{-1}e_1$. However, we show in the proof to Lemma E.22 that $v_{n,\text{post}} = l'$, so $\tilde{v}_n \bar{A}_{(\cdot,1)} = l'\Gamma^{-1}e_1$. The result then follows from the fact that $e_1\Gamma = l'$ by construction.
Lemma B.2.

\[ \frac{1}{\gamma} \left\{ \gamma \in \hat{V}_n, S_n = s, \bar{\theta} \in C_{\kappa,n}^{FLCI} \right\} \sim \frac{1}{\xi} \left\{ \xi \in [v_{FLCI}^{lo}, v_{FLCI}^{up}] \right\}, \]

where \( \xi \sim N \left( \gamma' \hat{\mu}, \gamma' \bar{\Sigma} \gamma \right) \), \( v_{FLCI}^{lo} := \max \{ v_{lo}, v_{FLCI}^{lo} \} \), \( v_{FLCI}^{up} := \min \{ v_{up}, v_{FLCI}^{up} \} \); \( v_{lo} \) and \( v_{up} \) are as defined in Section 4, and

\[
\begin{align*}
v_{FLCI}^{lo} &:= \max_{\{ j : (\bar{V}_n c_{n, \gamma} \}_j < 0 \}} \left\{ \bar{d}_n(\theta) - (\bar{V}_n S_n)_j \right\}, \\
v_{FLCI}^{up} &:= \min_{\{ j : (\bar{V}_n c_{n, \gamma} \}_j > 0 \}} \left\{ \bar{d}_n(\theta) - (\bar{V}_n S_n)_j \right\},
\end{align*}
\]

where \( c_{n, \gamma} = \frac{\bar{\Sigma} \gamma \gamma}{\gamma' \bar{\Sigma} \gamma} \) and \( S_n = (I - \frac{\bar{\Sigma} \gamma \gamma}{\gamma' \bar{\Sigma} \gamma} \gamma') \bar{Y}_n(\bar{\theta}) \), and \( \gamma \) is the vector of Lagrange multipliers for the primal problem (16).

Proof. The proof follows an analogous argument to Lemma 9 in ARP. Recall that conditional on \( \gamma \in \hat{V}_n, \hat{\eta} = \gamma' \hat{Y}_n \). Recall also that \( \theta \in C_{\kappa,n}^{FLCI} \) if and only if \( \bar{V}_n \hat{Y}_n \leq \bar{d}_n \). Next, observe that the set of values of \( \hat{Y}_n \) such that

\[
\gamma' \hat{Y}_n = \left( \max \gamma' \hat{Y}_n \text{ s.t. } \gamma \geq 0, \gamma' \hat{A}_{\kappa,n} = 0, \gamma' \bar{\sigma}_n = 1 \right)
\]

and \( \bar{V}_n \hat{Y}_n \leq \bar{d}_n \)

is convex. This follows from the fact that if the expression above holds for both \( \hat{Y}_n \) and \( \hat{Y}_n \), then \( \gamma' \hat{Y}_n \geq \gamma' \hat{Y}_n \) and \( \gamma' \hat{Y}_n \geq \gamma' \hat{Y}_n \) for all \( \gamma \) feasible in the maximization. It then follows that \( \gamma' (\alpha \hat{Y}_n + (1 - \alpha) \hat{Y}_n) \geq \gamma' (\alpha \hat{Y}_n + (1 - \alpha) \hat{Y}_n) \) for any \( \alpha \in (0,1) \). Thus, \( (\alpha \hat{Y}_n + (1 - \alpha) \hat{Y}_n) \) is also equal to the maximum. It is likewise clear that the second constraint holds for a convex combination of \( \hat{Y}_n \) and \( \hat{Y}_n \).

Thus, once we condition on \( S_n \), the set of values of \( \gamma' \hat{Y}_n \) such that \( \gamma \in \hat{Y}_n \) and \( \bar{V}_n \hat{Y}_n \leq \bar{d}_n \) is an interval. It thus suffices to find the endpoints. Without loss of generality, we focus on the lower bound. For ease of notation, let \( F := \{ \gamma \geq 0, \gamma' \hat{A}_{\kappa,n} = 0, \gamma' \bar{\sigma}_n = 1 \} \) denote the feasible region for the maximization. Then we are interested in

\[
\min_{\{ \hat{Y}_n : S_n = s \}} \left\{ \gamma' \hat{Y}_n : \gamma' \hat{Y}_n = \max_{\gamma \in F} \gamma' \hat{Y}_n, \bar{V}_n \hat{Y}_n \leq \bar{d}_n \right\}.
\]

Recalling that \( S_n = (I - c_{n, \gamma} \gamma') \hat{Y}_n \) for \( c_{n, \gamma} = \frac{\bar{\Sigma} \gamma \gamma}{\gamma' \bar{\Sigma} \gamma} \), the expression becomes

\[
\min_{\{ \hat{Y}_n : S_n = s \}} \left\{ \gamma' \hat{Y}_n : \gamma' \hat{Y}_n = \max_{\gamma \in F} \gamma' \left( s + c_{n, \gamma} \gamma' \hat{Y}_n \right), \bar{V}_n \hat{Y}_n \leq \bar{d}_n \right\}.
\]
which is equivalent to
\[
\min \left\{ x : x = \max_{\gamma \in F} \gamma' (s + cn, \gamma x) \right\} \cap \left\{ \gamma' Y_n : Y_n \text{ s.t. } S_n = s, \tilde{V}_n Y_n \leq \tilde{a}_n \right\}.
\]

However, the first set in the minimization above is the interval \([v_{lo}, v_{up}]\), and the polyhedral lemma in Lee et al. (2016) (Lemma 5.1) implies that second set is the interval \([v_{FLCI,lo}^{lo}, v_{FLCI,up}^{up}]\). Thus, the expression above is \(\max\{v_{lo}, v_{FLCI}^{lo}\}\), as desired. The argument for the lower bound of the interval is analogous. Finally, the independence of \(\gamma \tilde{Y}_n\) and \(S_n\) implies that the distribution of \(\gamma \tilde{Y}_n\) conditional on \(\{\gamma \in \hat{V}_n, S_n = s, \tilde{\theta} \in \mathcal{C}^{FLCI}_{n,\kappa}\}\) is truncated normal.

C Conditional Least-Favorable Hybrid Confidence Sets

In this section, we provide the technical details on the conditional least-favorable hybrid confidence sets that are briefly discussed in Section 7. The conditional-LF hybrid approach was originally introduced in ARP. These confidence sets are constructed by inverting a test that hybridizes the conditional test with tests using least-favorable critical values. We find that this hybridized test modestly improves upon the conditional test when the moments are not well-separated, while largely retaining the desirable properties of the conditional test when the moments are well-separated. As mentioned, we recommend this conditional least-favorable confidence set for choices of \(\Delta\) in which Assumption 4 is violated for all \(\delta \in \Delta\) and the FLCI is generally inconsistent, which includes \(\Delta^{RM1}(M)\).

C.1 Constructing conditional least-favorable hybrid confidence sets

Let \(\kappa \in (0, 1)\) and define \(\tilde{\alpha} = \frac{1 - \alpha}{1 - \kappa}\). The least favorable critical value \(c_{\kappa}^{LF}\) is the \(1 - \kappa\) quantile of
\[
\min_{\eta, \tilde{\eta}} \eta \text{ s.t. } \xi - \bar{X} \tilde{\eta} \leq \eta \tilde{\sigma}_n,
\]
for \(\xi \sim \mathcal{N}(0, \Sigma_n)\). That is, \(c_{\kappa}^{LF}\) is the \(1 - \kappa\) quantile of the distribution of \(\tilde{\eta}\) under the least-favorable assumption that \(E[\tilde{Y}(\tilde{\theta})] = 0\). Formally, we denote by \(\psi_{\kappa,\tilde{\alpha}}^{C-LF}(\bar{Y}_n(\tilde{\theta}), \bar{\Sigma}_n)\) an indicator for whether the \(1 - \alpha\) level conditional least-favorable hybrid test rejects when using a level \(1 - \kappa\) FLCI in the first-stage. We define \(\psi_{\kappa,\tilde{\alpha}}^{C-LF}\) by
\[
\psi_{\kappa,\tilde{\alpha}}^{C-LF}(\bar{Y}_n(\tilde{\theta}), \bar{\Sigma}_n) = 1 \iff F_{\xi | [v_{lo}, v_{C-LF}^{up}]}(\tilde{\eta}; \gamma_1 \Sigma_\gamma_s) > 1 - \tilde{\alpha},
\]
where \( v_{C-LF}^{up} = \min\{c_{n}^{LF}, v_{v}^{up}\} \) and \( F_{\xi|\xi \in [v_{lo}, v_{C-LF}^{up}]_n}(\cdot; \gamma^{\prime}_n \bar{S}_n \gamma_n) \) denotes the CDF of \( \xi \sim \mathcal{N}(0, \gamma^{\prime}_n \bar{S}_n \gamma_n) \) truncated to \([v_{lo}, v_{C-LF}^{up}]_n\).\(^{40}\) The hybrid test rejects whenever \( \hat{\eta} > c_{LF}^{\prime} \), and thus can be viewed as running a size-\( \kappa \) first-stage least favorable test, and then running a second-stage conditional test that conditions on both \( \gamma_n \) being an optimal vertex and the non-rejection of the least-favorable test. It follows immediately from Proposition 7 in ARP that the conditional least-favorable hybrid test controls size,

\[
\sup_{\delta \in \Delta, \tau \in \mathcal{S}(\Delta, \delta + \tau)} \mathbb{E}_{(\delta, \tau, \Sigma_n)} \left[ \psi^{C-LF}_{\kappa,\alpha} (\hat{Y}_n(\theta), \hat{\Sigma}_n) \right] \leq \alpha. \quad (32)
\]

We can thus construct a conditional least-favorable confidence set for the parameter \( \theta \) satisfying (11) by inverting this test,

\[
C_{C-LF}^{\kappa,\alpha,n} := \{ \theta : \psi^{C-LF}_{\kappa,\alpha} (\hat{Y}_n(\theta), \hat{\Sigma}_n) = 0 \}. \quad (33)
\]

### C.2 Asymptotic power of least-favorable conditional hybrid confidence sets

The least-favorable conditional hybrid confidence set maintains the asymptotic properties of the conditional confidence set under the same assumptions. Like the conditional test, these least-favorable conditional hybrid test are pointwise consistent. Additionally, the least-favorable conditional hybrid confidence sets satisfy a stronger notion of consistency in terms of expected excess length, where excess length is defined as the length of the part of the confidence set that falls outside the identified set

\[
EL(C; \delta_A, \tau_A) := \lambda(C \setminus \mathcal{S}(\Delta, \delta_A + \tau_A)), \quad (34)
\]

where \( \lambda(\cdot) \) denotes Lebesgue measure. The expected excess length of the least-favorable conditional hybrid confidence set converges to zero asymptotically.\(^{41}\)

**Proposition C.1.** The conditional least favorable hybrid test is pointwise consistent and consistent in terms of expected excess length. That is, for any \( \delta_A \in \Delta, \tau_A \in \mathbb{R}^T \) and \( \theta_{out} \notin \mathcal{S}(\Delta, \delta_A + \tau_A) \),

\[
\lim_{n \to \infty} \mathbb{P}_{\delta_A, \tau_A, \Sigma_n} (\theta_{out} \notin C_{C-LF}^{\kappa,\alpha,n}) = 1,
\]

\[
\lim_{n \to \infty} \mathbb{E}_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{C-LF}^{\kappa,\alpha,n}, \delta_A, \tau_A) \right] = 0.
\]

Similarly, the local asymptotic power of the hybrid test is at least as good as the optimum over tests of size \( \tilde{\alpha} = \frac{\alpha - \kappa}{1 - \alpha} \). Moreover, the expected excess length of the hybrid confidence set is at least

---

\(^{40}\)If \( v_{lo} > v_{C-LF}^{up} \), then we define \( F_{\xi|\xi \in [v_{lo}, v_{C-LF}^{up}]_n}(\hat{\eta}; \gamma^{\prime}_n \bar{S}_n \gamma_n) = 1 \), so that the test always rejects.

\(^{41}\)Due to its poor performance when the moments are not well-separated (Section 5.3), the expected excess length of the conditional confidence set need not be finite if the parameter space is unrestricted. If \( \theta \) is restricted to a compact set, the expected excess length of the conditional confidence set would be finite, and converge to zero asymptotically. Hybridization with the least-favorable critical value rectifies this issue without additionally assuming that the parameter space is compact.
as small as the optimum for level $1 - \tilde{\alpha}$ confidence sets.

**Proposition C.2.** Fix $\delta_A \in \Delta, \tau_A$, and $\Sigma^*$ positive definite. Suppose Assumption 3 holds. Let $\tilde{\alpha} = \frac{\alpha - \kappa}{1 - \kappa}$. Then,

$$
\liminf_{n \to \infty} \mathbb{P}_{(\delta_A, \tau_A, \Sigma_n)} \left( (\theta_A^b + \frac{1}{\sqrt{n}} x) \notin C^{C-LF}_{\kappa, \alpha, n} \right) \geq \lim_{n \to \infty} \sup_{C_{\tilde{\alpha}, n} \in I_{\tilde{\alpha}, n}(\Delta, \Sigma_n)} \mathbb{P}_{(\delta_A, \tau_A, \Sigma_n)} \left( (\theta_A^b + \frac{1}{\sqrt{n}} x) \notin C_{\tilde{\alpha}, n} \right).
$$

The analogous result holds replacing $\theta_A^b + \frac{1}{\sqrt{n}} x$ with $\theta_A^b - \frac{1}{\sqrt{n}} x$, for $\theta_A^b$ the lower bound of the identified set (although the constant $c^*$ may differ). Additionally,

$$
\limsup_{n \to \infty} \inf_{C_{\tilde{\alpha}, n} \in I_{\tilde{\alpha}, n}(\Delta, \Sigma^*)} \mathbb{E}_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C^{C-LF}_{\kappa, \alpha, n}; \delta_A, \tau_A) \right] \leq 1.
$$

Proposition C.2 implies that as the size of the first-stage least-favorable test becomes small, the limiting expected excess length of the hybridized ARP confidence set approaches the optimal bound. We state this as a corollary.

**Corollary C.1.** Suppose the assumptions of Proposition C.2 hold and let $\tilde{\alpha} = \frac{\alpha - \kappa}{1 - \kappa}$. Then,

$$
\limsup_{n \to \infty} \frac{\mathbb{E}_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C^{C-LF}_{\kappa, \alpha, n}; \delta_A, \tau_A) \right]}{\inf_{C_{\tilde{\alpha}, n} \in I_{\tilde{\alpha}, n}(\Delta, \Sigma_n)} \mathbb{E}_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{\tilde{\alpha}, n}; \delta_A, \tau_A) \right]} \leq \frac{z_{1-\tilde{\alpha}} (1 - \tilde{\alpha}) + \phi(z_{1-\tilde{\alpha}})}{z_{1-\alpha} (1 - \alpha) + \phi(z_{1-\alpha})}.
$$

This implies that

$$
\lim_{\kappa \to 0} \limsup_{n \to \infty} \frac{\mathbb{E}_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C^{C-LF}_{\kappa, \alpha, n}; \delta_A, \tau_A) \right]}{\inf_{C_{\tilde{\alpha}, n} \in I_{\tilde{\alpha}, n}(\Delta, \Sigma_n)} \mathbb{E}_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{\tilde{\alpha}, n}; \delta_A, \tau_A) \right]} = 1.
$$

ARP recommend using $\kappa = \alpha/10$ for conditional least-favorable hybrid confidence sets, following a similar recommendation from Romano et al. (2014). For $\alpha = 0.05$ and $\kappa = \alpha/10$, Corollary C.1 implies that the limiting expected excess length of the hybrid confidence set is within 3% of the optimal.

**D Optimal bounds on excess length**

In this section, we discuss the computation of optimal bounds on the (excess) length of confidence intervals that satisfy the uniform coverage requirement (11). In Section 8, we benchmark the performance of our proposed procedures in Monte Carlo simulations relative to these bounds.

The following result restates Theorem 3.2 of Armstrong and Kolesar (2018) in the notation of our paper, which provides a formula for the optimal expected length of a confidence set that satisfies the uniform coverage requirement.
Lemma D.1. Suppose that $\Delta$ is convex. Let $I_\alpha$ denote the set of confidence sets that satisfy the coverage requirement (11). Then, for any $\delta_A \in \Delta$ and $\tau_A \in \mathbb{R}^T$,

$$\inf_{\mathcal{C} \in I_\alpha} \mathbb{E}_{(\delta_A, \tau_A, \Sigma)_n}(\Lambda(C)) = (1 - \alpha)\mathbb{E}[\bar{\omega}(z_{1-\alpha} - Z) - \omega(z_{1-\alpha} - Z) | Z < z_{1-\alpha}],$$

where $Z \sim \mathcal{N}(0, 1)$, $z_{1-\alpha}$ is the $1 - \alpha$ quantile of $Z$, and

$$\bar{\omega}(b) := \sup\{l'\tau | \tau \in \mathbb{R}^T, \exists \delta \in \Delta \text{ s.t. } \|\delta + M_{post}\tau - \beta_A\|_{\Sigma_n}^2 \leq b^2\}$$

$$\omega(b) := \inf\{l'\tau | \tau \in \mathbb{R}^T, \exists \delta \in \Delta \text{ s.t. } \|\delta + M_{post}\tau - \beta_A\|_{\Sigma_n}^2 \leq b^2\},$$

for $\beta_A := \delta_A + M_{post}\tau_A$, and $\|x\|_{\Sigma} = x'\Sigma^{-1}x$.

The proof of this result follows from observing that the confidence set that optimally directs power against $(\delta_A, \tau_A)$ inverts Neyman-Pearson tests of $H_0 : \delta \in \Delta, \theta = \bar{\theta}$ against $H_A : (\delta, \tau) = (\delta_A, \tau_A)$ for each value $\bar{\theta}$. The formulas above are then obtained by integrating one minus the power function of these tests over $\bar{\theta}$. By the same argument, the optimal excess length for confidence sets that control size is the integral of one minus the power function over all points outside of the identified set. Additionally, for any value $\bar{\theta} \in S(\Delta, \beta_A)$, the null and alternative hypotheses are observationally equivalent, and so the most powerful test trivially has size $\alpha$. It follows that the lowest achievable expected excess length is $(1 - \alpha) \cdot \text{LID}(\Delta, \delta_{A,pre})$ shorter than the lower achievable expected length, where as in Section 6, LID denotes the length of the identified set.

Corollary D.1. Under the conditions of Lemma D.1,

$$\inf_{\mathcal{C} \in I_\alpha} \mathbb{E}_{(\delta_A, \tau_A, \Sigma)_n}(EL(C; \delta_A, \tau_A)) = \inf_{\mathcal{C} \in I_\alpha} \mathbb{E}_{(\delta_A, \tau_A, \Sigma)_n}(\Lambda(C)) - (1 - \alpha)\text{LID}(\Delta, \delta_{A,pre}).$$
E  Proofs of Main Finite Sample Normal Results

E.1  Proofs of Main Finite Sample Normal Results

Throughout the proofs, we will use the following notation. Let \( \tilde{Y}(\hat{\beta}_n, A, d, \tilde{\theta}) := A\hat{\beta}_n - d - \hat{A}_{(\cdot,1)}\tilde{\theta} \). Define \( \psi^C_\alpha(\hat{\beta}_n; A, d, \tilde{\theta}, \Sigma) := \psi^C_\alpha(\tilde{Y}(\hat{\beta}_n; A, d, \tilde{\theta}), A\Sigma A') \) to be an indicator for whether the conditional test constructed using \( (\hat{\beta}_n, A, d, \tilde{\theta}, \Sigma) \) rejects. Likewise, let \( \psi_{k,\alpha}^{C,LF}(\hat{\beta}_n; A, d, \tilde{\theta}, \Sigma) := \psi_{k,\alpha}^{C,LF}(\tilde{Y}(\hat{\beta}_n, A, d, \tilde{\theta}), A\Sigma A') \) and \( \psi^L_\alpha(\hat{\beta}_n, A, d, \tilde{\theta}, \Sigma) := 1\{ \hat{\eta} > c^L_{\alpha}(A\Sigma A') \} \) be indicators for whether the conditional least-favorable hybrid test and least-favorable test reject.

**Proof of Proposition 5.1**

Proof. We argued in Section 6.2 that the identified set is an interval, \( S(\Delta, \delta_A + \tau_A) = [\theta_{lb}, \theta_{ub}] \), and so if \( \theta_{out} \notin S(\Delta, \delta_A + \tau_A) \), then we must have either \( \theta_{out} = \theta_{ub} + x \) or \( \theta_{out} = \theta_{lb} - x \) for some \( x > 0 \). Without loss of generality, we consider the case where \( \theta_{out} = \theta_{ub} + x \). Using the notation introduced above, we have that

\[
\liminf_{n \to \infty} \mathbb{P}(\delta_{A,\tau_A},\Sigma_n) \left( \theta_{out} \notin C_{\alpha,n} \right) = \liminf_{n \to \infty} \mathbb{E}(\delta_{A,\tau_A},\Sigma_n) \left[ \psi^C_\alpha(\hat{\beta}_n; A, d, \theta_{ub} + x, \Sigma_n) \right].
\]

Lemma E.2 along with the fact that \( \Sigma_n = \frac{1}{n} \Sigma^* \) imply that \( \psi^C_\alpha(\hat{\beta}_n; A, d, \theta_{ub} + x, \Sigma_n) = \psi^C_\alpha(\sqrt{n}\hat{\beta}_n; A, \sqrt{n}d, \sqrt{n}\theta_{ub} + \sqrt{n}x, \Sigma^*) \). Thus,

\[
\liminf_{n \to \infty} \mathbb{P}(\delta_{A,\tau_A},\Sigma_n) \left( \theta_{out} \notin C_{\alpha,n} \right) = \liminf_{n \to \infty} \mathbb{E}(\delta_{A,\tau_A},\Sigma_n) \left[ \psi^C_\alpha(\sqrt{n}\hat{\beta}_n; A, \sqrt{n}d, \sqrt{n}\theta_{ub} + \sqrt{n}x, \Sigma^*) \right] = \liminf_{n \to \infty} \mathbb{E}(\sqrt{n}\delta_A,\sqrt{n}\tau_A,\Sigma^*) \left[ \psi^C_\alpha(\hat{\beta}_n; A, \sqrt{n}d, \sqrt{n}\theta_{ub} + \sqrt{n}x, \Sigma^*) \right],
\]

where in the second line we use the fact that \( \hat{\beta}_n \sim \mathcal{N}(\beta_A, \Sigma_n) \) implies \( \sqrt{n}\hat{\beta}_n \sim \mathcal{N}(\sqrt{n}\beta_A, \Sigma^*) \). Now, Lemma E.1 implies that \( \sqrt{n}\theta_{ub} = \theta_{ub} \), for \( \theta_{ub} = \sup S(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A) \) and \( \Delta_n = \{ \delta : A\delta \leq \sqrt{n}d \} \). It then follows from Lemma E.21 that

\[
\mathbb{E}(\sqrt{n}\delta_A,\sqrt{n}\tau_A,\Sigma^*) \left[ \psi^C_\alpha(\hat{\beta}_n; A, \sqrt{n}d, \sqrt{n}\theta_{ub} + \sqrt{n}x, \Sigma^*) \right] \geq \rho_{LB}(x, \Sigma^*),
\]

for \( \rho_{LB} \) a function with \( \lim_{x \to \infty} \rho_{LB}(x, \Sigma^*) = 1 \). It is then immediate from the previous two displays that

\[
\lim_{n \to \infty} \mathbb{P}(\delta_{A,\tau_A},\Sigma_n) \left( \theta_{out} \notin C_{\alpha,n} \right) = 1,
\]

as desired.

\hspace{1cm} \Box

**Proof of Proposition 5.2**
Proof. We will show that each of the limits in the statement of the proposition is equal to \( \Phi(c^* x - z_{1-\alpha}) \). The first limit of interest can be written as

\[
\lim_{n \to \infty} \mathbb{P}(\delta_{A, \tau_A, \Sigma_n}) \left( \theta^{ub} + \frac{1}{\sqrt{n}} x \right) \notin C_{\alpha,n} = \lim_{n \to \infty} \mathbb{E}(\delta_{A, \tau_A, \Sigma_n}) \left[ \psi^C_{\alpha} (\hat{\beta}_n; A, d, \theta^{ub} + \frac{1}{\sqrt{n}} x, \Sigma_n) \right].
\]

By Lemma E.2, we have that \( \psi^C_{\alpha} (\hat{\beta}_n; A, d, \theta^{ub} + \frac{1}{\sqrt{n}} x, \Sigma_n) = \psi^C_{\alpha} (\sqrt{n} \hat{\beta}_n; A, \sqrt{n} d, \sqrt{n} \theta^{ub} + x, \Sigma^*) \), where we use the fact that \( \Sigma_n = \frac{1}{n} \Sigma^* \). It follows that

\[
\lim_{n \to \infty} \mathbb{P}(\delta_{A, \tau_A, \Sigma_n}) \left( \theta^{ub} + \frac{1}{\sqrt{n}} x \right) \notin C_{\alpha,n} = \lim_{n \to \infty} \mathbb{E}(\delta_{A, \tau_A, \Sigma_n}) \left[ \psi^C_{\alpha} (\sqrt{n} \hat{\beta}_n; A, \sqrt{n} d, \sqrt{n} \theta^{ub} + x, \Sigma^*) \right]
\]

where in the second line we use the fact that \( \hat{\beta}_n \sim \mathcal{N} (\beta_A; \frac{1}{n} \Sigma^*) \) implies \( \sqrt{n} \hat{\beta}_n \sim \mathcal{N} (\sqrt{n} \beta_A; \Sigma^*) \). Lemma E.1 implies that \( \sqrt{n} \theta^{ub} = \theta^{ub}_n \), where we define \( \Delta_n = \{ \delta : A \delta \leq \sqrt{n} d \} \) and \( \theta^{ub}_n := \sup \mathcal{S}(\Delta_n, \sqrt{n} \delta_A + \sqrt{n} \tau_A) \). From this, we see that

\[
\lim_{n \to \infty} \mathbb{P}(\delta_{A, \tau_A, \Sigma_n}) \left( \theta^{ub} + \frac{1}{\sqrt{n}} x \right) \notin C_{\alpha,n} = \lim_{n \to \infty} \mathbb{E}(\sqrt{n} \delta_{A, \tau_A, \Sigma^*}) \left[ \psi^C_{\alpha} (\hat{\beta}_n; A, \sqrt{n} d, \theta^{ub}_n + x, \Sigma^*) \right]
\]

which converges to \( \Phi(c^* x - z_{1-\alpha}) \) by Lemma E.9.

We next turn our attention to the second limit. Consider the problem of testing \( H_0 : \delta \in \Delta = \{ \delta : A \delta \leq d \}, \theta = \theta \) against \( H_1 : (\delta, \tau) = (\delta_A, \tau_A) \). Observe that the null is equivalent to the null \( H_0 : \beta \in \mathcal{B}_0(\hat{\theta}) := \{ \beta : 3 \tau_{\text{post}} \text{ s.t. } l' \tau_{\text{post}} = \theta, A \beta - d - AM_{\text{post}} \tau_{\text{post}} \leq 0 \} \). Likewise, the alternative that \( (\delta, \tau) = (\delta_A, \tau_A) \) is equivalent to \( H_1 : \beta = \delta_A + \tau_A =: \beta_A \). It is clear from the definition of \( \mathcal{B}_0 \) that it is convex. It then follows from Lemma E.6 that the most powerful test that controls size is a one-sided t-test (Neyman-Pearson) that rejects for large values of \( (\beta_A - \beta_n) / \sqrt{n} \beta_n \), where \( \beta := \arg \min_{\beta \in \mathcal{B}_0(\hat{\theta})} || \beta_A - \beta || / \Sigma_n \). We will define \( \psi^M_{\alpha} (\beta_n; A, d, \hat{\theta}, \Sigma_n, \delta_A, \tau_A) \) to be an indicator for whether the Neyman-Pearson test rejects \( H_0 \) in favor \( H_1 \) given a draw \( \hat{\beta}_n \) that is assumed to be normally distributed with covariance \( \Sigma_n \). The second limit can thus be written as

\[
\lim_{n \to \infty} \sup_{C_{\alpha,n} \in \mathcal{L}_n (\Delta, \frac{1}{n} \Sigma^*)} \mathbb{P}(\delta_{A, \tau_A, \Sigma_n}) \left( \theta^{ub} + \frac{1}{\sqrt{n}} x \right) \notin C_{\alpha,n} = \lim_{n \to \infty} \mathbb{E}(\delta_{A, \tau_A, \Sigma_n}) \left[ \psi^M_{\alpha} (\hat{\beta}_n; A, d, \theta^{ub} + \frac{1}{\sqrt{n}} x, \Sigma_n, \delta_A, \tau_A) \right].
\]

From Lemma E.7 and the fact that \( \Sigma_n = \frac{1}{n} \Sigma^* \), we have that \( \psi^M_{\alpha} (\hat{\beta}_n; A, d, \theta^{ub} + \frac{1}{\sqrt{n}} x, \Sigma_n, \delta_A, \tau_A) = \psi^M_{\alpha} (\sqrt{n} \hat{\beta}_n; A, \sqrt{n} d, \sqrt{n} \theta^{ub} + x, \Sigma^*, \sqrt{n} \delta_A, \sqrt{n} \tau_A) \). It follows that
\[
\lim_{n \to \infty} \sup_{c_{\alpha,n} \in \mathcal{C}_{\alpha,n}} \mathbb{P}(\delta_{A,\tau_A,\Sigma_n}) \left( (\theta^{ub} + \frac{1}{\sqrt{n}} x) \notin C_{\alpha,n} \right) = \\
\mathbb{E}(\delta_{A,\tau_A,\Sigma_n}) \left[ \psi^{MP}_\alpha (\sqrt{n} \hat{\beta}_n; A, \sqrt{n} \theta^{ub} + x, \Sigma^*; \sqrt{n} \delta_A, \sqrt{n} \tau_A) \right] = \\
\mathbb{E}(\sqrt{n} \delta_{A,\tau_A,\Sigma^*}) \left[ \psi^{MP}_\alpha (\beta; A, \sqrt{n} \theta^{ub} + x, \Sigma^*; \sqrt{n} \delta_A, \sqrt{n} \tau_A) \right],
\]

where the second equality again uses the fact that if \( \hat{\beta}_n \sim \mathcal{N}(\beta_A, \frac{1}{n} \Sigma^*) \), then \( \sqrt{n} \hat{\beta}_n \sim \mathcal{N}(\sqrt{n} \beta_A, \Sigma^*) \).

Recalling from earlier in the proof that \( \theta^{ub} = \sqrt{n} \theta^{ub} \), we obtain that

\[
\lim_{n \to \infty} \sup_{c_{\alpha,n} \in \mathcal{C}_{\alpha,n}} \mathbb{P}(\delta_{A,\tau_A,\Sigma_n}) \left( (\theta^{ub} + \frac{1}{\sqrt{n}} x) \notin C_{\alpha,n} \right) = \\
\mathbb{E}(\sqrt{n} \delta_{A,\tau_A,\Sigma^*}) \left[ \psi^{MP}_\alpha (\hat{\beta}_n; A, \sqrt{n} \theta^{ub} + x, \Sigma^*; \sqrt{n} \delta_A, \sqrt{n} \tau_A) \right],
\]

which converges to \( \Phi(c^* x - z_{1-\alpha}) \) by Lemma E.13.

\[\square\]

**Proof of Proposition 6.1**

*Proof.* First, suppose Assumption 4 holds. Without loss of generality, we show \( \mathbb{P}((\theta^{ub} + x) \in C_{\alpha,n}) \to 0 \) for any \( x > 0 \). By Lemma E.23 there exists \((\bar{a}, \bar{v})\) such that \( \bar{b}(\bar{a}, \bar{v}) = \frac{1}{2} LID(\Delta, \delta_{pre}) =: \bar{b}_{min} \) and \( \mathbb{E}(\delta_{A,\tau_A,\Sigma_n}) \left[ \bar{a} + \bar{v} \hat{\beta}_n \right] = \frac{1}{2} (\theta^{ub} + \theta^{lbb}) =: \theta^{mid} \).

Let \( \tilde{c}_n := \bar{a} + \bar{v} \hat{\beta}_n \pm \chi_n(\bar{a}, \bar{v}) \) denote the fixed length confidence interval based on \((\bar{a}, \bar{v})\).

Recall that by construction, \( \tilde{\chi}_n := \chi_n(\bar{a}, \bar{v}) \) is the \( 1 - \alpha \) quantile of the \( |\mathcal{N}(\bar{b}_{min}, \sigma^2_{v,n})| \) distribution. Since \( \sigma^2_{v,n} = \frac{1}{n} \sigma^2_{v,1} \to 0 \), the \( \mathcal{N}(\bar{b}_{min}, \sigma^2_{v,n}) \) distribution collapses to a point mass at \( \bar{b}_{min} \), and thus \( \tilde{\chi}_n \to \bar{b}_{min} \). By construction the length of the shortest FLCI \( \chi_n := \chi_n(a_n, v_n) \) must be less than or equal to \( \tilde{\chi}_n \), and so \( \limsup_{n \to \infty} \chi_n \leq \bar{b}_{min} \). Let \( b_n := \bar{b}(a_n, v_n) \) be the worst-case bias of the optimal FLCI. Since \( \alpha \in (0, 0.5] \), Lemma E.24 implies that \( \chi_n \geq b_n \). Additionally, Lemma E.22 implies that \( b_n \geq \frac{1}{2} LID(\Delta, \delta_{pre}) = \bar{b}_{min} \), and thus \( \chi_n \geq \bar{b}_{min} \). Hence, \( \chi_n \to \bar{b}_{min} \) implies \( b_n \to \bar{b}_{min} \). Additionally, note that for \( \alpha \in (0, 0.5] \), \( \chi_n(a, v) \) is increasing in both \( \bar{b}(a, v) \) and \( \sigma_{v,n} \).

Since \( \bar{b}_{min} \leq b_n \) and \( \tilde{\chi}_n \leq \chi_n \), it must be that \( \sigma_{v,n} \leq \sigma_{\tilde{v},n} \), from which it follows that \( \sigma_{\tilde{v},n} \to 0 \).

Now, we claim that \( \mu_n := \mathbb{E}(\delta_{A,\tau_A,\Sigma_n}) \left[ a_n + v_n' \hat{\beta}_n \right] \) converges to \( \theta^{mid} := \frac{1}{2} (\theta^{ub} + \theta^{lbb}) \). To show this, note that \( \mu_n = a_n + v_n' \beta_A \) for \( \beta_A = \delta_A + \tau_A \). Since \( \theta^{ub}, \theta^{lb} \in S(\Delta, \beta_A) \), by the definition of the identified set there exist \( \delta^{ub}, \delta^{lb} \in \Delta \) and \( \tau^{ub}, \tau^{lb} \) such that \( \beta_A = \delta^{ub} + \tau^{ub} = \delta^{lb} + \tau^{lb}, \theta^{ub} = \theta_{ub}^{mid}, \theta^{lb} = \theta_{lb}^{mid} \), and \( \theta^{ub} = \theta_{ub}^{mid} \).

Thus, \( \theta^{ub} - \mathbb{E}(\delta^{ub}, \tau^{ub}, \Sigma_n) \left[ a_n + v_n' \hat{\beta}_n \right] = \theta^{ub} - \mu_n \) and \( \theta^{lb} = \theta_{lb}^{mid}, \theta^{ub} = \theta_{ub}^{mid} \).

This implies that \( b_n \geq \max \{ \theta^{ub} - \mu_n, \mu_n - \theta^{lb} \} = \bar{b}_{min} + | \mu_n - \theta^{mid} | \), where the equality uses the fact that \( \theta^{ub} - \theta^{lb} = LID(\Delta, \delta_{A,pre}) = 2 \bar{b}_{min} \). Since we’ve shown that \( b_n \to \bar{b}_{min} \), it follows that \( \mu_n \to \theta^{mid} \), as desired.

Next, note that if \( \hat{\beta}_n \sim \mathcal{N}(\beta_A + \tau_A, \Sigma_n) \), then \( a_n + v_n' \hat{\beta}_n \sim \mathcal{N}(\mu_n, \sigma^2_{v,n}) \). Observe that \( \bar{\theta} \in C^*_{\alpha,n} \) if and only if \( a_n + v_n' \bar{\beta}_n \in [\bar{\theta} - \chi_n, \bar{\theta} + \chi_n] \). Thus,
By Lemma E.22, summand converges to

\[ \lim_{n \to \infty} \frac{\theta + \chi_n - \mu_n}{\sigma_{v_n,n}} = \Phi \left( \frac{\theta - \chi_n - \mu_n}{\sigma_{v_n,n}} \right). \]

Now, recalling that \( \theta^{ub} = \theta^{mid} + \bar{b}_{min} \) by construction, we have

\[
\mathbb{P}(\delta_{A}, \bar{\tau}_{A}, \Sigma_n) \left( (\theta^{ub} + x) \in \mathcal{C}_{\alpha,n}^{FLCI} \right) = \Phi \left( \frac{\theta^{mid} + \bar{b}_{min} + x + \chi_n - \mu_n}{\sigma_{v_n,n}} \right) - \Phi \left( \frac{\theta^{mid} + \bar{b}_{min} + x - \chi_n - \mu_n}{\sigma_{v_n,n}} \right).
\]

Note that the term inside the second normal CDF in the previous display is equal to

\[ -\frac{\chi_n - b_n}{\sigma_{v_n,n}} + \frac{x + \theta^{mid} - \mu_n + \bar{b}_{min} - b_n}{\sigma_{v_n,n}}. \]

However, the first summand above is bounded between \(-z_{1-\alpha/2}\) and \(-z_{1-\alpha}\) by Lemma E.24. Additionally, we’ve shown that the \( \theta^{mid} - \mu_n \to 0 \) and \( \bar{b}_{min} - b_n \to 0 \), so the numerator of the second summand converges to \( x > 0 \). Since the denominator \( \sigma_{v_n,n} \to 0 \), the expression in the previous display diverges to \( \infty \), and hence the second normal CDF term in (35) converges to 1, which implies that \( \mathbb{P}(\theta^{ub} + x) \in \mathcal{C}_{\alpha,n}^{FLCI} \to 0 \), as needed.

Conversely, suppose Assumption 4 fails. Let \( L_A := LID(\Delta, \delta_{A,pre}) \) and \( \bar{L} := \sup_{\delta_{pre} \in \Delta_{pre}} LID(\Delta, \delta_{pre}) \). By Lemma E.22, \( b_n := \bar{b}(a_n, v_n) \geq \frac{1}{2} \bar{L} =: \bar{b}_{min} \). As argued earlier in the proof, since \( \alpha \in (0, 0.5) \), \( \chi_n \geq b_n = \frac{1}{2} \bar{L} \). If \( \bar{L} = \infty \), then \( \mathcal{C}_{\alpha,n}^{FLCI} \) is the real line, and thus never rejects, so \( \mathcal{C}_{\alpha,n}^{FLCI} \) is trivially inconsistent under the assumption that \( \mathcal{S}_{\theta}(\Delta, \delta_A + \tau_A) \neq \mathbb{R} \). For the remainder of the proof, we assume \( L_A < \bar{L} < \infty \). Now, as discussed in Section 6.2, we can write \( \mathcal{S}(\Delta, \delta_{A} + \tau_{A}) = [\theta^{lb}, \theta^{ub}] \), where \( \theta^{ub} - \theta^{lb} = LID(\Delta, \delta_{A,pre}) = L_A \). Let \( \epsilon = \frac{1}{4} (\bar{L} - L_A) \), and set \( \theta^{out} := \theta^{ub} + \epsilon \) and \( \theta^{out} = \theta^{lb} - \epsilon \). Let \( \theta^{mid} := \frac{1}{2}(\theta^{lb} + \theta^{ub}) \) be the midpoint of the identified set. By construction, \( \theta^{out} - \theta^{mid} = \theta^{mid} - \theta^{out} = \frac{1}{2} L_A \) and \( \epsilon < \frac{1}{2} \bar{L} \). Since \( \mathcal{C}_{\alpha,n}^{FLCI} \) is an interval with half-length at least \( \frac{1}{2} \bar{L} \), it follows that if \( \theta^{mid} \in \mathcal{C}_{\alpha,n}^{FLCI} \) then at least one of \( \theta^{out} \) is also in \( \mathcal{C}_{\alpha,n}^{FLCI} \). Hence,

\[
\mathbb{P} \left( \theta^{out} \in \mathcal{C}_{\alpha,n}^{FLCI} \right) \geq \mathbb{P} \left( \theta^{mid} \in \mathcal{C}_{\alpha,n}^{FLCI} \right) \geq 1 - \alpha,
\]

where the final bound follows since \( \mathcal{C}_{\alpha,n}^{FLCI} \) controls size. It follows that \( \limsup_{n \to \infty} \mathbb{P} \left( \theta^{out} \in \mathcal{C}_{\alpha,n}^{FLCI} \right) \geq \frac{1}{2} (1 - \alpha) > 0 \) for at least one \( j \in \{1, 2\} \).

**Proof of Proposition 6.2**

**Proof.** Following the same argument as in the proof to Proposition 6.1, we can show that

\[
\mathbb{P}(\delta_{A}, \bar{\tau}_{A}, \Sigma_n) \left( (\theta^{ub} + \frac{x}{\sqrt{n}}) \in \mathcal{C}_{\alpha,n}^{FLCI} \right) = \Phi \left( \frac{\theta^{mid} + \bar{b}_{min} + \frac{x}{\sqrt{n}} + \chi_n - \mu_n}{\sigma_{v_n,n}} \right) - \Phi \left( \frac{\theta^{mid} + \bar{b}_{min} + \frac{x}{\sqrt{n}} - \chi_n - \mu_n}{\sigma_{v_n,n}} \right),
\]
where $\bar{b}_{\min} = \frac{1}{2}LID(\Delta, \delta_{A, \text{pre}})$, $\mu_n = a_n + v'_n(\delta_A + \tau_A)$, and $\theta^{\text{mid}}$ is the midpoint of $S(\Delta, \delta_A + \tau_A)$.

Note that the term inside the second normal CDF in the previous display is equal to

$$- \frac{\chi_n - b_n}{\sigma_{v,n}} + \frac{x}{\sqrt{n} \sigma_{v,n}} + \frac{\theta^{\text{mid}} - \mu_n + \bar{b}_{\min} - b_n}{\sigma_{v,n}}. \tag{36}$$

We first show that the first term in the previous display converges to $-z_{1-\alpha}$. Since $\chi_n$ is the $1-\alpha$ quantile of the $|N(b_n, \sigma^2_{v,n})|$ distribution, we have that

$$\Phi\left(\frac{\chi_n - b_n}{\sigma_{v,n}}\right) - \Phi\left(\frac{-\chi_n - b_n}{\sigma_{v,n}}\right) = 1 - \alpha.$$

Lemma E.28 implies that $\bar{b}_{\min} = \frac{1}{2} \sup_{\delta_{\text{pre}}} LID(\Delta, \tilde{\delta}_{\text{pre}}) > 0$. We argued in the proof to Proposition 6.1 that $b_n \geq \bar{b}_{\min} > 0$, $\chi_n \geq 0$, and $\sigma_{v,n} \to 0$, from which we see that the term inside the second normal CDF in the previous display diverges to $-\infty$, and thus the second normal CDF term converges to 0. It follows that the first CDF term in the previous display converges to $1 - \alpha$, and hence

$$\frac{\chi_n - b_n}{\sigma_{v,n}} \to z_{1-\alpha}.$$

Next, we show that the second term in (36) converges to $c^*x$, for the constant $c^*$ defined in Proposition 5.2. Lemma E.26 implies that $\lim_{n \to \infty} \frac{x}{\sqrt{n} \sigma_{v,n}} = \lim_{n \to \infty} \frac{x}{\sqrt{n} \sigma_{\bar{v},n}} = \frac{x}{\sigma_{\bar{v},1}}$, where $\bar{v}$ is the unique value such that there exists $(\bar{a}, \bar{v})$ with $b(\bar{a}, \bar{v}) = \bar{b}_{\min}$. Moreover, Lemma E.25 implies that $1/\sigma_{\bar{v},1} = c^*$, from which we see that the limit of the second term is $c^*x$, as desired.

Now, we claim that the third term in (36) converges to 0. We argued in the proof to Proposition 6.1 that $|\mu_n - \theta^{\text{mid}}| \leq b_n - \bar{b}_{\min}$. It thus suffices to show that $\frac{b_n - \bar{b}_{\min}}{\sigma_{v,n}} \to 0$. Lemma E.25 implies that there is a unique pair $(\bar{a}, \bar{v})$ such that $b(\bar{a}, \bar{v}) = \bar{b}_{\min}$. Let $\bar{\chi}_n = \chi_n(\bar{a}, \bar{v})$ and $\chi_n = \chi_n(a_n, v_n)$. Note that $\chi_n \leq \bar{\chi}_n$ by construction, and $b_n \geq \bar{b}_{\min}$ by Lemma E.22. Hence, using the bounds from Lemma E.24, we have that

$$b_n + \sigma_{v,n} z_{1-\alpha} \leq \chi_n \leq \bar{\chi}_n = \sigma_{\bar{v},n} c\alpha \frac{\bar{b}_{\min}}{\sigma_{\bar{v},n}},$$

which, along with the inequality $b_n \geq \bar{b}_{\min}$, implies that

$$0 \leq \frac{b_n - \bar{b}_{\min}}{\sigma_{v,n}} \leq \frac{\sigma_{\bar{v},n} c\alpha}{\sigma_{v,n}} \left( \frac{\bar{b}_{\min}}{\sigma_{\bar{v},n}} \right) - \left( z_{1-\alpha} + \frac{\bar{b}_{\min}}{\sigma_{v,n}} \right)$$

$$= \left[ c\alpha \frac{\bar{b}_{\min}}{\sigma_{\bar{v},n}} - \left( z_{1-\alpha} + \frac{\bar{b}_{\min}}{\sigma_{v,n}} \right) \right] + \left[ \left( \frac{\sigma_{\bar{v},n}}{\sigma_{v,n}} - 1 \right) c\alpha \frac{\bar{b}_{\min}}{\sigma_{\bar{v},n}} - \frac{\bar{b}_{\min}}{\sigma_{v,n}} \right].$$

S-13
The first bracketed expression in the upper bound above converges to 0 by Lemma E.27. Applying the upper bound from Lemma E.24 to the $c v_\alpha$ term in the second bracketed expression, we obtain that the second bracketed expression is bounded above by $(\sigma_{\phi,n} - 1) z_{1-\alpha/2}$, which converges to 0 by Lemma E.26.

Combining the results above, we see that the expression in (36) converges to $c^* x - z_{1-\alpha}$. It follows that $\limsup_{n \to \infty} \mathbb{P} \left( (\theta_{A}^{ab} + \frac{x}{\sqrt{n}}) \notin C^{FLCI}_{\alpha,n} \right) \leq 1 - \Phi(c^* x - z_{1-\alpha})$, and hence $\liminf_{n \to \infty} \mathbb{P} \left( (\theta_{A}^{ab} + \frac{x}{\sqrt{n}}) \notin C^{FLCI}_{\alpha,n} \right) \geq \Phi(c^* x - z_{1-\alpha})$. Proposition 5.2 gives that $\Phi(c^* x - z_{1-\alpha})$ is the optimal local asymptotic power over procedures that control size, from which the result follows.

**Proof of Proposition 7.1**

**Proof.** The proof follows from the same argument as for Proposition 5.1, replacing Lemma E.2 with Lemma E.4, and Lemma E.19 with Lemma E.20.

**Proof of Proposition 7.2**

**Proof.** By an invariance to scale argument analogous to that in the proof of Proposition 5.2,

$$\liminf_{n \to \infty} \mathbb{P}_{(\delta_A, \tau_A, \Sigma_A)} \left( (\theta_{A}^{ab} + \frac{x}{\sqrt{n}}) \notin C^{FLCI}_{\alpha,n} \right) = \liminf_{n \to \infty} \mathbb{E}_{(\sqrt{\eta}\delta_A, \sqrt{\eta}\tau_A, \Sigma_A)} \left[ \psi_{\alpha}^{C_{FLCI}}(\hat{\beta}_n, A, \sqrt{\eta}d, \theta_{n}^{ub} + x, \Sigma^*) \right].$$

Proposition 5.2 implies that

$$\liminf_{n \to \infty} \mathbb{E}_{(\sqrt{\eta}\delta_A, \sqrt{\eta}\tau_A, \Sigma_A)} \left[ \psi_{\alpha}^{C_{FLCI}}(\hat{\beta}_n, A, \sqrt{\eta}d, \theta_{n}^{ub} + x, \Sigma^*) \right] = \limsup_{n \to \infty} \mathbb{P}_{(\delta_A, \tau_A, \Sigma_A)} \left( (\theta_{A}^{ab} + \frac{x}{\sqrt{n}}) \notin C_n \right).$$

It thus suffices to show that

$$\liminf_{n \to \infty} \mathbb{E}_{(\sqrt{\eta}\delta_A, \sqrt{\eta}\tau_A, \Sigma_A)} \left[ \psi_{\alpha}^{C_{FLCI}}(\hat{\beta}_n, A, \sqrt{\eta}d, \theta_{n}^{ub} + x, \Sigma^*) \right] \geq \liminf_{n \to \infty} \mathbb{E}_{(\sqrt{\eta}\delta_A, \sqrt{\eta}\tau_A, \Sigma_A)} \left[ \psi_{\alpha}^{C_{FLCI}}(\hat{\beta}_n, A, \sqrt{\eta}d, \theta_{n}^{ub} + x, \Sigma^*) \right].$$

Note that the second stage of the test $\psi_{\alpha}^{C_{FLCI}}$ is nearly identical to $\psi_{\alpha}^{c_{FLCI}}$ except it uses $v_{C_{FLCI}}^{lo} := \max\{v_{F_{C_{FLCI}}}^{lo}, v_{FLCI}^{lo}\}$ and $v_{C_{FLCI}}^{up} := \min\{v_{F_{C_{FLCI}}}^{up}, v_{FLCI}^{up}\}$ instead of $v^{lo}$ and $v^{up}$. Since $F_{\tilde{\xi}; \tilde{\xi}_0^{lo}, \tilde{\xi}_0^{up}}(\tilde{\eta})$ is decreasing in $v^{lo}$ and $v^{up}$, it suffices to show that $v_{C_{FLCI}}^{lo} \to p - \infty$, where $P_n$ denotes the sequence of distributions under which $(\delta, \tau, \Sigma) = (\sqrt{\eta}\delta_A, \sqrt{\eta}\tau_A, \Sigma^*)$.

Let $\Delta_n = \{ \delta : A \delta \leq \sqrt{\eta}d \}$. Let $\nu_n = \nu_n(\Delta_n, \Sigma^*)$ and $\tilde{v}_n = v_n(\Delta, \Sigma_n)$. Define $a_n$ and $\bar{a}_n$, and $\tilde{\chi}_n$ and $\chi_n$ analogously. By Lemma E.3, $\nu_n = \nu_n(a_n, \bar{a}_n)$, $\chi_n = \sqrt{\eta}\tilde{\chi}_n$, and $b(a, v; \Delta_n) = \sqrt{\eta}b(\bar{a}, v, \Delta)$. We argued in the proof to Lemma E.26 that $\tilde{v}_n \to \tilde{v}$. Further, we showed in the proof to Lemma E.25 that $\tilde{v} = -\tilde{\gamma} A$, where $\tilde{\gamma}_{-B} = 0$ and $\tilde{\gamma}_{B}$ is the unique vector such that $\tilde{\gamma}'_{B} \hat{A}_{(B,1)} = 1$, $\tilde{\gamma}'_{B} \hat{A}_{(B,-1)} = 0$. Likewise, we argued in the proof to Lemma E.26 that $\bar{a}_n \to \bar{a}$, for $\bar{a}$ the unique value such that $\tilde{b}(\bar{a}, \tilde{v}; \Delta) = b_{min}(\Delta)$. We also showed in the proof to Proposition 6.1 that $\tilde{\chi}_n \to b_{min}(\Delta)$. 

S-14
Let $\tilde{v}_n$ be a vector such that $\tilde{v}'_n A = v'$ (which exists by Lemma B.1). Observe that

$$\tilde{v}^*_n \tilde{\Sigma} \gamma = \left( \tilde{v}'_n A \Sigma^* A' \gamma \right) \rightarrow \tilde{v}' \Sigma^* A' \gamma = -\gamma' A \Sigma^* A' \gamma = -\gamma' \tilde{\Sigma} \gamma,$$

where we use the fact that $v_n \rightarrow \tilde{v} = -\gamma' A$ as shown above, and the identity $\tilde{\Sigma} = A \Sigma^* A'$.

Now, Lemma E.30 implies that there is a constant $c > 0$ such that, with probability approaching one under $P_n$, $c \cdot \gamma$ is an optimal vertex of the dual problem for $\psi_{CFLCI}^e(\beta_n, A, \sqrt{nd}, \theta_{\tau n}^{ub} + x, \Sigma^*)$. Observe from Lemma B.2 that if $\gamma_a$ is an optimal vertex, and $\frac{v_n' \Sigma \gamma_a}{\gamma_a' \Sigma \gamma} < 0$, then the value of $v_{FLCI}^{lo}$ used in $\psi_{CFLCI}^e(\beta_n, A, \sqrt{nd}, \theta_{\tau n}^{ub} + x, \Sigma^*)$ is

$$v_{FLCI}^{lo} = \frac{d_{n,1} - v_n' \left( I - \frac{\hat{\Sigma} \gamma_a}{\gamma_a' \Sigma \gamma} \right) \tilde{Y}_n}{v_n' \Sigma \gamma_a} = -\left( v_n' \tilde{Y}_n - d_{n,1} \right) + \gamma_a' \tilde{Y}_n,$$

(37)

where $\tilde{Y}_n = A \beta_n - \sqrt{nd} - \tilde{A}(\gamma_{\tau n}) A(\theta_{\tau n}^{ub} + x)$. Since $c \gamma$ is optimal in the dual with probability approaching one under $P_n$, and

$$\frac{v_n' \Sigma \gamma_a}{c^2 \gamma' \Sigma \gamma} \rightarrow -\frac{1}{c} < 0$$

(38)

by the argument above, we have that with probability approaching 1 under $P_n$,

$$v_{FLCI}^{lo} = \frac{-\left( v_n' \tilde{Y}_n - d_{n,1} \right)}{v_n' \Sigma \gamma_a} + c \gamma' \tilde{Y}_n.$$

(39)

Now, we showed in the proof to Lemma E.9 that $\mathbb{E}_{(\sqrt{nd} a, \sqrt{nd} \tau A, \Sigma^*)} \left[ \tilde{Y}_{n,B} \right] = -\tilde{A}(B,1) x$ regardless of $n$, where $\tilde{Y}_n = \tilde{Y}_n - \tilde{A}(\gamma_{\tau n}) A(\theta_{\tau n}^{ub} + x)$ for a vector $\tilde{\gamma}_{1}^{ub}$. Since $\gamma_{\tau B} = 0$ and $\gamma_B' \tilde{A}(B,-1) = 0$, it follows that $\mathbb{E}_{(\sqrt{nd} a, \sqrt{nd} \tau A, \Sigma^*)} \left[ \gamma' \tilde{Y}_n \right] = -\gamma_B' \tilde{A}(B,1) x$ regardless of $n$. Thus,

$$c \gamma' \tilde{Y}_n \mathcal{P}_\Omega \mathcal{N} \left( -c \gamma_B' \tilde{A}(B,1) x, c^2 \gamma' \Sigma \gamma \right).$$

(40)

Now, note that by construction, $v_n' \tilde{Y}_n - d_{n,1} = a_n + v_n' \beta - (\theta_{\tau n}^{ub} + x) - \chi_n$. Further, we have that $\beta_n \mathcal{P}_\Omega \mathcal{N} (\sqrt{n} \beta_A, \Sigma^*)$, where $\beta_A = \delta_A + \tau_A$. It follows that under $P_n$,

$$v_n' \tilde{Y}_n - d_{n,1} = a_n + v_n' \sqrt{n} \beta_A - (\theta_{\tau n}^{ub} + x) - \chi_n + v_n' \xi,$$

where $\xi \sim \mathcal{N} (0, \Sigma^*)$. Applying the equalities $v_n = \tilde{v}_n$, $a = \sqrt{n} \tilde{a}_n$, $\chi_n = \sqrt{n} \tilde{\chi}_n$ derived above, along with the fact that $\theta_{\tau n}^{ub} = \sqrt{n} \theta_1^{ub}$ by Lemma E.1, we see that under $P_n$,

$$v_n' \tilde{Y}_n - \tilde{d}_1 = \sqrt{n} \left( \tilde{a}_n + v_n' \beta_A - \theta_1^{ub} - \tilde{\chi}_n \right) - x + v_n' \xi,$$

(41)
Since \( v_n' \to \bar{v}, \bar{v}' \xi \to_d \bar{v}' \xi \) by Slutsky’s lemma. Additionally, the results above imply that

\[
\bar{a} + \bar{v}' \beta_A - \theta_1^{ub} - \bar{\chi}_n \to \bar{a} + \bar{v}' \beta_A - \theta_1^{ub} - \bar{b}_{min}(\Delta).
\]

Now, we claim that the limit in the previous display is strictly negative. To show this, note that since Assumption 3 holds, Lemma E.28 implies that \( LID(\Delta, \delta_A) > 0 \). Hence, for \( \epsilon > 0 \) sufficiently small, we have that \( \theta_1^{ub} - \epsilon \in S(\Delta, \beta_A) \). If the limit above were weakly positive, then we would have

\[
\bar{a} + \bar{v}' \beta_A - (\theta_1^{ub} - \epsilon) - \bar{b}_{min}(\Delta) > 0.
\]

However, this implies that \( \bar{b}(\bar{a}, \bar{v}) > \bar{b}_{min}(\Delta) \), which is a contradiction. The limit must thus be strictly negative, as desired. We then see from (41) that

\[
\bar{v}_n' \bar{Y}_n(\theta_1^{ub} + x) - d_{n,1} \frac{E_p}{p} - \infty. \tag{42}
\]

Displays (38), (39), (40), and (42) together give that \( v_{FLCI}^0 \frac{E_p}{p} - \infty \), as desired.

\[\square\]

**Proof of Proposition C.1**

**Proof.** First, by construction the least-favorable hybrid test rejects whenever the \( \bar{\alpha} \)-level conditional test rejects. That is, \( \psi_{C}(\tilde{Y}, \tilde{\Sigma}) \leq \psi_{C:LF}(\tilde{Y}, \tilde{\Sigma}) \) for any \( \tilde{Y} \) and \( \tilde{\Sigma} \). This is because the hybrid test is identical to the size \( \bar{\alpha} \) conditional test, except that it replaces \( v^0 \) with \( \min\{v^0, \psi_{C:LF}\} \), and \( \Pr(\xi < \xi \mid \xi \in [v^0, v^{up}], \xi \sim \mathcal{N}(0, \sigma_n^2)) \) is decreasing in \( v^0 \). The consistency of the hybrid thus follows from the consistency of the conditional test.

Second, from the definition of excess length, we have

\[
\mathbb{E}(\delta_A, \tau_A, \Sigma_n) \left[ EL(C_{\alpha, \delta_A}^0; \delta_A, \tau_A) \right] = \mathbb{E}(\delta_A, \tau_A, \Sigma_n) \left[ \int_{\mathbb{R} \setminus S(\Delta, \delta_A + \tau_A)} 1 - \psi_{C, \alpha}^{C:LF}(\beta_n; A, d, \theta, \Sigma_n) d\theta \right].
\]

From Lemma E.5 and \( \Sigma_n = \frac{1}{n} \Sigma^* \), note \( \psi_{C, \alpha}^{C:LF}(\beta_n; A, d, \theta, \Sigma_n) = \psi_{C, \alpha}^{C:LF}(\sqrt{n}\beta_n; A, \sqrt{n}d, \sqrt{n}\theta, \Sigma^*) \). Thus,

\[
\mathbb{E}(\delta_A, \tau_A, \Sigma_n) \left[ EL(C_{\alpha, \delta_A}^0; \delta_A, \tau_A) \right] = \mathbb{E}(\delta_A, \tau_A, \Sigma_n) \left[ \int_{\mathbb{R} \setminus S(\Delta, \delta_A + \tau_A)} 1 - \psi_{C, \alpha}^{C:LF}(\beta_n; A, \sqrt{n}d, \sqrt{n}\theta, \Sigma^*) d\theta \right] = \mathbb{E}(\sqrt{n}\delta_n, \sqrt{n}\tau_A, \Sigma^*) \left[ \int_{\mathbb{R} \setminus S(\Delta, \delta_A + \tau_A)} 1 - \psi_{C, \alpha}^{C:LF}(\beta_n; A, \sqrt{n}d, \sqrt{n}\theta, \Sigma^*) d\theta \right],
\]

where the second line uses the fact that if \( \beta_n \sim \mathcal{N}(\beta_A, \frac{1}{n} \Sigma^*) \), then \( \sqrt{n}\beta_n \sim \mathcal{N}(\sqrt{n}\beta_A, \Sigma^*) \). From Lemma E.1, \( \sqrt{n}S(\Delta, \delta_A + \tau_A) = S(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A) \), where \( \Delta_n = \{\delta : A\delta \leq \sqrt{n}d\} \). Applying
the change of variables $\bar{\theta}_{\text{new}} = \sqrt{n}\bar{\theta}_{\text{old}}$ to the integral inside the expectation in the previous display, we obtain

$$
E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{k, \alpha, n}; \delta_A, \tau_A) \right] = 
\frac{1}{\sqrt{n}} E(\sqrt{n}\delta_A, \sqrt{n}\tau_A, \Sigma^*) \left[ \int \mathbb{R}_+(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A) 1 - \psi_{k, \alpha}^{C-LF}(\hat{\theta}_n; A, \sqrt{n}d, \bar{\theta}, \Sigma^*) \, d\bar{\theta} \right].
$$

Recall from Section 6.2 that the identified set is an integral, $\mathcal{S}(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A) = [\theta_n^{lb}, \theta_n^{ub}]$, and so we can split the integral inside the expectation above into the parts above and below the identified set.

$$
E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{k, \alpha, n}; \delta_A, \tau_A) \right] = 
\frac{1}{\sqrt{n}} E(\sqrt{n}\delta_A, \sqrt{n}\tau_A, \Sigma^*) \left[ \int_0^\infty 1 - \psi_{k, \alpha}^{C-LF}(\hat{\theta}_n; A, \sqrt{n}d, \theta_n^{ub} + x, \Sigma^*) \, d\bar{\theta} \right] + 
\frac{1}{\sqrt{n}} E(\sqrt{n}\delta_A, \sqrt{n}\tau_A, \Sigma^*) \left[ \int_0^\infty 1 - \psi_{k, \alpha}^{C-LF}(\hat{\theta}_n; A, \sqrt{n}d, \theta_n^{lb} - x, \Sigma^*) \, d\bar{\theta} \right].
$$

For brevity here and in the associated lemmas, we will show that the first summand converges to 0; the convergence of the second summand can be obtained analogously. As discussed in Section C, the conditional least-favorable hybrid test rejects whenever the size-$\kappa$ least-favorable test rejects, so $0 \leq 1 - \psi_{k, \alpha}^{C-LF}(\hat{\theta}_n; A, \sqrt{n}d, \theta_n^{ub} + x, \Sigma^*) \leq 1 - \psi_{k}^{LF}(\hat{\theta}_n; A, \sqrt{n}d, \theta_n^{ub} + x, \Sigma^*)$. Lemma E.17 gives that there exists a function $\bar{\rho}_{LF}(x, A\Sigma^* A')$ such that for all $n$,

$$
0 \leq E(\sqrt{n}\delta_A, \sqrt{n}\tau_A, \Sigma^*) \left[ 1 - \psi_{k}^{LF}(\hat{\theta}_n; A, \sqrt{n}d, \theta_n^{ub} + x, \Sigma^*) \right] \leq 1 - \bar{\rho}_{LF}(x, A\Sigma^* A'),
$$

and $\int_0^\infty (1 - \rho_{LF}(x, A\Sigma^* A')) \, dx < \infty$. Thus, we have that

$$
0 \leq \limsup_{n \to \infty} E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{k, \alpha, n}; \delta_A, \tau_A) \right] \leq \lim_{n \to \infty} \frac{1}{\sqrt{n}} \int_0^\infty 1 - \rho_{LF}(x) \, dx = 0,
$$

completing the proof.

Proof of Proposition C.2

Proof. Note that the LF hybrid rejects whenever the size-$\bar{\alpha}$ conditional test rejects. This is because the “second-stage” of the hybrid is equivalent to the size-$\bar{\alpha}$ conditional test, except it uses a weakly lower value of $v^{up}$, and the critical value for the test is decreasing in $v^{up}$. The first result then follows immediately from Proposition 5.2.
Following Pratt (1961) and Armstrong and Kolesar (2018), observe that for any confidence set \( C \) with finite expected excess length,

\[
\mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C; \delta_A, \tau_A) \right] = \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ \int_{\mathbb{R}} \left| S(\Delta, \delta_A, \tau_A) \right| \mathbb{1} \{ \hat{\theta} \in C \} \right] = \int_{\mathbb{R}} \mathbb{E} \left( \delta_A, \tau_A, \Sigma_n \right) \mathbb{1} \{ \hat{\theta} \in C \} \right] d\hat{\theta},
\]

by Fubini’s theorem. It follows that the confidence set that minimizes expected excess length inverts most powerful tests of \( H_0 : \delta \in \Delta, \theta = \hat{\theta} \) against \( H_1 : (\delta, \tau) = (\delta_A, \tau_A) \) for each point \( \hat{\theta} \).

Note that this testing problem is equivalent to testing \( H_0 : \beta \in B_0(\hat{\theta}, d) = \{ \beta : \exists \tau_{post} \text{ s.t. } \tau_{post} = \hat{\theta}, A\beta - d - AM_{post} \tau_{post} \leq 0 \} \), the set of values for \( \beta \) consistent with the null that \( \theta = \hat{\theta} \), against the alternative \( H_1 : \beta = \beta_A := \delta_A + \tau_A \). We thus wish to show that

\[
\limsup_{n \to \infty} \frac{\sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C^{C-LF}; \delta_A, \tau_A) \right]}{\mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C^{MP}; \delta_A, \tau_A) \right]} \leq 1,
\]

for \( C^{MP}_{\hat{\alpha}, n} \) the confidence interval that inverts the most powerful size-\( \hat{\alpha} \) tests when \( \hat{\beta}_n \sim \mathcal{N} (\beta_A, \Sigma_n) \).

Note that we can multiply both the numerator and denominator by \( \sqrt{n} \), so it suffices to show that

\[
\limsup_{n \to \infty} \frac{\sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C^{C-LF}; \delta_A, \tau_A) \right]}{\sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C^{MP}; \delta_A, \tau_A) \right]} \leq 1.
\]

To do this, we will show that the limit superior of the numerator is bounded above by the limit of the denominator, which is strictly positive.

First, we turn our attention to the numerator. By definition, we have

\[
\sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C^{C-LF}; \delta_A, \tau_A) \right] = \sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ \int_{\mathbb{R}} \left| S(\Delta, \delta_A + \tau_A) \right| 1 - \psi^{C-LF}_{\kappa, \alpha} (\hat{\beta}_n; A, \bar{d}, \hat{\theta}, \frac{1}{n} \Sigma^*) \right] d\hat{\theta},
\]

Lemma E.5 along with the fact that \( \Sigma_n = \frac{1}{n} \Sigma^* \) give that \( \psi^{C-LF}_{\kappa, \alpha} (\hat{\beta}_n; A, d, \hat{\theta}, \Sigma) = \psi^{C-LF}_{\kappa, \alpha} (\sqrt{n} \hat{\beta}_n; A, \sqrt{n} d, \sqrt{n} \hat{\theta}, n \Sigma) \), which implies

\[
\sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ EL(C^{C-LF}; \delta_A, \tau_A) \right] = \sqrt{n} \mathbb{E}_{\delta_A, \tau_A, \Sigma_n} \left[ \int_{\mathbb{R}} \left| S(\Delta, \delta_A + \tau_A) \right| 1 - \psi^{C-LF}_{\kappa, \alpha} \left( \sqrt{n} \hat{\beta}_n; A, \sqrt{n} d, \sqrt{n} \hat{\theta}, \Sigma^* \right) \right] d\hat{\theta}
\]

\[
= \sqrt{n} \mathbb{E}_{\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*} \left[ \int_{\mathbb{R}} \left| S(\Delta, \delta_A + \tau_A) \right| 1 - \psi^{C-LF}_{\kappa, \alpha} (\hat{\beta}_n; A, \sqrt{n} d, \sqrt{n} \hat{\theta}, \Sigma^*) \right] d\hat{\theta},
\]

where the second line uses the fact that if \( \hat{\beta}_n \sim \mathcal{N} (\beta_A, \frac{1}{n} \Sigma^*) \), then \( \sqrt{n} \hat{\beta}_n \sim \mathcal{N} (\sqrt{n} \beta_A, \Sigma^*) \). From Lemma E.1, \( \sqrt{n} S(\Delta, \delta_A + \tau_A) = S(\Delta_n, \sqrt{n} \delta_A + \sqrt{n} \tau_A) \), where \( \Delta_n = \{ \delta : A\delta \leq \sqrt{n} d \} \). Applying the change of variables \( \hat{\theta}_{\text{new}} = \sqrt{n} \hat{\theta}_{\text{old}} \) to the integral inside the expectation in the previous display,
we obtain

\[ \sqrt{n}\mathbb{E}(\delta_{A, \tau_A}, \Sigma) \left[ EL(C_{C, LF}; \delta_A, \tau_A) \right] = \\
\mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta, \Sigma^*) d\theta \right]. \]

Since the integrand is weakly positive, by Fubini’s theorem we can reverse the order of integration provided that the resulting integral is finite (as we will show below). We thus have

\[ \sqrt{n}\mathbb{E}(\delta_{A, \tau_A}, \Sigma) \left[ EL(C_{C, LF}; \delta_A, \tau_A) \right] = \\
\int_{\mathbb{R}\setminus S(\Delta_n, \sqrt{n}\delta_A, \sqrt{n}\tau_A)} \mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta, \Sigma^*) \right] d\theta = \\
\int_0^\infty \mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \right] dx + \\
\int_0^\infty \mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{lb} - x, \Sigma^*) \right] dx, \]

where the second equality splits the integral into the parts above and below the identified set (which recall we have shown in Section 6.2 to be an interval), and we define \( \theta_{ub} \) and \( \theta_{lb} \) to be the upper and lower bounds of \( S(\Delta_n, \sqrt{n}\delta_A, \sqrt{n}\tau_A) \). To see why the application of Fubini’s theorem is valid, note that \( 0 \leq 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \leq 1 - \psi_{LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \), since by construction the hybrid test rejects whenever the \( \kappa \)-level least favorable test rejects. However, Lemma E.17 gives that there exists a function \( \rho_{LF}(x, A\Sigma^* A') \) such that for all \( n \), \( 0 \leq \mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \right] \leq 1 - \rho_{LF}(x, A\Sigma^* A') \), and \( \int_0^\infty \rho_{LF}(x, A\Sigma^* A') dx < \infty \). An analogous argument for the part below the identified set gives that the integral is finite.

For ease of exposition in this proof and the associated lemmas, we show that the limsup of the integral for the part above the identified set is bounded above by the limit of the analogous integral for the most powerful test as \( n \to \infty \); the proof for the part below the identified set is analogous.

We showed above that there exists a positive integrable function \( 1 - \rho_{LF}(x, A\Sigma^* A') \) that is weakly greater than \( 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \) for all \( x \) and all \( n \). We can thus apply the Reverse Fatou’s Lemma to obtain

\[ \limsup_{n \to \infty} \int_0^\infty \mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \right] dx \leq \\
\int_0^\infty \limsup_{n \to \infty} \mathbb{E}(\sqrt{n}\delta_{A, \tau_A}, \Sigma^*) \left[ 1 - \psi_{C, LF}^*(\hat{\beta}_n; A, \sqrt{n}d, \theta_{ub} + x, \Sigma^*) \right] dx. \]

Now, we argued in the proof to Proposition C.1 that the hybrid test rejects whenever the size-\( \alpha \)
conditional test rejects. Thus,

\[
\limsup_{n \to \infty} E_{\sqrt{n} \delta_n, \sqrt{n} \tau_n, \Sigma^*} \left[ 1 - \psi_{\delta_n, \alpha}^{C-LF} (\hat{\beta}; A, \sqrt{n}d, \theta_{\alpha}^{ub} + x, \Sigma^*) \right] \leq \limsup_{n \to \infty} E_{\sqrt{n} \delta_n, \sqrt{n} \tau_n, \Sigma^*} \left[ 1 - \psi_{\delta_n}^{C} (\hat{\beta}; A, \sqrt{n}d, \theta_{\alpha}^{ub} + x, \Sigma^*) \right].
\]

However, by Lemma E.9, as \( n \to \infty \), the upper bound in the previous display is equal to \( \Phi(z_{1-\bar{\alpha}} - c^* x) \) for a positive constant \( c^* \). We have thus shown that

\[
\limsup_{n \to \infty} \sqrt{n} E_{\delta_n, \tau_n, \Sigma_n} \left[ EL(C_{mp}^{C-LF}; A_n, \tau_n) \right] \leq \int_{0}^{\infty} \Phi(z_{1-\bar{\alpha}} - c^* x) dx.
\]

Applying the change of variables \( s = z_{1-\bar{\alpha}} - c^* x \), we see that

\[
\int_{0}^{\infty} \Phi(z_{1-\bar{\alpha}} - c^* x) dx = \frac{1}{c^*} \int_{-\infty}^{z_{1-\bar{\alpha}}} \Phi(s) ds = \frac{1}{c^*} [z_{1-\bar{\alpha}} (1 - \bar{\alpha}) + \phi(z_{1-\bar{\alpha})}],
\]

where the second equality uses Lemma E.16.

Now, we turn our attention to the denominator of the limit of interest, and follow similar steps as we did for the numerator. We have

\[
\sqrt{n} E_{\delta_n, \tau_n, \Sigma_n} \left[ EL(C_{mp}^{MP}; \delta_n, \tau_n) \right] = \sqrt{n} E_{\delta_n, \tau_n, \Sigma_n} \left[ \int_{\mathbb{R} \times S(\Delta, \delta_n, \tau_n)} 1 - \psi_{\delta_n}^{MP} (\hat{\beta}_n; A, d, \bar{\theta}, \frac{1}{n} \Sigma^*, \delta_n, \tau_n) d\bar{\theta} \right],
\]

where, as in the proof to Proposition 5.2, \( \psi_{\delta_n}^{MP} (\hat{\beta}_n; A, d, \bar{\theta}, \Sigma, \delta_n, \tau_n) \) is an indicator for whether the most powerful size-\( \bar{\alpha} \) test between the null hypothesis \( H_0 : \delta \in \Delta, \theta = \bar{\theta} \) and the alternative \( H_A : (\delta, \tau) = (\delta_A, \tau_A) \) rejects the null.

Using Lemma E.7, we obtain

\[
\sqrt{n} E_{\delta_n, \tau_n, \Sigma_n} \left[ EL(C_{mp}^{MP}; \delta_n, \tau_n) \right] = \sqrt{n} E_{\delta_n, \tau_n, \Sigma_n} \left[ \int_{\mathbb{R} \times S(\Delta, \delta_n, \tau_n)} 1 - \psi_{\delta_n}^{MP} (\hat{\beta}_n; A, \sqrt{n}d, \sqrt{n}\bar{\theta}, \Sigma^*, \sqrt{n}\delta_n, \sqrt{n}\tau_n) d\bar{\theta} \right] = \sqrt{n} E_{\delta_n, \tau_n, \Sigma_n} \left[ \int_{\mathbb{R} \times S(\Delta, \delta_n, \tau_n)} 1 - \psi_{\delta_n}^{MP} (\hat{\beta}_n; A, \sqrt{n}d, \sqrt{n}\bar{\theta}, \Sigma^*, \sqrt{n}\delta_n, \sqrt{n}\tau_n) d\bar{\theta} \right].
\]
Proof of Corollary C.1

By Lemma E.13, analogously. Let \( \rho \) for Fubini’s theorem are satisfied. An analogous argument can be applied to the integral below the identified set, and so the conditions Lemma E.17 implies, however, that the integral of the upper bound in the previous display is finite.

Since the integrand is weakly positive, by Fubini’s theorem we can reverse the order of integration provided that the resulting integral is finite (as we will show below). We thus have

\[
\sqrt{n}E_{(\delta, \tau; \Sigma_n)}[E|\mathcal{C}_p; \delta, \tau_A] = \\
\int_{\mathbb{R}, \mathcal{S}(\Delta_n; \sqrt{n}d + \sqrt{n}\tau_A)} E_{(\sqrt{n}\delta, \sqrt{n}\tau_A; \Sigma^*)} \left[ 1 - \psi_G^{MP}(\sqrt{n}d; A, \sqrt{n}\delta, \sqrt{n}\tau_A) \right] d\theta \\
\int_0^\infty E_{(\sqrt{n}\delta, \sqrt{n}\tau_A; \Sigma^*)} \left[ 1 - \psi_G^{MP}(\sqrt{n}d; A, \sqrt{n}\delta, \sqrt{n}\tau_A) \right] dx \\
\int_0^\infty E_{(\sqrt{n}\delta, \sqrt{n}\tau_A; \Sigma^*)} \left[ 1 - \psi_G^{MP}(\sqrt{n}d; A, \sqrt{n}\delta, \sqrt{n}\tau_A) \right] dx,
\]

where as before we split the integral into the parts above and below the identified set. To see why the application of Fubini’s theorem is valid, note that by construction, the power of the most powerful \( \alpha \)-level test must be weakly greater than that of the \( \alpha \)-level least favorable test, and so

\[
0 \leq E_{(\sqrt{n}\delta, \sqrt{n}\tau_A; \Sigma^*)} \left[ 1 - \psi_G^{MP}(\sqrt{n}d; A, \sqrt{n}\delta, \sqrt{n}\tau_A) \right] \\
\leq E_{(\sqrt{n}\delta, \sqrt{n}\tau_A; \Sigma^*)} \left[ 1 - \psi_G^{LF}(\sqrt{n}d; A, \sqrt{n}\delta, \sqrt{n}\tau_A) \right].
\]

Lemma E.17 implies, however, that the integral of the upper bound in the previous display is finite. An analogous argument can be applied to the integral below the identified set, and so the conditions for Fubini’s theorem are satisfied.

We again focus on the part above the identified set; the proof for the part below proceeds analogously. Let \( \rho^{MP}(\theta_n + x) = E_{(\sqrt{n}\delta, \sqrt{n}\tau_A; \Sigma^*)} \left[ \psi_G^{MP}(\sqrt{n}d; A, \sqrt{n}\delta, \sqrt{n}\tau_A) \right] \). By Lemma E.13, \( 1 - \rho^{MP}(\theta_n + x) \rightarrow \Phi(z_{1-\alpha} - c^* x) \) as \( n \rightarrow \infty \). Additionally, Lemma E.17 implies that there exists a function \( \rho^{LF}(x, A\Sigma^* A') \) such that \( 0 \leq 1 - \rho^{MP}(\theta_n + x) \rightarrow \Phi(z_{1-\alpha} - c^* x) \), and so by the dominated convergence theorem \( \int_0^\infty 1 - \rho^{MP}(\theta_n + x) dx \leq \Phi(z_{1-\alpha} - c^* x) \). The limit of the denominator is thus equal to the upper bound that we obtained for the numerator, which completes the proof.

Proof of Corollary C.1
Proof. Observe that

$$\limsup_{n \to \infty} \frac{E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{C, LF}; \delta_A, \tau_A) \right]}{\inf_{c_{\alpha, n} \in I_\alpha(\Delta, \Sigma_n)} E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{C, \alpha, n}; \delta_A, \tau_A) \right]} =$$

$$\limsup_{n \to \infty} \frac{\sqrt{n} E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{C, LF}; \delta_A, \tau_A) \right]}{\inf_{c_{\alpha, n} \in I_\alpha(\Delta, \Sigma_n)} \left[ EL(C_{C, \alpha, n}; \delta_A, \tau_A) \right]} \leq$$

$$\limsup_{n \to \infty} \sqrt{n} \inf_{c_{\alpha, n} \in I_\alpha(\Delta, \Sigma_n)} E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{C, \alpha, n}; \delta_A, \tau_A) \right] \leq \limsup_{n \to \infty} \sqrt{n} \inf_{c_{\alpha, n} \in I_\alpha(\Delta, \Sigma_n)} E_{(\delta_A, \tau_A, \Sigma_n)} \left[ EL(C_{C, \alpha, n}; \delta_A, \tau_A) \right].$$

Proposition C.2 gives that the first limsup in the product in the previous display is bounded above by 1. Additionally, the argument in the proof to Proposition C.2 implies that the numerator in the second limsup converges to

$$\left( \frac{1}{C^*} + \frac{1}{c^{**}} \right) (z_{1-\alpha} (1 - \bar{\alpha}) + \phi(z_{1-\bar{\alpha}}))$$

for positive constants $c^*$ and $c^{**}$.\footnote{In particular, we show that the term in the numerator can be decomposed into two integrals, corresponding with the parts above and below the identified set. We show the former converges to $\frac{1}{C^*} (z_{1-\alpha} (1 - \bar{\alpha}) + \phi(z_{1-\bar{\alpha}}))$. An analogous argument yields that the second integral converges to $\frac{1}{c^{**}} (z_{1-\alpha} (1 - \bar{\alpha}) + \phi(z_{1-\bar{\alpha}}))$, for a (possibly different) positive constant $c^{**}$.}

Analogously, the denominator of the second limsup converges to

$$\left( \frac{1}{C^*} + \frac{1}{c^{**}} \right) (z_{1-\alpha} (1 - \alpha) + \phi(z_{1-\alpha})).$$

The first result is then immediate from dividing the expression in the previous two displays. Since the quantile function and density of the normal distribution are continuous in $\alpha$ and $\bar{\alpha} \to \alpha$ as $\kappa \to 0$, the second result follows from the first.

\[\square\]

E.2 Auxiliary Lemmas for Finite Sample Normal Results

Lemma E.1. For any $n > 0$, let $\Delta_n = \{ \delta : A\delta \leq \sqrt{nd} \}$. Fix $\delta_A \in \Delta_1$ and $\tau_A$. Then $S(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A) = \sqrt{n}S(\Delta_1, \delta_A + \tau_A)$. This implies that $\theta_{n}^b = \sqrt{n}\theta_{1}^b$, where $\theta_{n}^b := \sup_{\delta} S(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A)$, and likewise $\theta_{n}^b = \sqrt{n}\theta_{1}^b$, for $\theta_{n}^b := \inf_{\delta} S(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A)$.

Proof. For ease of notation, let $S_n = S(\Delta_n, \sqrt{n}\delta_A + \sqrt{n}\tau_A)$ and $B_n = \delta_A + \tau_A$. By definition, $\bar{\theta}_n \in S_n$ iff there exists a vector $\tau_{post} \in \mathbb{R}^T$ such that $l^T \tau_{post} = \bar{\theta}_n$ and $A(\sqrt{n}\beta_A - M_{post} \tau_{post} - \sqrt{nd}) \leq 0$. Using the change of basis described in Section 4.2, it follows that $\bar{\theta}_n \in S_n$ iff there exists $\bar{\tau}_n \in \mathbb{R}^{T-1}$ such that

S-22
\[ A \sqrt{n} \beta_A - \sqrt{n} d - \tilde{A}_{(\cdot,1)} \tilde{\theta}_n - \tilde{A}_{(\cdot,-1)} \tilde{\tau}_n \leq 0. \]  

(43)

Thus, \( \tilde{\theta}_1 \in \mathcal{S}_1 \) iff there exists \( \tilde{\tau}_1 \) such that

\[ A \beta_A - d - \tilde{A}_{(\cdot,1)} \tilde{\theta}_1 - \tilde{A}_{(\cdot,-1)} \tilde{\tau}_1 \leq 0. \]  

(44)

If there exists a \( \tilde{\tau}_1 \) such that (44) holds for \( \tilde{\theta}_1 \), then multiplying both sides of (44) by \( \sqrt{n} \) implies that (43) holds with \( \tilde{\theta}_n = \sqrt{n} \tilde{\theta}_1 \) and \( \tilde{\tau}_n = \sqrt{n} \tilde{\tau}_1 \). Likewise, if there exists a \( \tilde{\tau}_n \) such that (43) holds for \( \tilde{\theta}_n \), then multiplying both sides of (43) by \( \frac{1}{\sqrt{n}} \) implies that (44) holds with \( \tilde{\theta}_1 = \frac{1}{\sqrt{n}} \tilde{\theta}_n \) and \( \tilde{\tau}_1 = \frac{1}{\sqrt{n}} \tilde{\tau}_n \). The desired result follows immediately.

\[ \Box \]

**Lemma E.2.** For any \( n > 0 \) and \((\beta; A, d, \tilde{\theta}, \Sigma)\), \( \psi_C^C(\beta; A, d, \tilde{\theta}, \Sigma) = \psi_C^C(\sqrt{n} \tilde{\beta}; A, \sqrt{n} d, \sqrt{n} \tilde{\theta}_n, n \Sigma) \).

**Proof.** Using the change of basis described in Section 4.2, the test statistic used to calculate \( \psi_C^C(\beta; A, d, \tilde{\theta}, \Sigma) \) is

\[
\min_{\eta, \hat{\tau}} \eta \quad \text{s.t.} \quad A \hat{\beta} - d - \tilde{A}_{(\cdot,1)} \hat{\theta} - \tilde{A}_{(\cdot,-1)} \hat{\tau} \leq \eta \hat{\sigma},
\]

where \( \hat{\sigma} \) is the vector containing the square roots of the diagonal elements of \( \hat{\Sigma} = A \Sigma A' \). Since multiplying the constraints by \( \sqrt{n} \) does not change the feasible set, this optimization is equivalent to

\[
\min_{\eta, \hat{\tau}} \eta \quad \text{s.t.} \quad A \sqrt{n} \hat{\beta} - \sqrt{n} d - \tilde{A}_{(\cdot,1)} \sqrt{n} \hat{\theta} - \tilde{A}_{(\cdot,-1)} \sqrt{n} \hat{\tau} \leq \eta \sqrt{n} \hat{\sigma}.
\]

However, since \( \hat{\tau} \) enters only in the constraint, and \( \{ \sqrt{n} \tilde{\tau} : \tilde{\tau} \in \mathbb{R}^{T-1} \} = \{ \tilde{\tau} \in \mathbb{R}^{T-1} \} \), this linear program is equivalent to

\[
\min_{\eta, \tilde{\tau}} \eta \quad \text{s.t.} \quad A \sqrt{n} \tilde{\beta} - \sqrt{n} d - \tilde{A}_{(\cdot,1)} \sqrt{n} \tilde{\theta} - \tilde{A}_{(\cdot,-1)} \tilde{\tau} \leq \eta \sqrt{n} \tilde{\sigma},
\]

which is the test statistic used to calculate \( \psi_C^C(\sqrt{n} \tilde{\beta}; A, \sqrt{n} d, \sqrt{n} \tilde{\theta}_n, n \Sigma) \). Thus, the test statistics used for the two problems are the same. Additionally, the feasible set for the dual for the unscaled problem is \( F_1 = \{ \gamma : \gamma' \tilde{A}_{(\cdot,1)} = 0, \gamma' \tilde{\sigma} = 1, \gamma \geq 0 \} \), whereas for the problem scaled by \( \sqrt{n} \) it is \( F_n = \{ \gamma : \gamma' \tilde{A}_{(\cdot,1)} = 0, \gamma' \sqrt{n} \tilde{\sigma} = 1, \gamma \geq 0 \} = \frac{1}{\sqrt{n}} F_1 \). It follows that \( V_n = \frac{1}{\sqrt{n}} V_1 \), for \( V_1 \) and \( V_n \), respectively the vertices of \( F_1 \) and \( F_n \). Moreover, it is immediate that if \( \gamma_1 \) is an optimal vertex of the unscaled problem, then \( \gamma_n = \frac{1}{\sqrt{n}} \gamma_1 \) will be an optimal vertex of the problem scaled by \( \sqrt{n} \).

Recall that the critical value for the conditional test depends on \( \gamma_1' \tilde{\Sigma} \gamma_*, \) where \( \gamma_* \) is an optimal vertex, and the values \( v^{lo} \) and \( v^{up} \) which are functions of \( \gamma_*, \tilde{\Sigma} \), and a sufficient statistic \( S \). However, for \( \gamma_n = \frac{1}{\sqrt{n}} \gamma_1 \), we have that \( \gamma_n'(n \tilde{\Sigma}) \gamma_n = \frac{1}{\sqrt{n}} \gamma_1'(n \tilde{\Sigma}) \frac{1}{\sqrt{n}} \gamma_1 = \gamma_1' \tilde{\Sigma} \gamma_1 \), and so the variances are the
same. Let $\tilde{Y}_1 = A\hat{\beta} - d - \tilde{A}_{(\cdot,1)}\tilde{\theta}$ and $\tilde{Y}_n = A\sqrt{n}\tilde{\beta} - \sqrt{n}d - \tilde{A}_{(\cdot,1)}\sqrt{n}\tilde{\theta} = \sqrt{n}\tilde{Y}_1$. The sufficient statistic used to construct $v^{lo}$ and $v^{up}$ in the first problem is $S_1 = (I - \frac{\Sigma^\gamma_n\gamma_n'}{\gamma_n^2n})\tilde{Y}_1$, whereas for the second problem it is $S_n = (I - \frac{n\tilde{\Sigma}^\gamma_n\gamma_n'}{\gamma_n^2n})\tilde{Y}_n$. The identities $\tilde{Y}_n = \sqrt{n}\tilde{Y}_1$ and $\gamma_1 = \sqrt{n}\gamma_n$ immediately imply that $S_n = \sqrt{n}S_1$. The values $v^{lo}$ and $v^{up}$ for the first problem are then the minimum and maximum of

$$C_1 = \{c : c = \max_{\gamma_1 \in V_1} \bar{z}'_1(S_1 + \frac{\tilde{\Sigma}\gamma_1}{\gamma_1'\Sigma}\gamma_1 c)\}.$$ 

Likewise, the values $v^{lo}$ and $v^{up}$ for the second problem are the the minimum and maximum of

$$C_n = \{c : c = \max_{\tilde{\gamma}_n \in V_n} \bar{z}'_n(S_n + \frac{n\tilde{\Sigma}\gamma_n}{\gamma_n'\Sigma}\gamma_n c)\},$$

However, since $V_n = \sqrt{n}V_1$, $S_n = \sqrt{n}S_1$, and $\gamma_n = \frac{1}{\sqrt{n}}\gamma_1$, we have that for any $c$,

$$\max_{\tilde{\gamma}_n \in V_n} \bar{z}'_n(S_n + \frac{n\tilde{\Sigma}\gamma_n}{\gamma_n'\Sigma}\gamma_n c) = \max_{\gamma_1 \in V_1} \bar{z}'_1(S_1 + \frac{\tilde{\Sigma}\gamma_1}{\gamma_1'\Sigma}\gamma_1 c) = \max_{\gamma_1 \in V_1} \bar{z}'_1(S_1 + \frac{\tilde{\Sigma}\gamma_1}{\gamma_1'\Sigma}\gamma_1 c),$$

from which it is immediate that $C_1 = C_n$, and hence the values of $v^{lo}$ and $v^{up}$ are the same across the two problems as well. Since the test statistics and critical values of the two problems are the same, they are equivalent. 

**Lemma E.3.** Let $v_n(\Delta, \Sigma_n)$ be the vector $v_n$ used in the optimal FLCI as defined in Section 6, making the dependence on $(\Delta, \Sigma_n)$ explicit. Define, $\chi_n(\Delta, \Sigma_n)$, $a_n(\Delta, \Sigma_n)$, and $b_n(\Delta, \Sigma_n)$ analogously. Let $\Sigma_n = \frac{1}{n}\Sigma^*$ and $\Delta_n = \sqrt{n}\Delta$. Then

1. $v_n(\Delta_n, \Sigma^*) = v_n(\Delta, \frac{1}{n}\Sigma_n)$
2. $a_n(\Delta_n, \Sigma^*) = \sqrt{n}a_n(\Delta, \frac{1}{n}\Sigma_n)$
3. $\chi_n(\Delta_n, \Sigma^*) = \sqrt{n}\chi_n(\Delta, \frac{1}{n}\Sigma_n)$
4. $\tilde{b}(a_n(\Delta_n, \Sigma^*), v_n(\Delta_n, \Sigma^*); \Delta_n) = \sqrt{n}\tilde{b}(a_n(\Delta, \Sigma_n), v_n(\Delta, \Sigma_n); \Delta)$.

**Proof.** Recall the proof to Lemma E.22 that $\tilde{b}(a, v; \Delta_n)$ is finite only if $v_{post} = l$, in which case

$$\tilde{b}(a, v; \Delta) = \max_{\delta \in \Delta} |a + v'\delta|.$$ 

Likewise, $\tilde{b}(|\sqrt{n}a, v; \Delta_n)$ is finite only if $v_{post} = l$, in which case

$$\tilde{b}(|\sqrt{n}a, v; \Delta) = \max_{\delta_n \in \Delta_n} |\sqrt{n}a + v'\delta_n| = \max_{\delta \in \Delta} |\sqrt{n}a + v'\sqrt{n}\delta| = \sqrt{n}\tilde{b}(a, v, \Delta).$$
Next, observe that using the invariance above and the fact that \( \Sigma_n = \frac{1}{n} \Sigma^* \),

\[
\chi(\sqrt{n}a, v; \Sigma^*, \Delta_n) = \sqrt{n}^{\Sigma^*}v \cdot cv_{\alpha} \left( \frac{\tilde{b}(\sqrt{n}a, v; \Delta_n)}{\sqrt{n}^{\Sigma^*}v} \right) = \sqrt{n} \cdot \sqrt{n^{\Sigma^*}v} \cdot cv_{\alpha} \left( \frac{\tilde{b}(a, v; \Delta)}{\sqrt{n}^{\Sigma^*}v} \right) = \sqrt{n} \chi(a, v; \Sigma^*, \Delta). \]

It is then immediate that if \( (a^*, v^*) = \arg\min_{(a,v)} \chi(a, v; \Delta, \Sigma_n) \), then \( (\sqrt{n}a^*, v^*) = \arg\min_{(a,v)} \chi(a, v; \Delta_n, \Sigma^*) \), from which the first two results follow. The second two results then follow from the two invariances derived above.

**Lemma E.4.** For any \( n > 0 \) and \((\hat{\beta}; A, d, \hat{\theta}, \Sigma)\),

\[
\psi_{\kappa, \alpha}^{C-FLCI}(\hat{\beta}; A, d, \hat{\theta}, \Sigma) = \psi_{\kappa, \alpha}^{C-FLCI}(\sqrt{n}\hat{\beta}; A, \sqrt{n}d, \sqrt{n}\hat{\theta}, n\Sigma).
\]

**Proof.** Lemma E.3 implies that if

\[
C^F_{\kappa}(\hat{\beta}; A, d, \hat{\theta}, \Sigma) = a_1 + v'_1\hat{\beta} \pm \chi_1,
\]

then

\[
C^F_{\kappa}(\sqrt{n}\hat{\beta}; A, \sqrt{n}d, \sqrt{n}\hat{\theta}, \sqrt{n}\Sigma) = \sqrt{n} \left( a_1 + v'_1\hat{\beta} \pm \chi_1 \right),
\]

and thus \( \hat{\theta} \in C^F_{\kappa}(\hat{\beta}; A, d, \hat{\theta}, \Sigma) \) if and only if \( \sqrt{n}\hat{\theta} \in C^F_{\kappa}(\sqrt{n}\hat{\beta}; A, \sqrt{n}d, \sqrt{n}\hat{\theta}, \sqrt{n}\Sigma) \), so the first stage tests are equivalent. Note that the second stage test is almost identical to \( \psi_{\kappa, \alpha}^C \), which we showed in Lemma E.2 to be invariant to scale, except it replaces \( v^{lo} \) with \( \max\{v^{lo}, v^{lo}_{FLCI}\} \) and \( v^{up} \) with \( \min\{v^{up}, v^{up}_{FLCI}\} \). It thus suffices to show that \( v^{lo}_{FLCI} \) and \( v^{up}_{FLCI} \) are invariant to scale. We show in the proof to Lemma E.2 that \( S_n = \sqrt{n}S_1 \), and if \( \gamma_n \in \hat{V}_n \) then \( \gamma = \sqrt{n}\gamma_n \in \hat{V}_n \); where objects subscripted by 1 indicate values based on \((\hat{\beta}; A, d, \hat{\theta}, \Sigma)\) and values subscripted by \( n \) indicate those based on \((\sqrt{n}\hat{\beta}; A, \sqrt{n}d, \sqrt{n}\hat{\theta}, \sqrt{n}\Sigma)\). Additionally, Lemma E.3 implies that \( \hat{V}_1 = \hat{V}_n \), and \( \hat{\alpha}_n = \sqrt{n}\hat{\alpha}_d \). The desired invariance is then immediate from the formulas in Lemma B.2.

**Lemma E.5.** For any \( n > 0 \) and \((\hat{\beta}; A, d, \hat{\theta}, \Sigma)\), \( \psi_{\kappa, \alpha}^{C-FL} (\hat{\beta}; A, d, \hat{\theta}, \Sigma) = \psi_{\kappa, \alpha}^{C-FL} (\sqrt{n}\hat{\beta}; A, \sqrt{n}d, \sqrt{n}\hat{\theta}, n\Sigma) \).

**Proof.** Note that the hybrid test is a function of \( \hat{\eta} \), the least favorable critical value, and the values \( v^{lo} \) and \( v^{up} \) and \( \gamma'\Sigma\gamma \) from the size \( \frac{\alpha}{1-\alpha} \) conditional test. We showed in the proof to Lemma E.2 that the values \( \hat{\eta}, v^{lo}, v^{up} \) and \( \gamma'\Sigma\gamma \) are equivalent across the two problems. It thus suffices to show that the least favorable critical values are also the same.

Recall that the critical value for the size \( \kappa \) least favorable test used in the first-stage of \( \psi_{\kappa, \alpha}^{C-FL} (\hat{\beta}; A, d, \hat{\theta}, \Sigma) \) is the \( 1 - \kappa \) quantile of

\[
\min_{\eta, \bar{\eta}} \eta \quad \text{s.t.} \quad \xi - \bar{A}(-1)\bar{\bar{\tau}} \leq \eta \bar{\sigma},
\]

for \( \xi_1 \sim \mathcal{N}(0, A\Sigma A') \). Rewriting in terms of the dual formulation, the critical value is the \( 1 - \kappa \) quantile of \( \max_{\gamma_1 \in V_1} \gamma' \xi_1 \), where \( V_1 \) is the set of vertices of the feasible region for the dual, \( F_1 := \)
\{γ : γ' A_{\{ -1 } = 0, γ' σ = 1, γ ≥ 0\}.

Analogously, the critical value for the least favorable test used in \(ψ_{\alpha}^{C-LF}(\sqrt{n} \hat{δ}; A, \sqrt{n} d, \sqrt{n} \hat{θ}, n Σ)\) can be written as the \(1 - κ\) quantile of \(max_{γ_n \in V_n} γ_n' ξ_n\), where \(ξ_n \sim N(0, A(n Σ)A')\) and \(V_n\) is the set of vertices of \(F_n\) defined in Lemma E.2, however, that \(V_n = \frac{1}{\sqrt{n}} V_1\). It follows that

\[
q_{1-α}\left(\max_{γ_n \in V_n} γ_n' ξ_n, ξ_n \sim N(0, A(n Σ)A')\right) = q_{1-α}\left(\max_{γ_n \in V_1} \left(\frac{1}{\sqrt{n}} γ_n\right)' ξ_n, ξ_n \sim N(0, A(n Σ)A')\right)
\]

\[
= q_{1-α}\left(\max_{γ_n \in V_1} γ_n' ξ_n, ξ_n \sim N(0, A(n Σ)A')\right)
\]

and thus the critical values for the two least-favorable tests coincide.

\[\square\]

**Lemma E.6.** Suppose \(\hat{δ} \sim N(δ, Σ)\) for \(Σ\) known. Let \(B_0\) be a closed, convex set. Then the most-powerful size \(α\) test of \(H_0 : β ∈ B_0\) against the point alternative \(H_A : β = \beta_A\) is equivalent to the most powerful test of \(H_0 : β = \hat{δ}\) against \(H_A : β = \beta_A\), where \(\hat{δ} = arg\ min_{β} ||β - β_A||_Σ\) and \(||·||_Σ\) is the Mahalanobis norm in \(Σ\). \(||x||_Σ = \sqrt{x' Σ^{-1} x}\). The most powerful test rejects for values of \((β_A - \hat{δ})' Σ^{-1} \hat{δ}\) greater than \((β_A - \hat{δ})' Σ^{-1} \hat{δ} + z_{1-α} ||β_A - \hat{δ}||_Σ\), and has power against the alternative of \(Φ(||β_A - \hat{δ}||_Σ - z_{1-α})\), for \(z_{1-α}\) the \(1 - α\) quantile of the standard normal.

**Proof.** Define \(\prec, \succ, \succeq, \preceq, >Σ, >_Σ\) by \(x, y >_Σ x' Σ^{-1} y\), and observe that \(\prec, \succ, \succeq, \preceq, >_Σ\) is an inner product. The result then follows immediately from the discussion in Section 2.4.3 of Ingster and Suslina (2003), replacing all instances of the usual euclidean inner product with \(\prec, \succ, \succeq, \preceq, >_Σ\).

\[\square\]

**Lemma E.7.** Suppose \(Δ = \{δ : Aδ ≤ d\}\). As in the proof to Proposition 5.2, let \(ψ_{α}^{MP}(\hat{δ}; A, d, \tilde{θ}, Σ, δ_A, τ_A)\) be an indicator for whether the most powerful (Neyman-Pearson) test between the null hypothesis \(H_0 : δ ∈ Δ, θ = \tilde{θ}\) and the alternative \(H_A : (δ, τ) = (δ_A, τ_A)\) rejects the null, given the realization \(\hat{δ}\) which is assumed to come from a normal distribution with variance \(Σ\). Then for any \(n > 0\),

\[
ψ_{α}^{MP}(\hat{δ}; A, d, \tilde{θ}, Σ, δ_A, τ_A) = ψ_{α}^{MP}(\sqrt{n} \hat{δ}; A, \sqrt{n} d, \sqrt{n} \tilde{θ}, n Σ, \sqrt{n} δ_A, \sqrt{n} τ_A)
\]

**Proof.** As argued in the proof to Proposition 5.2, the null hypothesis \(\hat{δ} ∈ Δ, θ = \tilde{θ}\) is equivalent to the null \(H_0 : β ∈ B_0(\tilde{θ}, d) = \{β : ∃ τ_{post} \text{ s.t. } l' τ_{post} = \tilde{θ}, Aβ - d - AM_{post} τ_{post} ≤ 0\}\). Likewise, the alternative that \((δ, τ) = (δ_A, τ_A)\) is equivalent to \(H_A : β = δ_A + τ_A = ψ_A\). It is clear from the definition that \(B_0\) is convex. Thus, by Lemma E.6, the most powerful test of \(H_0\) against \(H_A\) when the covariance of \(\hat{δ}\) is \(Σ\) is a t-test that rejects for large values of \((β_A - \hat{δ})' Σ^{-1} \hat{δ}\), where \(\hat{δ}_1 = arg\ min_{β ∈ B_0(\tilde{θ}, d)} ||β - \hat{δ}||_Σ\). Its critical value is \((β_A - \hat{δ})' Σ^{-1} \hat{δ}_1 + z_{1-α} ||β_A - \hat{δ}_1||_Σ\), for \(z_{1-α}\) the \(1 - α\) quantile of the standard normal distribution.

Similarly, the null hypothesis \(δ ∈ Δ, Aδ ≤ d\) is equivalent to \(H_0 : β ∈ B_0(\sqrt{n} \tilde{θ}, \sqrt{n} d) = \{β : ∃ τ_{post} \text{ s.t. } l' τ_{post} = \sqrt{n} \tilde{θ}, Aβ - \sqrt{n} d - AM_{post} τ_{post} ≤ 0\}\). Likewise, the
alternative that \((\delta, \tau) = (\sqrt{n}\delta_A, \sqrt{n}\tau_A)\) is equivalent to \(H_A : \beta = \sqrt{n}\delta_A + \sqrt{n}\tau_A = \sqrt{n}\beta_A\). It is clear from the definition that \(B_0(\sqrt{n}\bar{\theta}, \sqrt{n}d)\) is convex. Thus, by Lemma E.6, the most powerful test of \(H_0\) against \(H_A\) when the covariance of \(\tilde{\beta}\) is \(n\Sigma\) is a \(t\)-test that rejects for large values of \((\sqrt{n}\beta_A - \tilde{\beta}_2)'(n\Sigma)^{-1}\tilde{\beta}_2\), where \(\tilde{\beta}_2 = \arg\min_{\beta \in B_0(\sqrt{n}\bar{\theta}, \sqrt{n}d)} \|\sqrt{n}\beta_A - \beta\|_{(n\Sigma)}\). Its critical value is \((\sqrt{n}\beta_A - \tilde{\beta}_2)'(n\Sigma)^{-1}\tilde{\beta}_2 + z_{1-\alpha}\|\beta_A - \tilde{\beta}_2\|_{(n\Sigma)}\).

Now, define

\[\eta(\hat{\beta}, A, d, \bar{\theta}, \Sigma) := \min_{\hat{\eta}, \hat{\tau}} \eta \text{ s.t. } A\hat{\beta} - d - \bar{A}_{(,1)}\bar{\beta} - \bar{A}_{(,-1)}\bar{\tau} \leq \eta\bar{\sigma},\]

where \(\bar{\sigma}\) is the square root of the diagonal elements of \(A\Sigma A'\). It follows immediately from the definition of \(B_0\) and the function \(\eta\) that we can write

\[B_0(\bar{\theta}, d) = \{\beta : \eta(\beta, A, d, \bar{\theta}, \Sigma) \leq 0\}\]
\[B_0(\sqrt{n}\bar{\theta}, \sqrt{n}d) = \{\beta : \eta(\beta, A, \sqrt{n}d, \sqrt{n}\bar{\theta}, n\Sigma) \leq 0\}\]

As argued in the proof to Lemma E.2 above, for any \(n > 0, \eta(\beta, A, d, \bar{\theta}, \Sigma) = \eta(\sqrt{n}\beta, A, \sqrt{n}d, \sqrt{n}\bar{\theta}, n\Sigma),\) from which it follows that \(\sqrt{n}B_0(\bar{\theta}, d) = B_0(\sqrt{n}\bar{\theta}, \sqrt{n}d)\). Thus,

\[\tilde{\beta}_2 = \arg\min_{\beta \in B_0(\sqrt{n}\bar{\theta}, \sqrt{n}d)} \|\sqrt{n}\beta_A - \beta\|_{(n\Sigma)} = \sqrt{n} \arg\min_{\beta \in B_0(\theta, d)} \|\sqrt{n}\beta_A - \sqrt{n}\beta\|_{(n\Sigma)} = \sqrt{n} \arg\min_{\beta \in B_0(\theta, d)} \|\beta_A - \beta\|_{\Sigma} = \sqrt{n}\tilde{\beta}_1;\]

where the third equality uses the fact that \(\|\sqrt{n}x\|_{(n\Sigma)} = \|x\|_\Sigma\), since \(\sqrt{n}x'(n\Sigma)^{-1}\sqrt{n}x = x'S^{-1}x\).

Thus, the test statistic used for \(\psi_\alpha^{MP}(\sqrt{n}\beta; A, \sqrt{n}d, \sqrt{n}\bar{\theta}, \sqrt{n}\delta_A, \sqrt{n}\tau_A)\) is

\[(\sqrt{n}\beta_A - \tilde{\beta}_2)'(n\Sigma)^{-1}(\sqrt{n}\beta) = (\sqrt{n}\beta_A - \sqrt{n}\tilde{\beta}_1)'(n\Sigma)^{-1}(\sqrt{n}\beta) = (\beta_A - \tilde{\beta}_1)'\Sigma^{-1}\tilde{\beta},\]

which is the test statistic used for \(\psi_\alpha^{MP}(\hat{\beta}; A, d, \bar{\theta}, \delta_A, \tau_A)\).

Likewise, the critical value used for \(\psi_\alpha^{MP}(\sqrt{n}\hat{\beta}; A, \sqrt{n}d, \sqrt{n}\bar{\theta}, \sqrt{n}\delta_A, \sqrt{n}\tau_A)\) is

\[\phi\alpha^{MP}(\beta_A - \sqrt{n}\tilde{\beta}_1)'\Sigma^{-1}(\sqrt{n}\tilde{\beta}_1 + z_{1-\alpha}\|\beta_A - \sqrt{n}\tilde{\beta}_1\|_{(n\Sigma)} \leq \phi\alpha^{MP}(\beta_A - \sqrt{n}\tilde{\beta}_1)'\Sigma^{-1}(\sqrt{n}\tilde{\beta}_1 + z_{1-\alpha}\|\beta_A - \sqrt{n}\tilde{\beta}_1\|_{(n\Sigma)}\)\]

which is the critical value used for \(\psi_\alpha^{MP}(\hat{\beta}; A, d, \bar{\theta}, \delta_A, \tau_A)\). We have thus shown that the test statistics and critical values for the two tests align, which gives the desired result.
Lemma E.8. Suppose Assumption 3 holds. Let $\theta^{ub} := \sup_{\theta} \mathcal{S}(\Delta, \delta_{A} + \tau_{A})$ and $\beta_{A} = \delta_{A} + \tau_{A}$. Then there exists a vector $\bar{\tau}^{ub} \in \mathbb{R}^{T-1}$ such that for $B = B(\delta^{**})$ as defined in Assumption 3,

\[
A_{(B,.)} \beta_{A} - d_{B} - \bar{A}_{(B,1)}(\theta^{ub} - \bar{A}_{(B,-1)} \bar{\tau}^{ub}) = 0
\]
\[
A_{(-B,.)} \beta_{A} - d_{-B} - \bar{A}_{(-B,1)}(\theta^{ub} - \bar{A}_{(-B,-1)} \bar{\tau}^{ub}) = -\epsilon < 0,
\]

for $\epsilon$ a vector with strictly positive entries. Additionally, the matrix $\bar{A}_{(B,-1)}$ has rank equal to $|B| - 1$, and $\{\gamma_{B} : \gamma'_{B} \bar{A}_{(B,-1)} = 0\} = \{c \bar{\gamma}_{B} : c \in \mathbb{R}\}$ for a non-zero vector $\bar{\gamma}_{B} \geq 0$.

Proof. From (10), we have that $\theta^{ub} = l' \beta_{A,post} - l' \delta^{**}_{post}$, for $\delta^{**}$ a solution to

\[
\min_{\delta} l' \delta_{post} \text{ s.t. } A \delta \leq d, \delta_{pre} = \delta_{A,pre}.
\]

(45)

Let $B = B(\delta^{**})$ index the binding inequalities of the optimization above at $\delta^{**}$, so that

\[
A_{(B,.)} \delta^{**} - d_{B} = 0
\]

(46)
\[
A_{(-B,.)} \delta^{**} - d_{-B} = -\epsilon < 0.
\]

(47)

By Assumption 3, $A_{(B,post)}$ has rank $|B|$.

Now, let $\tau^{**} = (\delta_{A,post} + \tau_{A,post}) - \delta^{**}_{post}$. Since by construction $\theta^{ub} \in \mathcal{S}_{\theta}(\Delta, \beta_{A})$, we have

\[
\begin{pmatrix}
\delta^{**}_{pre} \\
\delta^{**}_{post} + \tau^{**}
\end{pmatrix} = \begin{pmatrix}
\delta_{A,pre} \\
\delta_{A,post} + \tau_{A,post}
\end{pmatrix} = \beta_{A}.
\]

It follows that

\[
A \delta^{**} = A \beta_{A} - A M_{post} \tau^{**}
\]
\[
= A \beta_{A} - A M_{post} \Gamma^{-1} \Gamma \tau^{**}
\]
\[
= A \beta_{A} - A \bar{\bar{A}}_{(.,1)}(l' \tau^{**}) - A \bar{\bar{A}}_{(.,-1)} \Gamma_{(-1,.)} \tau^{**}
\]
\[
= A \beta_{A} - A \bar{\bar{A}}_{(.,1)}(\theta^{ub}) - A \bar{\bar{A}}_{(.,-1)} \bar{\tau}^{ub},
\]

where the third equality uses the definition of $\bar{\bar{A}}$ and the fact that the first row of $\Gamma$ is $l'$; and the fourth equality uses the fact that $\theta^{ub} = l'((\delta_{A,post} + \tau_{A,post}) - \delta^{**}_{post}) = l' \tau^{**}$ and defines $\bar{\tau}^{ub} := -\Gamma_{(-1,.)} \tau^{**}$. The first result then follows immediately from the previous display along with (46) and (47).

To show the second set of results, note that $\bar{A}_{(B,.)} = A_{(B,post)} \Gamma^{-1}$. Since $A_{(B,post)}$ has rank $|B|$ by assumption and $\Gamma^{-1}$ is full rank, $\bar{A}_{(B,.)}$ also has rank $|B|$. This implies that $\bar{A}_{(B,-1)}$ has rank of either $|B| - 1$ or $|B|$. To show that the rank must be $|B| - 1$, note that the optimization (45) can

S-28
and the set
\[ \tilde{\gamma} \]
\[ \text{s.t.} \quad A_{(\gamma_0)} \tilde{\gamma} = d - A_{(\gamma_0)} \tilde{\gamma}, \tilde{\gamma} \geq 0. \]

Since the optimization is assumed to have a finite solution, it is equivalent to its dual formulation,
\[ \max_{\gamma} \gamma' (A_{(\gamma_0)} \delta_{A_{(\gamma_0)}} - d) \quad \text{s.t.} \quad -\gamma' A_{(\gamma_0)} = l', \gamma \geq 0. \]

Let \( \tilde{\gamma} \) be a solution to the dual problem. Since \( \tilde{\gamma} \) is feasible in the dual,
\[ -\tilde{\gamma}' A_{(\gamma_0)} = l' \quad \text{and} \quad \tilde{\gamma} \geq 0. \]

Additionally, by the complementary slackness conditions, it must be that \( \gamma_{\text{res}} = 0 \). Hence, we have
\[ -\tilde{\gamma}' A_{(\gamma_0)} = l' \].

Multiplying on the right by \( \Gamma^{-1} \), we obtain \( -\tilde{\gamma}' \tilde{A}_{(\gamma_0)} = l' \Gamma^{-1} \). Recall that by construction the first row of \( \Gamma \) is \( l' \), so \( l' = e_1' \Gamma \), and thus \( -\tilde{\gamma}' \tilde{A}_{(\gamma_0)} = l' \Gamma^{-1} = e_1' \). This shows, however, that \( \gamma_{\text{res}} \) is in the nullspace of \( \tilde{A}_{(\gamma_0)} \) but not in the null space of \( \tilde{A}_{(\gamma_0)} \). It follows that the rank of \( \tilde{A}_{(\gamma_0)} \) is strictly less than that of \( \tilde{A}_{(\gamma_0)} \), and thus must be equal to \( |B| - 1 \). Since \( \tilde{A}_{(\gamma_0)} \) has \( |B| \) rows and rank \( |B| - 1 \), by the rank nullity theorem the set \( \{ \gamma_B : \gamma_B' \tilde{A}_{(\gamma_0)} = 0 \} \) must be one dimensional. We’ve shown that \( \gamma_B' \tilde{A}_{(\gamma_0)} = 0 \), and \( \gamma_B \neq 0 \) since \( \gamma_B' A_{(\gamma_0)} = -l' \neq 0 \), which implies that \( \{ \gamma_B : \gamma_B' \tilde{A}_{(\gamma_0)} = 0 \} = \{ c \gamma_B : c \in \mathbb{R} \} \), as needed.

\[ \square \]

**Lemma E.9.** Let \( \Delta = \{ \delta : A \delta \leq d \} \), and fix \( \delta_A \in \Delta, \tau_A, \) and \( \Sigma^* \) positive definite. If Assumption 3 holds, then for any \( x > 0 \),
\[ E_{(\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*)} \left[ \psi^C_{\alpha}(\tilde{\beta}; A, \sqrt{n} d, \theta_{ub} + x, \Sigma^*) \right] \rightarrow 1 - \Phi(z_{1-\alpha} - c^* x), \]
where \( \theta_{ub} := \sup S(\Delta_n, \sqrt{n} \delta_A + \sqrt{n} \tau_A) \), \( \Delta_n = \{ \delta : A \delta \leq \sqrt{n} d \} \), and \( c^* \) is a positive constant (not depending on \( x \) or \( \alpha \)). In particular, \( c^* = -\gamma_B' \tilde{A}_{(\gamma_0)} / \sigma_B \), where \( \sigma_B = \sqrt{\gamma_B' A_{(\gamma_0)} \Sigma^* A_{(\gamma_0)'}} \gamma_B \) and \( \gamma_B \) is the unique vector such that \( \gamma_B' \tilde{A}_{(\gamma_0)} = 0 \), \( \gamma_B \geq 0 \), \( \| \gamma_B \| = 1 \).

**Proof.** From Lemma E.8, there exists a vector \( \tilde{\gamma}_{ub}^* \) and a set of indices \( B \) such that
\[ A_{(\gamma_0)} \beta_A - d_B - \tilde{A}_{(\gamma_0)} \tilde{\gamma}_{ub}^* = 0 \]
and
\[ A_{(\gamma_0)} \beta_A - d_B - \tilde{A}_{(\gamma_0)} \tilde{\gamma}_{ub}^* = -\epsilon < 0, \]

and the set \( \{ \gamma_B \in \mathbb{R}^{|B|} : \gamma_B' \tilde{A}_{(\gamma_0)} = 0 \} = \{ c \gamma_B : c \in \mathbb{R} \} \) for some non-zero vector \( \gamma_B \geq 0 \), which without loss of generality we can normalize so that \( \| \gamma_B \| = 1 \). Let \( \tilde{\sigma} \) be the vector containing the square roots of the diagonal elements of \( A \Sigma^* A' \). It follows that the set \( \{ \gamma_B \in \mathbb{R}^{|B|} : \gamma_B' \tilde{A}_{(\gamma_0)} = 0, \gamma_B' \tilde{\sigma} = 1, \gamma_B \geq 0 \} \) is a singleton. In particular, its lone element is \( \gamma_B' := (\tilde{\sigma}^* \gamma_B)^{-1} \gamma_B \). Note that \( (\tilde{\sigma}^* \gamma_B)^{-1} \) is well-defined since \( \gamma_B \geq 0 \) and has at least one strictly positive element, and \( \tilde{\sigma} > 0 \) since by assumption \( A \) has no all-zero rows and \( \Sigma^* \) is positive definite.

Now, consider \( \psi^C_{\alpha}(\tilde{\beta}; A_{(\gamma_0)}, \sqrt{n} d_B, \theta_{ub} + x, \Sigma^*) \), the conditional test that uses only the moments
in $B$. The test statistic for the conditional test that uses only the moments in $B$ is

$$
\eta(\hat{\beta}_n, A_{(B,.)}, \sqrt{n}d_B, \theta^n_{ub} + x, \Sigma^*) = \min_{\eta, \hat{\tau}} \eta \\
\text{s.t. } A_{(B,.)}\hat{\beta}_n - \sqrt{n}d_B - \bar{A}_{(B,1)}(\theta^n_{ub} + x) - \bar{A}_{(B,-1)}\hat{\tau} \leq \eta \tilde{\sigma}_B.
$$

The equivalent dual problem is

$$
\max_{\gamma_B} \gamma_B' \bar{Y}_{B,n} \text{s.t. } \gamma_B' \bar{A}_{(B,-1)} = 0, \gamma_B \tilde{\sigma}_B = 1, \gamma_B \geq 0,
$$

where $\bar{Y}_{B,n} = A_{(B,.)}\hat{\beta}_n - \sqrt{n}d_B - \bar{A}_{(B,1)}(\theta^n_{ub} + x)$. We have shown, however, that there is a single value, $\gamma_B^*$, that satisfies the constraints of the dual problem, and so the solution to the problem in the previous display is $\gamma_B^* \bar{Y}_{B,n}$. Additionally, since the set of dual vertices is a singleton, the conditioning event that $\gamma_B^*$ is optimal is trivial, so $v^{lo} = -\infty$ and $v^{up} = \infty$. It follows that the conditional test using only the moments $B$ is a one-sided t-test that rejects for large values of $\gamma_B^* \bar{Y}_{B,n}$. Specifically, the critical value is $z_{1-\alpha}\sigma^*_B$, for $\sigma^*_B = \sqrt{\gamma_B^* A_{(B,.)} \Sigma^* A_{(B,.)}^*} \gamma_B^*$ the standard deviation of the test statistic $\gamma_B^* \bar{Y}_{B,n}$. We claim that $\sigma^*_B > 0$. To see why this is the case, observe that Assumption 3 implies that $A_{(B,.)}$ has full row rank, and by construction $\gamma_B^* \neq 0$, so $\gamma_B^* A_{(B,.)} \neq 0$. That $\sigma^*_B > 0$ then follows from the fact that $\Sigma^*$ is positive definite.

Additionally, observe that

$$
\mathbb{E}(\sqrt{n}d_A, \sqrt{n}d_A, \Sigma^*) [\gamma_B^* \bar{Y}_{B,n}] = \gamma_B^* \left[ A_{(B,.)} \sqrt{n}d_A - \sqrt{n}d_B - \bar{A}_{(B,1)}(\theta^n_{ub} + x) \right] \\
= \gamma_B^* \sqrt{n}d \left[ A_{(B,.)} \hat{\beta}_n - d_B - \bar{A}_{(B,1)}(\theta^n_{ub}) \right] - \gamma_B^* \bar{A}_{(B,1)}x \\
= \gamma_B^*[\sqrt{n}\bar{A}_{(B,-1)}\hat{\tau}^n_{ub}] - \gamma_B^* \bar{A}_{(B,1)}x \\
= -\gamma_B^* \bar{A}_{(B,1)}x
$$

where the second equality uses $\theta^n_{ub} = \sqrt{n}d^{ub}$ from Lemma E.1, the third equality uses (49) to substitute for the term in brackets, and the final equality follows from the fact that $\gamma_B^* \bar{A}_{(B,-1)} = 0$ by construction. Thus, regardless of $n$, the conditional test using only the moments in $B$ rejects with probability

$$
\mathbb{E}(\sqrt{n}d_A, \sqrt{n}d_A, \Sigma^*) [\psi_\alpha^C(\hat{\beta}_n, A_{(B,.)}, \sqrt{n}d_B, \theta^n_{ub} + x, \Sigma^*)] = 1 - \Phi(z_{1-\alpha} - (-\gamma_B^* \bar{A}_{(B,1)}/\sigma^*_B \cdot x)). \quad (51)
$$

Note also that we showed in the proof of Lemma E.8 that $-\gamma_B^* \bar{A}_{(B,1)} = \epsilon'_1$, which implies that $-\gamma_B^* \bar{A}_{(B,1)} = 1$, and hence $c^* := -\gamma_B^* \bar{A}_{(B,1)}/\sigma^*_B > 0$ since $\gamma_B^*$ is a positive multiple of $\tilde{\gamma}_B$. Moreover, observe that if we define $\sigma_B = \sqrt{\gamma_B^* A_{(B,.)} \Sigma^* A_{(B,.)}^*} \gamma_B$, then $\gamma_B^*/\sigma_B = \gamma_B/\sigma_B$, so $c^* = -\gamma_B^* \bar{A}_{(B,1)}/\sigma_B$.

Recall that $\psi_\alpha^C(\hat{\beta}_n; A, \sqrt{n}d, \theta^n_{ub} + x, \Sigma^*) = \psi_\alpha^C(\tilde{Y}_n, A \Sigma^* A')$ for $\tilde{Y}_n = A \hat{\beta}_n - \sqrt{n}d - \bar{A}_{(.,1)}(\theta^n_{ub} + x)$. Since the conditional test optimizes over $\hat{\tau}$, and $\hat{\tau}$ appears in this optimization only in the term $\bar{A}_{(.,-1)}\hat{\tau}$, the result of the conditional test using $\tilde{Y}_n$ is equivalent to the result of the conditional test.
Differentiating with respect to the norm, it follows that we must have any $\theta \in B$ moments

\[
\Delta \left[ \frac{\partial}{\partial \theta} \psi^C_n \right] = \Delta \left[ \frac{\partial}{\partial \theta} \psi^C_n \right] \times B
\]

for all $B$, whereas $\frac{\partial}{\partial \theta} \psi^C_n \left( \hat{Y}_n, A \Sigma^* A' \right)$. The expectation of the elements of $\hat{Y}_n$ corresponding with the rows $B$ is

\[
\mathbb{E}_{\Delta [\tilde{\theta}]} \left[ \frac{\partial}{\partial \theta} \psi^C_n \right] = A_{(B, \cdot)} \Delta [\tilde{\theta}] - A_{(B, \cdot)} \Delta [\tilde{\theta}] - A_{(B, \cdot)} \Delta [\tilde{\theta}]
\]

where the second line uses the fact that $\hat{\theta}_{\Delta B} = \sqrt{n} \hat{\theta}_{\Delta B}$ from Lemma E.1, and the third uses (49). Similarly, the expectation of the elements of $\hat{Y}_n$ corresponding to the rows other than $B$ is

\[
\mathbb{E}_{\Delta [\tilde{\theta}]} \left[ \frac{\partial}{\partial \theta} \psi^C_n \right] = A_{(-B, \cdot)} \Delta [\tilde{\theta}] - A_{(-B, \cdot)} \Delta [\tilde{\theta}]
\]

where the last line uses (50). Since $-\epsilon < 0$, all of the elements of $\mathbb{E} \left[ \hat{Y}_n_{(-B)} \right]$ converge to $-\infty$ as $n \to \infty$, whereas $\mathbb{E} \left[ \hat{Y}_n_{(B)} \right]$ does not depend on $n$. It follows from Proposition 3 in ARP that the conditional test based on the full set of moments is equal to the conditional test that only uses the moments $B$ with probability approaching one,

\[
\lim_{n \to \infty} \mathbb{P} \left[ \Delta [\tilde{\theta}] \left( \psi^C_n (\hat{\beta}_n; \hat{Y}_n) \left( \psi^C_n (\hat{\beta}_n, A_{(B, \cdot)} \hat{Y}_n, \theta_{\Delta B} + x, \Sigma^*) \right) = \psi^C_n (\hat{\beta}_n, A_{(B, \cdot)} \hat{Y}_n, \theta_{\Delta B} + x, \Sigma^*) \right) = 1.
\]

This, combined with (51), gives the desired result.

\[\square\]

**Lemma E.10.** Let $B$ be a closed, convex subset of $\mathbb{R}^K$, and $\beta_A \notin B$. Let $\beta = \arg \min_{\beta \in B} ||\beta - \beta_A||_\Sigma$, where $||x||_\Sigma^2 = x^T \Sigma^{-1} x$ for some positive definite matrix $\Sigma$. Then for any $\beta \in B$, $(\beta - \beta_A)^T \Sigma^{-1} (\beta - \beta) \geq 0$.

**Proof.** Consider any $\beta \in B$. Define $\beta_0 = \theta (\beta - \beta) + \tilde{\beta}$, and note that since $B$ is convex $\beta_0 \in B$ for any $\theta \in [0,1]$. Further,

\[
||\beta_0 - \beta_A||_\Sigma^2 = \theta^2 ||\beta - \beta||_\Sigma^2 + 2 \theta (\beta - \beta_A)^T \Sigma^{-1} (\beta - \tilde{\beta}) + ||\tilde{\beta} - \beta_A||_\Sigma^2.
\]

Differentiating with respect to $\theta$, we have

\[
\frac{\partial}{\partial \theta} ||\beta_0 - \beta_A||_\Sigma^2 = 2 \theta ||\beta - \tilde{\beta}||_\Sigma^2 + 2 (\beta - \beta_A)^T \Sigma^{-1} (\beta - \tilde{\beta}),
\]

from which we see that the derivative evaluated at $\theta = 0$ is $2 (\beta - \beta_A)^T \Sigma^{-1} (\beta_A - \tilde{\beta})$. Since $\beta$ minimizes the norm, it follows that we must have $2 (\beta - \beta_A)^T \Sigma^{-1} (\beta_A - \tilde{\beta}) \geq 0$, else we could achieve a lower
value of the norm at $\beta_0$ by choosing $\theta$ sufficiently small.

Lemma E.11. Let $B = \{\beta \in \mathbb{R}^K : v^T \beta \leq d\}$ for some $v \in \mathbb{R}^K \setminus \{0\}$ and $d \in \mathbb{R}$. Let $\tilde{\beta} = \arg\min_{\beta \in B} ||\beta - \beta_A||_{\Sigma}$ for some $\beta_A \notin B$, where $||x||_{\Sigma}^2 = x^T \Sigma^{-1} x$ and $\Sigma$ is positive definite. Then $(\beta_A - \tilde{\beta})^T \Sigma^{-1} = c \cdot v^T$ for the positive constant $c = \frac{v^T \beta_A - d}{\sqrt{\Sigma}}$.

Proof. Note that we can form a basis $v, \tilde{v}_2, ..., \tilde{v}_K$ such that $v^T \tilde{v}_j = 0$ for $j = 2, ..., K$. It follows by construction that for any $j = 2, ..., K$ and any $t \in \mathbb{R}$, $\tilde{\beta} + t \cdot \tilde{v}_j \in B$. Hence, from Lemma E.10, $-(\beta_A - \tilde{\beta})^T \Sigma^{-1} (t \tilde{v}_j) \geq 0$. Since we can choose $t$ both positive and negative, it follows that $(\beta_A - \tilde{\beta})^T \Sigma^{-1} \tilde{v}_j = 0$ for all $j$. Since $(\beta_A - \tilde{\beta})^T \Sigma^{-1}$ is orthogonal to $\{\tilde{v}_2, ..., \tilde{v}_K\}$, and $\{v, \tilde{v}_2, ..., \tilde{v}_K\}$ form a basis, we have that $(\beta_A - \tilde{\beta})^T \Sigma^{-1} = c \cdot v^T$, for some $c \in \mathbb{R}$. Multiplying both sides of the equation on the right by $\Sigma v$, we obtain that $(\beta_A - \tilde{\beta})^T v = c \cdot v^T \Sigma v$. However, since $\tilde{\beta}$ is the closest point to $\beta_A$ in Mahalanobis distance, it must be on the boundary of $B$, and so $v^T \tilde{\beta} = d$. It follows that $c = (v^T \beta_A - d)/(v^T \Sigma v)$, which is clearly positive since $\beta_A \notin B$ and thus $v^T \beta_A > d$. 

Lemma E.12. Let $B = \{\beta \in \mathbb{R}^K : v^T \beta \leq d\}$ for some $v \in \mathbb{R}^K \setminus \{0\}$ and $d \in \mathbb{R}$. Suppose $\tilde{\beta} \sim \mathcal{N}(\beta, \Sigma)$ for $\Sigma$ positive definite known, and consider the problem of testing $H_0 : \beta \in B$ against $H_A : \beta = \beta_A$ for some $\beta_A \notin B$. Then the most powerful size-$\alpha$ test of $H_0$ against $H_A$ is a one-sided t-test that rejects for large values of $v^T \tilde{\beta}$, and has power equal to $\Phi((v^T \beta_A - d)/\sqrt{v^T \Sigma v} - z_{1-\alpha})$.

Proof. From Lemma E.6, the most powerful test rejects for large values of $(\beta_A - \tilde{\beta})^T \Sigma^{-1} \tilde{\beta}$, where $\tilde{\beta} = \arg\min_{\beta \in B} ||\beta - \beta_A||_{\Sigma}$, and has power $\Phi(||\beta_A - \tilde{\beta}||_{\Sigma} - z_{1-\alpha})$. By Lemma E.11, $(\beta_A - \tilde{\beta})^T \Sigma^{-1} = c v^T$, for $c = (v^T \beta_A - d)/(v^T \Sigma v)$. It follows that

$$||\beta_A - \tilde{\beta}||_{\Sigma}^2 = (\beta_A - \tilde{\beta})^T \Sigma^{-1} (\beta_A - \tilde{\beta}) = c v^T (\beta_A - \tilde{\beta}) = c (v^T \beta_A - d) = (v^T \beta_A - d)^2/(v^T \Sigma v),$$

where we use the fact that $v^T \tilde{\beta} = d$, since $\tilde{\beta}$ must be on the boundary of $B$, as argued in the proof to Lemma E.11. The result then follows immediately.

Lemma E.13. Let $\Delta = \{\delta : A\delta \leq d\}$, and fix $\delta_A \in \Delta$, $\tau_A$, and $\Sigma^*$ positive definite. If Assumption 3 holds, then for any $x > 0$,

$$\mathbb{E}_{(\bar{\nu}_\delta A, \bar{\nu}_\tau A, \Sigma^*)} \left[ \psi^{\text{MP}}_\alpha (\hat{\beta}_n; A, \sqrt{n} d, \theta_{\text{ub}}^n + x, \Sigma^*) \right] \to 1 - \Phi(z_{1-\alpha} - c^* x),$$

where $\psi^{\text{MP}}_\alpha$ is as defined in in the proof to Proposition 5.2, $\theta_{\text{ub}}^n := \sup S(\Delta, \sqrt{n} \delta_A + \sqrt{n} \tau_A)$, $\Delta = \{\delta : A\delta \leq \sqrt{n} d\}$, and $c^*$ is the same positive constant as in Lemma E.9.

Proof. As argued in the proof to Lemma E.7, the null hypothesis $H_0 : \theta = \bar{\theta}, \delta \in \{A\delta \leq d\}$ is equivalent to the null $H_0 : \beta \in B(\bar{\theta}, d) = \{\beta : \exists \tau_{\text{post}} s.t. v^T \tau_{\text{post}} = \bar{\theta}, A\beta - d = A\beta_{\text{post}} \tau_{\text{post}} \leq 0\}$, which we showed in Lemma E.7 to be equivalent to $B_{0}(\bar{\theta}, d) = \{\beta : \eta(\beta, A, d, \bar{\theta}, \Sigma^*) \leq 0\}$ for the function $\eta$ as
where \( t \) is defined in (45). Thus, the null hypothesis for the test associated with \( \psi_{\alpha}^{MF}(\beta_n; A, \sqrt{n}d, \theta^{ub}_n + x, \Sigma^*) \) can be written as \( H_0 : \beta_n \in B_{n,0} := \{ \beta : \eta(\beta, A, \sqrt{n}d, \theta^{ub}_n + x, \Sigma^*) \leq 0 \} \). Under the alternative for this test, \( \beta_n = \sqrt{n}\beta_A \), so by Lemma E.6 the most powerful test uses the test statistic 
\[
(\sqrt{n}\beta_A - \bar{\beta}_n)'\Sigma^{-1}\bar{\beta}_n,
\]
where \( \bar{\beta}_n = \arg\min_{\beta \in B_n} ||\beta - \sqrt{n}\beta_A||_{\Sigma^*} \).

Now, from Lemma E.8, there exists a vector \( \tilde{\gamma}_B \) and a set of indices \( B \) such that
\[
A_{(B)}\beta_A - d_B - \tilde{A}_{(B,1)}\gamma^{ub} - \tilde{A}_{(B,1)}\tilde{\gamma}^{ub} = 0
\]
(52)
\[
A_{(B,-)}\beta_A - d_B - \tilde{A}_{(B,-)}\gamma^{ub} - \tilde{A}_{(B,-)}\tilde{\gamma}^{ub} = -\epsilon < 0,
\]
(53)
where \( \{\gamma_B \in \mathbb{R}^{|B|} : \gamma'_B \tilde{A}_{(B,-)} = 0\} = \{c\tilde{\gamma}_B : c \in \mathbb{R}\} \) for some non-zero vector \( \tilde{\gamma}_B \geq 0 \). Define
\[
B^B_n := \{ \beta : \eta(\beta, A_{(B)}, \sqrt{n}d_B, \theta^{ub}_n + x, \Sigma^*) \leq 0 \},
\]
the analog to \( B_{n,0} \) that restricts attention only to the set of moments \( B \). By an argument analogous to that in the proof to Lemma E.9 (replacing \( \tilde{Y} \) with \( \mu \)), we can show that \( \eta(\beta, A_{(B)}, \sqrt{n}d_B, \theta^{ub}_n + x, \Sigma^*) = \gamma^{st}_B\mu_{B,n}(\beta) \), where \( \mu_{B,n}(\beta) = A_{(B)}\beta - \sqrt{n}d_B - \tilde{A}_{(B,1)}(\gamma^{ub} + x) \) and \( \gamma^{st}_B = (\tilde{\gamma}_B\tau)^{-1}\tilde{\gamma}_B \). Note also that (52) implies that \( \tilde{A}_{(B,1)}\gamma^{ub} = A_{(B)}\beta_A - d_B - \tilde{A}_{(B,1)}\tilde{\gamma}^{ub} \). Substituting into the expression for \( \mu_{B,n}(\beta) \) and using the fact that \( \theta^{ub}_n = \sqrt{n}\theta^{ub}_1 \) by Lemma E.1, we obtain \( \mu_{B,n}(\beta) = A_{(B)}(\beta - \sqrt{n}\beta_A - \tilde{A}_{(B,1)}x + \sqrt{n}\tilde{A}_{(B,1)}\tilde{\gamma}^{ub} \). Since \( \gamma^{st}_B \tilde{A}_{(B,-)} = 0 \) by construction, this implies that \( \gamma^{st}_B\mu_{B,n}(\beta) = \gamma^{st}_B(A_{(B)}(\beta - \sqrt{n}\beta_A - \tilde{A}_{(B,1)}x) \).

Hence,
\[
B^B_n = \{ \beta : \eta(\beta, A_{(B)}, \sqrt{n}d_B, \theta^{ub}_n + x, \Sigma^*) \leq 0 \} = \{ \beta : \gamma^{st}_B(A_{(B)}(\beta - \sqrt{n}\beta_A - \tilde{A}_{(B,1)}x) \leq 0 \} = \{ \beta : (\beta - (\sqrt{n} - 1)\beta_A) \in B^B_1 \} = (\sqrt{n} - 1)\beta_A + B^B_1.
\]

Now, define \( \beta^*_n = \arg\min_{\beta \in B^B_n} ||\beta - \sqrt{n}\beta_A||_{\Sigma^*} \). The results above imply that
\[
\beta^*_n = \arg\min_{\beta \in B^B_n} ||\beta - \sqrt{n}\beta_A||_{\Sigma^*} = \arg\min_{\beta \in B^B_1} ||\beta - \sqrt{n}\beta_A||_{\Sigma^*} = (\sqrt{n} - 1)\beta_A + \beta^*_1.
\]
Observe that \( B^B_n \equiv B_{n,0} \) since \( B^B_n \) is the set of values \( \beta \) that are consistent with a subset of the \( m \) moments used in \( B_{n,0} \) (formally, \( \eta(\beta, A_{(B)}, \sqrt{n}d_B, \theta^{ub}_n + x, \Sigma^*) \leq \eta(\beta, A, \sqrt{n}d_B, \theta^{ub}_n + x, \Sigma^*) \) since the RHS minimizes the same objective function subject to additional constraints). Thus, \( \beta^*_n = \bar{\beta}_n \) iff
\( \beta_n^* \in B_{n, 0} \). From the definition of \( B_{n, 0} \), this occurs if there exists a value \( \tilde{\tau}_n \) such that

\[
A_{(B, \cdot)}\beta_n^* - \sqrt{n}d_B - \tilde{A}_{(B, 1)}(\theta_{n}^{ub} + x) - \tilde{A}_{(B, -1)}\tilde{\tau}_n \leq 0
\]

\[
A_{(-B, \cdot)}\beta_n^* - \sqrt{n}d_B - \tilde{A}_{(-B, 1)}(\theta_{n}^{ub} + x) - \tilde{A}_{(-B, -1)}\tilde{\tau}_n \leq 0
\]

Now, since \( \beta_n^* \in B_1^B \), there exists a value \( \tilde{\tau}_n^* \) such that

\[
A_{(B, \cdot)}\beta_1^* - d_B - \tilde{A}_{(B, 1)}(\theta_{1}^{ub} + x) - \tilde{A}_{(B, -1)}\tilde{\tau}_n^* \leq 0.
\]

It follows that

\[
A_{(B, \cdot)}\beta_n^* - \sqrt{n}d_B - \tilde{A}_{(B, 1)}(\theta_{n}^{ub} + x) - \tilde{A}_{(B, -1)}(\tilde{\tau}_1^* + (\sqrt{n} - 1)\tilde{\tau}_n^{ub})
\]

\[
= A_{(B, \cdot)}\beta_n^* - d_B - \tilde{A}_{(B, 1)}(\theta_{1}^{ub} + x) - \tilde{A}_{(B, -1)}\tilde{\tau}_1^* + (\sqrt{n} - 1)\left[ A_{(B, \cdot)}\beta_A - d_B - \tilde{A}_{(B, 1)}\theta_{1}^{ub} - \tilde{A}_{(B, -1)}\tilde{\tau}_1^{ub} \right]
\]

\[
= A_{(B, \cdot)}\beta_n^* - d_B - \tilde{A}_{(B, 1)}(\theta_{1}^{ub} + x) - \tilde{A}_{(B, -1)}\tilde{\tau}_1^* \leq 0,
\]

where the first equality uses the fact that \( \theta_n^{ub} = \sqrt{n}\theta_1^{ub} \) by Lemma E.1 and \( \beta_n^* = \beta_1^* + (\sqrt{n} - 1)\beta_A \) as shown above, and the second equality uses (52).

Similarly, we have

\[
A_{(-B, \cdot)}\beta_n^* - \sqrt{n}d_B - \tilde{A}_{(-B, 1)}(\theta_{n}^{ub} + x) - \tilde{A}_{(-B, -1)}(\tilde{\tau}_1^* + (\sqrt{n} - 1)\tilde{\tau}_n^{ub})
\]

\[
= A_{(-B, \cdot)}\beta_n^* - d_B - \tilde{A}_{(-B, 1)}(\theta_{1}^{ub} + x) - \tilde{A}_{(-B, -1)}\tilde{\tau}_1^* + (\sqrt{n} - 1)\left[ A_{(-B, \cdot)}\beta_A - d_B - \tilde{A}_{(-B, 1)}\theta_{1}^{ub} - \tilde{A}_{(-B, -1)}\tilde{\tau}_1^{ub} \right]
\]

\[
= \left[ A_{(-B, \cdot)}\beta_1^* - d_B - \tilde{A}_{(-B, 1)}(\theta_{1}^{ub} + x) - \tilde{A}_{(-B, -1)}\tilde{\tau}_1^* \right] - (\sqrt{n} - 1)\epsilon,
\]

for \( \epsilon \) a vector with strictly positive elements, where the first equality again uses that \( \theta_n^{ub} = \sqrt{n}\theta_1^{ub} \) and \( \beta_n^* = \beta_1^* + (\sqrt{n} - 1)\beta_A \), and the second equality uses (53). Since the term in brackets in the final expression in the previous display does not depend on \( n \) and all elements of the final term go to \(-\infty\), for \( n \) sufficiently large the expression in the previous display will be less than or equal to 0. Thus, for \( n \) sufficiently large, \( \beta_n^* = \tilde{\beta}_n \), and hence the MP test of \( H_0 : \beta \in B_{n, 0} \) against \( H_A : \beta = \sqrt{n}\beta_A \) is equivalent to the most powerful test of \( H_0 : \beta \in B_1^B \) against \( H_A : \beta = \sqrt{n}\beta_A \).

We showed earlier in the proof that \( B_1^B = \{ \beta : v'\beta \leq d_n \} \), for \( v = \gamma_B^*A_{(B, \cdot)} \) and \( d_n = \gamma_B^*A_{(B, 1)}x + v'\sqrt{n}\beta_A \). From Lemma E.12, the MP test of \( H_0 : \beta \in B_n^B \) against \( H_A : \beta = \beta_A \) has power equal to \( \Phi((v'\sqrt{n}\beta_A - d_n)/(v'S^*v) - z_{1-\alpha}) \). Plugging in the definitions of \( v \) and \( d \) and cancelling like terms, we obtain that the power of the test is \( \Phi(-\gamma_B^*A_{(B, 1)}x/\sigma_B^* - z_{1-\alpha}) \), for \( \sigma_B^* = \sqrt{\gamma_B^*A_{(B, \cdot)}S^*A_{(-B, \cdot)}^*\beta_A^*}/\beta_A^* \), which coincides with the expression for the limiting power of the conditional test in Lemma E.9, as needed.

**Lemma E.14.** Let \( \eta(\beta, A, d, \theta, \Sigma) \) be as defined in (45). Fix \( \Sigma^* \) positive definite. For any \( \delta_A, \tau_A, \) and \( d \), let \( \beta_A(\delta_A, \tau_A) = \delta_A + \tau_A \) and \( \theta^{ub}(\delta_A, \tau_A, d) = \sup S(\Delta, \delta_A + \tau_A) \) for \( \Delta = \{ \delta : A\delta \leq d \} \). Let
\( \eta^*(x; \delta_A, \tau_A, d) := \eta(\beta_A(\delta_A, \tau_A), A, d, \theta^{ub}(\delta_A, \tau_A, d) + x, \Sigma^*) \). Then there exists a scalar \( c(\Sigma^*, A) > 0 \) such that \( \eta^*(x; \delta_A, \tau_A, d) \geq c(\Sigma^*, A) \cdot x \) for all \( \delta_A, \tau_A \), and \( d \).

**Proof.** Observe that \( \eta(\beta_A, A, d, \bar{\theta}, x) \) is equivalent to the linear program

\[
\min_{\eta, \tau} \eta \text{ s.t. } A\beta_A - d - A\tau \leq \eta \bar{\sigma}, \ l^T = \bar{\theta}.
\]

The dual formulation for this problem is

\[
\max_{\gamma} \left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} \end{array} \right) \text{ s.t. } \gamma_A A + \gamma_{\theta} \bar{\sigma} = 0, \ \gamma_A A^T = 1, \ \gamma_A \geq 0,
\]

where \( \gamma_A \) is a vector with length equal to the number of rows of \( A \), and \( \gamma_{\theta} \) is a scalar. Note that the feasible set for the dual depends on \( A \) and \( \Sigma^* \) but not on \( d, \delta_A, \) or \( \tau_A \). Let \( V_D \) denote the set of vertices of the dual, which is finite, and recall that maximizing over the feasible set is equivalent to maximizing over the set of vertices.

Now, we first claim that \( \eta(\beta_A, A, d, \theta^{ub}, \Sigma^*) = 0 \). Note that since \( \theta^{ub} \) is in the identified set, it must be that \( \eta(\beta_A, A, d, \theta^{ub}, \Sigma^*) \leq 0 \). Towards contradiction, suppose that \( \eta(\beta_A, A, d, \theta^{ub}, \Sigma^*) = -\epsilon_1 < 0 \). Then for all \( \gamma = \left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right) \in V_D,
\]

\[
\left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} \end{array} \right) \leq -\epsilon_1.
\]

Since \( V_D \) is finite, \( \gamma_{\theta} := \max_{\gamma \in V_D} \gamma_{\theta} \) is finite. But then for \( \epsilon_2 > 0 \),

\[
\left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} + \epsilon_2 \end{array} \right) \leq -\epsilon_1 + \gamma_{\theta} \epsilon_2.
\]

By choosing \( \epsilon_2 \) sufficiently small, we can make the upper bound in the previous display less than or equal to 0. However, this implies that \( \eta(\beta_A, A, d, \theta^{ub} + \epsilon_2, \Sigma^*) \leq 0 \). But this in turn implies \( \theta^{ub} + \epsilon_2 \) is in the identified set, which contradicts \( \theta^{ub} \) being maximal. Therefore, \( \eta(\beta_A, A, d, \theta^{ub}, \Sigma^*) = 0 \).

Additionally, we claim that for \( \theta = \theta^{ub} \), there must be an optimal dual vertex with \( \gamma_{\theta} > 0 \). Towards contradiction, suppose not. Then there exists \( \epsilon_3 > 0 \) such that for all \( \gamma = \left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right) \in V_{D,+} := \{ \gamma \in V_D : \gamma_{\theta} > 0 \},
\]

\[
\left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} \end{array} \right) < -\epsilon_3. \text{ Letting } \epsilon_4 = \epsilon_3/\max_{\gamma \in V_{D,+}} \gamma_{\theta}, \text{ it follows that for all } \gamma \in V_{D,+},
\]

\[
\left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} + \epsilon_4 \end{array} \right) < 0. \text{ Additionally, for } \gamma = \left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right) \in V_D \setminus V_{D,+}, \gamma_{\theta} \leq 0, \text{ and so}
\]

\[
\left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} + \epsilon_4 \end{array} \right) \leq \left( \begin{array}{c} \gamma_A \\ \gamma_{\theta} \end{array} \right)' \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} \end{array} \right) \leq 0.
\]
Thus, \( \left( \begin{array}{c} \gamma_A \\ \gamma_\theta \end{array} \right) \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} + \epsilon_4 \end{array} \right) \) \( \leq 0 \) for all \( \gamma \in V_D \), and so \( \eta(\beta_A, A, d, \theta^{ub} + \epsilon_4, \Sigma^*) \) \( \leq 0 \). However, this implies that \( \theta^{ub} + \epsilon_4 \) is in the identified set, which contradicts \( \theta^{ub} \) being maximal. Thus, there must be at least one \( \gamma^* \in V_{D,+} \) such that

\[
\left( \begin{array}{c} \gamma^*_A \\ \gamma^*_\theta \end{array} \right) \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} \end{array} \right) = 0.
\]

Since \( \gamma^* \) remains feasible in the dual with \( \hat{\theta} = \theta^{ub} + x \), it follows that \( \eta(\beta_A, A, d, \theta^{ub} + x, \Sigma^*) \) is lower bounded by

\[
\left( \begin{array}{c} \gamma^*_A \\ \gamma^*_\theta \end{array} \right) \left( \begin{array}{c} A\beta_A - d \\ \theta^{ub} + x \end{array} \right) = \gamma^*_\theta \cdot x.
\]

Note that the choice of \( \gamma^* \in V_{D,+} \) depended on \( d, \delta_A, \) and \( \tau_A \). However, as noted earlier in the proof, the set \( V_{D,+} \) depends on \( A \) and \( \Sigma^* \) but does not on \( d, \delta_A, \tau_A \). Since \( V_{D,+} \) is finite and \( \gamma_\theta > 0 \) for all \( \gamma \in V_{D,+} \), there is a value \( c > 0 \) such that \( \gamma_\theta \geq c \) for all \( \gamma \in V_{D,+} \). Hence, \( \eta^*(x; \delta_A, \tau_A, d) \geq c \cdot x \) for all \( \delta_A, \tau_A, d \), as needed.

\[\Box\]

**Lemma E.15.** Let \( \alpha \in (0, 1) \) and \( c > z_{1-\alpha} \). Then there exists a unique constant \( \zeta(c) > 0 \) such that

\[
\frac{\Phi(c) - \Phi(c - \zeta(c))}{1 - \Phi(c - \zeta(c))} = 1 - \alpha.
\]

Additionally, for any values \( z_{lo} < z_{up} \), with \( z_{lo} \) and \( z_{up} \) potentially infinite-valued, and \( \eta > \max\{c, z_{lo} + \zeta(c)\} \),

\[
F_{\xi, \eta} [z_{lo}, z_{up}] (\cdot) > 1 - \alpha,
\]

where \( F_{\xi, \eta} [z_{lo}, z_{up}] (\cdot) \) is the CDF of \( \xi \sim \mathcal{N}(0, 1) \) truncated to \([z_{lo}, z_{up}]\).

**Proof.** First, we show that \( F_{\xi, \eta} [z_{lo}, z_{up}] (t) \) is increasing in \( t \) and decreasing in \( z_{lo} \) and \( z_{up} \), and these comparative statics are strict for \( t \in (z_{lo}, z_{up}) \). To see this, note that

\[
F_{\xi, \eta} [z_{lo}, z_{up}] (t) = \begin{cases} 0 & \text{for } t \leq z_{lo} \\ \frac{\Phi(t) - \Phi(z_{lo})}{\Phi(z_{up}) - \Phi(z_{lo})} & \text{for } t \in (z_{lo}, z_{up}) \\ 1 & \text{for } t \geq z_{up} \end{cases}
\]

It is immediate that \( F_{\xi, \eta} [z_{lo}, z_{up}] (t) \) is increasing in \( t \) and decreasing in \( z_{up} \), and strictly so when \( t \in (z_{lo}, z_{up}) \). Additionally, we have

\[
\frac{\partial}{\partial z_{lo}} \frac{\Phi(t) - \Phi(z_{lo})}{\Phi(z_{up}) - \Phi(z_{lo})} = -\frac{\phi(z_{lo}) (\Phi(z_{up}) - \Phi(t))}{(\Phi(z_{up}) - \Phi(z_{lo}))^2},
\]

which is clearly negative for \( t \in (z_{lo}, z_{up}) \), which gives the desired result for \( z_{lo} \).
Next, consider the function
\[ f(\zeta) = \frac{\Phi(c) - \Phi(c - \zeta)}{1 - \Phi(c - \zeta)}. \]

Observe that \( f(0) = 0 \) and \( \lim_{\zeta \to \infty} f(\zeta) = \Phi(c) > 1 - \alpha \). Additionally, the derivative in the previous paragraph (with \( z_{up} = \infty \)) implies that \( \frac{d}{d\zeta} f(\zeta) > 0 \) for \( \zeta > 0 \). It follows that there is a unique value \( \zeta(c) > 0 \) such that
\[ f(c, \zeta(c)) = \frac{\Phi(c) - \Phi(c - \zeta(c))}{1 - \Phi(c - \zeta(c))} = 1 - \alpha, \]
which gives the first result.

Next, we claim that for \( z_{lo} \in (-\infty, \infty) \) and \( \zeta > 0 \), \( F_{\xi|\xi \in [z_{lo}, \infty)}(z_{lo} + \zeta) \) is increasing in \( z_{lo} \). To see why this is the case, note that
\[ F_{\xi|\xi \in [z_{lo}, \infty)}(z_{lo} + \zeta) = \frac{\Phi(z_{lo} + \zeta) - \Phi(z_{lo})}{1 - \Phi(z_{lo})}. \]

Differentiating with respect to \( z_{lo} \), we obtain
\[ \frac{\phi(z_{lo} + \zeta)(1 - \Phi(z_{lo})) - \phi(z_{lo})(1 - \Phi(z_{lo} + \zeta))}{[1 - \Phi(z_{lo})]^2}, \]
which is greater than zero iff
\[ \frac{\phi(z_{lo} + \zeta)}{1 - \Phi(z_{lo} + \zeta)} > \frac{\phi(z_{lo})}{1 - \Phi(z_{lo})}, \]
which holds since the normal hazard function is strictly increasing.

Now, suppose that \( \eta > \max\{c, z_{lo} + \zeta(c)\} \). Then
\[
F_{\xi|\xi \in [z_{lo}, z_{up})}(\eta) \geq F_{\xi|\xi \in [z_{lo}, \infty)}(\eta) \\
> F_{\xi|\xi \in [z_{lo}, \infty)}(\max\{c, z_{lo} + \zeta(c)\}) \\
= \max\{F_{\xi|\xi \in [z_{lo}, \infty)}(c), F_{\xi|\xi \in [z_{lo}, \infty)}(z_{lo} + \zeta(c))\},
\]
where the first inequality uses the fact that \( F_{\xi|\xi \in [z_{lo}, z_{up})}(t) \) is decreasing in \( z_{up} \) and the second inequality uses the fact that \( F_{\xi|\xi \in [z_{lo}, z_{up})}(t) \) is strictly increasing in \( t \) when \( t \in (z_{lo}, \infty) \), and that \( \max\{c, z_{lo} + \zeta(c)\} > z_{lo} \) since \( \zeta(c) > 0 \). The final equality again uses the fact that \( F_{\xi|\xi \in [z_{lo}, z_{up})}(t) \) is increasing in \( t \).

However, if \( z_{lo} < c - \zeta(c) \), then
\[ F_{\xi|\xi \in [z_{lo}, \infty)}(c) \geq F_{\xi|\xi \in [c-\zeta(c), \infty)}(c) = 1 - \alpha, \]
since we’ve shown that the expression on the left hand side is decreasing in \( z_{lo} \). On the other hand,
if \( z_{lo} \geq c - \zeta(c) \), then

\[
F_{\xi \mid \xi \in [z_{lo}, x)}(z_{lo} + \zeta(c)) \geq F_{\xi \mid \xi \in [c - \zeta(c), x)}(c) = 1 - \alpha,
\]

since we’ve shown that \( F_{\xi \mid \xi \in [z_{lo}, x)}(z_{lo} + \zeta) \) is increasing in \( z_{lo} \). We have thus shown that the max on the right-hand side of (54) is at least \( 1 - \alpha \), which gives the desired result. \( \square \)

Lemma E.16. For any \( t \in \mathbb{R} \), \( \int_{-\infty}^{t} \Phi(x) dx \) is finite. In particular,

\[
\int_{-\infty}^{t} \Phi(x) dx = t\Phi(t) + \phi(t).
\]

Proof. We have

\[
\int_{-\infty}^{t} \Phi(x) dx = \int_{-\infty}^{\infty} 1[ x \leq t ] \int_{-\infty}^{\infty} 1[ s \leq x ] \phi(s) ds dx
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1[ s \leq x \leq t ] \phi(s) ds dx
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{t} 1[ s \leq x \leq t ] dx \phi(s) ds
\]

\[
= \int_{-\infty}^{\infty} (t - s) 1[ s \leq t ] \phi(s) ds
\]

\[
= t \int_{-\infty}^{\infty} 1[ s \leq t ] \phi(s) ds - \int_{-\infty}^{\infty} s 1[ s \leq t ] \phi(s) ds
\]

\[
= t\Phi(t) - \Phi(t) \mathbb{E} [ \xi \mid \xi \leq t, \xi \sim \mathcal{N}(0, 1)]
\]

\[
= t\Phi(t) - \Phi(t) \frac{-\phi(t)}{\Phi(t)};
\]

where the last line uses the formula for the mean of a truncated normal distribution. Note that we exchange the order of integration via Fubini’s theorem, which is valid since the integrand is weakly positive everywhere and thus equal to its absolute value, and we’ve shown that the integral after switching the order is finite. \( \square \)

Lemma E.17. Let

\[
\rho_{\nu}^{\text{LF}}(\theta^{\text{ub}}_{\nu} + x) := \mathbb{E}(\sqrt{n} \delta_{A}, \sqrt{n} \tau_{A}, \Sigma^{*}) \left[ \psi_{\nu}^{\text{LF}}(\tilde{\beta}_{n}; A, \sqrt{n} d, \theta^{\text{ub}}_{\nu} + x, \Sigma^{*}) \right],
\]

where \( \theta^{\text{ub}}_{\nu} = \sup \mathcal{S}(\Delta_{n}, \sqrt{n} \delta_{A} + \sqrt{n} \tau_{A}) \) and \( \Delta_{n} = \{ \delta : A \delta \leq \sqrt{n} d \} \). Then there exists a function \( \rho_{\nu}^{\text{LF}}(x, A \Sigma^{*} A') \) such that \( 0 \leq 1 - \rho_{\nu}^{\text{LF}}(\theta^{\text{ub}}_{\nu} + x) \leq 1 - \rho_{\nu}^{\text{LF}}(x, A \Sigma^{*} A') \) for all \( x \) and \( n \), and \( \int_{0}^{\infty} 1 - \rho_{\nu}^{\text{LF}}(x, A \Sigma^{*} A') dx < \infty \).

Proof. By definition, \( \psi_{\nu}^{\text{LF}}(\tilde{\beta}_{n}; A, \sqrt{n} d, \theta^{\text{ub}}_{\nu} + x, \Sigma^{*}) = \psi_{\nu}^{\text{LF}}(\tilde{\gamma}_{n}(x), A \Sigma^{*} A) \), where \( \tilde{\gamma}_{n}(\theta^{\text{ub}}_{\nu} + x) = A \tilde{\beta}_{n} - \sqrt{n} d - \tilde{A}(-1)(\theta^{\text{ub}}_{\nu} + x) \). Note that when \( (\delta, \tau, \Sigma) = (\sqrt{n} \delta_{A}, \sqrt{n} \tau_{A}, \Sigma^{*}) \), \( \tilde{\gamma}_{n}(\theta^{\text{ub}}_{\nu} + x) \sim \mathcal{N}(\bar{\mu}_{n}, \bar{\Sigma}) \), where
\[ \tilde{\mu}_n := \mathbb{E}(\sqrt{n}\delta_{A^\tau}^*, \sqrt{n}\tau_{A^\tau}^* \Sigma^*) \left[ \dot{Y}_n(\theta_n + x) \right] = A\sqrt{n}\beta_A - \sqrt{n}d - \dot{A}_{(1)}(\theta_n + x), \]

and \( \hat{\Sigma} = A\Sigma^* A' \). From Lemma E.14, there exists a constant \( c(\Sigma^*, A) > 0 \) such that for all \( n \),

\[ c \cdot x \leq \eta^*(x; \sqrt{n}\delta_{A^\tau}, \sqrt{n}\tau_{A^\tau}, \Sigma^*) := \min_{\eta, \tilde{\tau}} \eta \]

s.t. \( A\sqrt{n}\beta_A - \sqrt{n}d - \dot{A}_{(1)}(\theta_n + x) - \dot{A}_{(1)}(\tilde{\tau}) \leq \eta \tilde{\tau} \)

for \( \tilde{\tau} \) the vector containing the square roots of the diagonal elements of \( \hat{\Sigma} \). Substituting the definition of \( \tilde{\mu}_n \) from above, this implies that

\[ c \cdot x \leq \min_{\eta, \tilde{\tau}} \eta \quad \text{s.t.} \quad \tilde{\mu}_n - \dot{A}_{(1)}(\tilde{\tau}) \leq \eta \tilde{\tau}. \]

Reformulating in terms of the dual linear program, we have \( c \cdot x \leq \max_{\gamma \in V(\hat{\Sigma})} \gamma^T \tilde{\mu}_n \), where \( V(\hat{\Sigma}) \) is the set of vertices of \( F = \{ \gamma : \gamma^T \dot{A}_{(-1)} = 0, \gamma^T \hat{\Sigma} = 1 \} \), and \( \hat{\Sigma} \) contains the square roots of the diagonal elements of \( \hat{\Sigma} \). The result then follows immediately from Lemma E.18.

**Lemma E.18.** Suppose \( \dot{Y}(x) \sim \mathcal{N}(\bar{\mu}(x), \hat{\Sigma}) \) for some function \( \bar{\mu}(x) \) such that \( \max_{\gamma \in V(\hat{\Sigma})} \gamma^T \bar{\mu}(x) \geq x \geq 0 \), where \( V(\hat{\Sigma}) \) is the set of vertices of the dual feasible set, \( F = \{ \gamma : \gamma^T \dot{A}_{(1)} = 0, \gamma^T \hat{\Sigma} = 1 \} \), and \( \hat{\Sigma} \) contains the square roots of the diagonal elements of \( \hat{\Sigma} \). Then there exists a function \( \rho_{LF}(x, \hat{\Sigma}) \), not depending on \( \bar{\mu}(x) \), such that \( \mathbb{E}\left[ \psi_{\bar{\mu}}^{LF}(\dot{Y}, \hat{\Sigma}; \alpha) \right] \geq \rho_{LF}(x, \hat{\Sigma}) \geq 0 \), and \( \int_0^\infty 1 - \rho_{LF}(x, \hat{\Sigma}) dx < \infty \).

**Proof.** Recall that \( \hat{\eta} \) can be written as the solution to the dual problem that restricts to vertices of the feasible set, \( \hat{\eta} = \max_{\gamma \in V(\hat{\Sigma})} \gamma^T \dot{\gamma} \). By assumption, there exists \( \dot{\gamma} \in V(\hat{\Sigma}) \) such that \( \dot{\gamma}^T \bar{\mu}(x) \geq x \). Since \( \dot{\gamma} \) is feasible in the dual problem for \( \hat{\eta} \), we have that \( \psi_{\bar{\mu}}^{LF} = 1 \) whenever \( \dot{\gamma}^T \dot{\gamma} \geq c_{LF}(\hat{\Sigma}) \). Further, observe that \( \dot{\gamma}^T \dot{\gamma} \sim \mathcal{N}(\dot{\gamma}^T \bar{\mu}(x), \dot{\gamma}^T \hat{\Sigma} \dot{\gamma}) \). Suppose first that \( \dot{\gamma}^T \hat{\Sigma} \dot{\gamma} > 0 \). Then, \( \dot{\gamma}^T \dot{\gamma} > c_{LF}^2(\hat{\Sigma}) \) with probability \( \Phi \left( \frac{\dot{\gamma}^T \bar{\mu}(x) - c_{LF}^2(\hat{\Sigma})}{\sqrt{\dot{\gamma}^T \hat{\Sigma} \dot{\gamma}}} \right) \). Since the set of vertices \( V(\hat{\Sigma}) \) is finite, \( \bar{\sigma}^2 := \max_{\gamma \in V(\hat{\Sigma})} \gamma^T \hat{\Sigma} \gamma \) is finite. Thus, when \( x > c_{LF}^2 \), we have

\[
\mathbb{P} \left( \dot{\gamma}^T \dot{\gamma} > c_{LF}^2(\hat{\Sigma}) \right) = \Phi \left( \frac{\dot{\gamma}^T \bar{\mu}(x) - c_{LF}^2(\hat{\Sigma})}{\sqrt{\dot{\gamma}^T \hat{\Sigma} \dot{\gamma}}} \right) \geq \Phi \left( \frac{x - c_{LF}^2(\hat{\Sigma})}{\bar{\sigma}} \right) \geq \Phi \left( \frac{x - c_{LF}^2(\hat{\Sigma})}{\bar{\sigma}} \right),
\]

where the first inequality uses the fact that \( \dot{\gamma}^T \bar{\mu}(x) \geq x \), the second uses \( x > c_{LF}^2 \) by assumption,
and both inequalities use the fact that the normal CDF is increasing in its argument. Likewise, if \( \tilde{\gamma}' \tilde{\Sigma} \gamma = 0 \) and \( x > c_k^{LF} \), then \( \tilde{\gamma}' \tilde{Y} = \tilde{\gamma}' \tilde{\mu}(x) > x \) with probability 1. We have thus shown that for \( x > c_k^{LF} \), \( \mathbb{E} \left[ \psi_k^{LF}(\tilde{Y}, \tilde{\Sigma}) \right] \geq \Phi \left( \frac{x - c_k^{LF}}{\tilde{\sigma}} \right) \).

Now, define \( \rho_{LF}(x, \tilde{\Sigma}) = 0 \) for \( x \leq c_k^{LF}(\tilde{\Sigma}) \) and \( \rho_{LF}(x) = \Phi \left( \frac{x - c_k^{LF}(\tilde{\Sigma})}{\tilde{\sigma}} \right) \) for \( x > c_k^{LF}(\tilde{\Sigma}) \). By construction, we have \( 0 \leq \mathbb{E} \left[ 1 - \psi_k^{LF}(\tilde{Y}, \tilde{\Sigma}) \right] \leq 1 - \rho_{LF}(x, \tilde{\Sigma}) \). Moreover, we have

\[
\int_0^\infty 1 - \rho_{LF}(x) dx = c_k^{LF} + \int_{c_k^{LF}}^\infty \Phi \left( \frac{c_k^{LF} - x}{\tilde{\sigma}} \right) dx
\]

\[
= c_k^{LF} + \tilde{\sigma} \int_{-\infty}^0 \Phi(t) dt,
\]

where the second line applies a change of variables to the integral in the first. The integral in the previous display is finite by Lemma E.16, which completes the proof.

**Lemma E.19.** Suppose \( \tilde{Y}(x) \sim \mathcal{N}(\tilde{\mu}(x), \tilde{\Sigma}) \) for some \( \tilde{\mu}(x) \) such that \( \max_{\gamma \in V(\tilde{\Sigma})} \gamma' \tilde{\mu}(x) \geq x > 0 \), where \( V(\tilde{\Sigma}) \) is the set of vertices of the dual feasible set, \( F = \{ \gamma : \gamma \geq 0, \gamma' \tilde{A}_{(-1)} = 0, \gamma' \tilde{\sigma} = 1 \} \), and \( \tilde{\sigma} \) contains the square root of the diagonal elements of \( \tilde{\Sigma} \). Then there exists a function \( \rho(x, \tilde{\Sigma}) \), not depending on \( \tilde{\mu}(x) \), such that \( \mathbb{E} \left[ \psi_{\alpha}^C(\tilde{Y}, \tilde{\Sigma}) \right] \geq \rho(x, \tilde{\Sigma}) \) and \( \lim_{n \to \infty} \rho(x, \tilde{\Sigma}) = 1 \).

**Proof.** Recall that \( \psi_{\alpha}^C \) is based on the solution to the dual problem, \( \hat{\eta} = \max_{\gamma \in V(\tilde{\Sigma})} \gamma' \tilde{Y} \). Specifically, \( \psi_{\alpha}^C(\tilde{Y}, \tilde{\Sigma}) = 1 \) iff

\[
F_{\xi | \xi \in [v^{lo}_{\gamma_*}, v^{up}_{\gamma_*}]}(\hat{\eta}; \gamma_*, \tilde{\Sigma})) > 1 - \alpha,
\]

where \( \gamma_* \) is an optimal solution to the dual \( (\gamma_* \in \hat{V}) \), and \( v^{lo}, v^{up} \) are functions of \( \gamma_* \), \( \tilde{\Sigma} \), and a sufficient statistic \( S_{\gamma_*}(\tilde{Y}) \) that by construction is independent of \( \gamma_*' \tilde{Y} \). (In this proof only, we make the dependence of \( v^{lo} \) and \( v^{up} \) on \( \gamma_* \) explicit.) If \( \gamma_*' \tilde{\Sigma} \gamma_* \neq 0 \), then using the standard formula for the CDF of a truncated normal distribution, we have that the conditional test rejects if

\[
\frac{\Phi(\hat{\eta}/\sigma_{\gamma_*}) - \Phi(z^{lo}_{\gamma_*}/\gamma_*)}{\Phi(z^{up}_{\gamma_*}/\gamma_*) - \Phi(z^{lo}_{\gamma_*}/\gamma_*)} > 1 - \alpha,
\]

where \( \sigma_{\gamma_*} = \sqrt{\gamma_*' \tilde{\Sigma} \gamma_*} \) and \( z^{lo}_{\gamma_*} = \hat{\eta}^{lo}_{\gamma_*}/\sigma_{\gamma_*}, z^{up}_{\gamma_*} = \hat{\eta}^{up}_{\gamma_*}/\sigma_{\gamma_*} \). By Lemma E.15, for any \( c > z_{1-\alpha} \), (55) holds whenever \( \hat{\eta}/\sigma_{\gamma_*} > \max\{c, z^{lo}_{\gamma_*} + \zeta(c)\} \), where \( \zeta(c) \) is the unique value that solves

\[
\frac{\Phi(c) - \Phi(c - \zeta(c))}{1 - \Phi(c - \zeta(c))} = 1 - \alpha.
\]

Thus, when \( \sigma_{\gamma_*} \neq 0 \), \( \psi_{\alpha} = 1 \) whenever \( \hat{\eta}/\sigma_{\gamma_*} > \max\{c, z^{lo}_{\gamma_*} + \zeta(c)\} \), or equivalently, whenever \( \hat{\eta} > \sigma_{\gamma_*} c \) and \( \hat{\eta}/\sigma_{\gamma_*} - z^{lo}_{\gamma_*} > \zeta(c) \). Additionally, if \( \sigma_{\gamma_*} = 0 \), then \( \psi_{\alpha} = 1 \) whenever \( \hat{\eta} > 0 \).
Let \( \bar{\sigma} = \max_{\gamma \in V(\tilde{\Sigma})} \sigma_\gamma \), which is finite since \( V(\tilde{\Sigma}) \) is finite. Then the preceding discussion implies that for any \( c > \max\{z_{1-\alpha}, 0\} \), \( \psi^{C}_\alpha = 1 \) whenever

1) \( \hat{\eta} > \bar{\sigma} c \), AND

2) \( \exists \gamma_\ast \in \hat{V} \) such that either i) \( \sigma_{\gamma_\ast} = 0 \), OR ii) \( \sigma_{\gamma_\ast} > 0 \) and \( \gamma_\ast \hat{Y}/\sigma_{\gamma_\ast} - z_{\gamma_\ast}^{lo} > \zeta(c) \),

where for the second part of condition 2) we use the fact that \( \hat{\eta} = \gamma_\ast \hat{Y} \) when \( \gamma_\ast \in \hat{V} \). Hence, \( \psi^{C}_\alpha = 0 \) only if either

A) \( \hat{\eta} \leq \bar{\sigma} c \), OR

B) \( \exists \gamma_\ast \in \hat{V} \) such that \( \sigma_{\gamma_\ast} > 0 \) and \( \gamma_\ast \hat{Y}/\sigma_{\gamma_\ast} - z_{\gamma_\ast}^{lo} \leq \zeta(c) \),

Now, by assumption there exists some \( \hat{\gamma} \in V(\tilde{\Sigma}) \) such that \( \hat{\gamma}' \hat{\mu}(x) \geq x \). Since \( \hat{\gamma} \) is feasible in the problem for \( \hat{\eta} \), we see that \( \hat{\eta} \) is lower bounded by \( \hat{\gamma}' \hat{Y} \), which is distributed \( \mathcal{N} (\hat{\gamma}' \hat{\mu}(x), \sigma_\gamma^2) \). Thus, the probability that condition A) holds is bounded above by the probability that \( \hat{\gamma}' \hat{Y} \leq \bar{\sigma} c \). If \( \sigma_\gamma = 0 \), then the probability condition A) holds is 0 so long as \( c < \bar{\sigma} \frac{\sigma}{\bar{\sigma}} \). If \( \sigma_\gamma > 0 \), then \( \Pr (\hat{\gamma}' \hat{Y} \leq \bar{\sigma} c) = \Phi \left( \frac{\bar{\sigma} c - \hat{\gamma}' \hat{\mu}(x)}{\sigma_\gamma} \right) \). If \( \sigma_\gamma > 0 \), then the set \( V^+(\tilde{\Sigma}) := \{ \gamma \in V(\tilde{\Sigma}) \colon \sigma_\gamma > 0 \} \) is non-empty. In this case, let \( \sigma = \min_{\gamma \in V^+} \sigma_\gamma \) and note that \( \sigma > 0 \) since \( V^+ \) is finite. Then

\[
\Phi \left( \frac{\bar{\sigma} c - \hat{\gamma}' \hat{\mu}(x)}{\sigma_\gamma} \right) \leq \Phi \left( \frac{\bar{\sigma} c - x}{\sigma} \right) \leq \Phi \left( \frac{\bar{\sigma} c - \hat{\gamma}' \hat{\mu}(x)}{\sigma} \right),
\]

where we use the fact that \( \Phi(\cdot) \) is increasing, \( c \geq 0 \) and \( \hat{\gamma}' \hat{\mu}(x) \geq x \). Thus, if \( c < \bar{\sigma} \frac{\sigma}{\bar{\sigma}} \), we have that condition A) holds with probability bounded above by \( \Phi \left( \frac{\bar{\sigma} c - \bar{\sigma} \frac{\sigma}{\bar{\sigma}}}{\bar{\sigma} \frac{\sigma}{\bar{\sigma}}} \right) \).

Now, the probability that condition B) holds is equal to

\[
\Pr (\exists \gamma_\ast \in \hat{V} \text{ s.t. } \sigma_{\gamma_\ast} > 0 \text{ and } \gamma_\ast \hat{Y}/\sigma_{\gamma_\ast} - z_{\gamma_\ast}^{lo} \leq \zeta(c)) = \\
\Pr (\exists \gamma_\ast \in \hat{V} \text{ s.t. } \sigma_{\gamma_\ast} > 0 \text{ and } |\gamma_\ast \hat{Y}/\sigma_{\gamma_\ast} - z_{\gamma_\ast}^{lo}| \leq \zeta(c)) \leq \\
\Pr (\exists \gamma_+ \in V^+ \text{ s.t. } |\gamma_+ \hat{Y}/\sigma_{\gamma+} - z_{\gamma+}^{lo}| \leq \zeta(c)) \leq \\
\sum_{\gamma_+ \in V^+} \Pr (|\gamma_+ \hat{Y}/\sigma_{\gamma+} - z_{\gamma+}^{lo}| \leq \zeta(c)).
\]

The equality above uses the fact that \( \gamma_\ast \in \hat{V} \) implies that \( \gamma_\ast \hat{Y}/\sigma_{\gamma_\ast} - z_{\gamma_\ast}^{lo} \geq 0 \) since \( \hat{\eta} \geq v^{lo} \) by construction; and the remaining inequalities follow from standard properties of probability. Next, observe that for \( \gamma_+ \in V^+(\tilde{\Sigma}) \), \( \gamma_+ \hat{Y}/\sigma_{\gamma+} \) is normally distributed with variance 1. Additionally, the random variable \( z_{\gamma_+}^{lo} \) is by construction independent of \( \gamma_+ \hat{Y}/\sigma_{\gamma+} \). However, for any variable \( \xi \) that is normally distributed with variance 1 and any variable \( Z \) independent of \( \xi \),

\[
\Pr(\xi, Z) (|\xi - Z| \leq \zeta) = \mathbb{E}_Z \left[ \Pr_{\xi|Z} (\xi \in [z - \zeta, z + \zeta] \mid Z = z) \right] \\
\leq \max_{v \in \mathbb{R}} \Pr_{\xi} (\xi \in [v - \zeta, v + \zeta]) = \Phi(\zeta) - \Phi(-\zeta),
\]

S-41
where the first equality follows from iterated expectations, the inequality uses the fact that the distribution of \( \xi \) is independent of \( Z \), and the final equality uses the fact that the normal distribution is single-peaked at its mean, so the maximal probability that a normal distribution with variance 1 falls in an interval of length 2\( \zeta \) is \( \Phi(\zeta) - \Phi(-\zeta) \). Additionally, observe that \( \Phi(\zeta) - \Phi(-\zeta) = \int_{-\zeta}^{\zeta} \phi(t) dt \leq 2\phi(0)\zeta \). It follows that for any constant \( c > \max\{z_{1-\alpha}, 0\} \), the probability condition B) holds is bounded above by \( \kappa \zeta(c) \), where we define the constant \( \kappa = 2|V^+|\phi(0) \).

Since \( \psi^c = 0 \) only if either condition A) or condition B) holds, the probability that \( \psi^c = 0 \) is bounded above by

\[
\Phi\left( \frac{\sigma}{\sigma} - x \right) + \kappa \zeta(c),
\]

for any \( c \in [\max\{z_{1-\alpha}, 0\}, \frac{x}{\sigma}] \). Let \( c(x) = c_0 \cdot x \) for \( c_0 = \frac{1}{2} \sigma \). Note that \( c(x) > \max\{z_{1-\alpha}, 0\} \) for \( x > \max\{z_{1-\alpha}/c_0, 0\} =: x_{\text{min}} \). Note also that \( c_0 = \frac{1}{2} \sigma \) if \( \frac{x}{\sigma} < \frac{1}{2} \), so \( c(x) < \frac{1}{2} x \). For \( x > x_{\text{min}} \), we then have that the probability \( \psi^c = 0 \) is bounded above by

\[
\Phi\left( -\frac{1}{2\sigma} x \right) + \kappa \zeta(c_0 x).
\]

Define \( \rho(x, \Sigma) = 1 - \Phi\left( -\frac{1}{2\sigma} x \right) - \kappa \zeta(c_0 x) \) for \( x > x_{\text{min}} \) and \( \rho(x, \Sigma) = 0 \) otherwise. By construction, \( \mathbb{E}\left[ \psi^c(\hat{Y}, \Sigma) \right] \geq \rho(x, \Sigma) \). Note that as \( x \to \infty \), \( \Phi\left( -\frac{1}{2\sigma} x \right) \to 0 \). To complete the proof that \( \rho \to 1 \), we show that \( \kappa \zeta(c) \to 0 \) as \( c \to \infty \). To show this, observe that for any \( \epsilon > 0 \), by L'Hopital's rule,

\[
\lim_{c \to \infty} \frac{\Phi(c) - \Phi(c - \epsilon)}{1 - \Phi(c - \epsilon)} = \lim_{c \to \infty} \frac{\phi(c) - \phi(c - \epsilon)}{-\phi(c - \epsilon)} = 1 - \lim_{c \to \infty} \frac{\phi(c)}{\phi(c - \epsilon)} = 1 - \lim_{c \to \infty} \exp\left( -\frac{1}{2}(2\epsilon - \epsilon^2) \right) = 1.
\]

Additionally, as shown in the proof to Lemma E.15, \( \frac{\Phi(c) - \Phi(c - \epsilon)}{1 - \Phi(c - \epsilon)} \) is increasing in \( \zeta \). It is then immediate that \( \limsup_{c \to \infty} \zeta(c) < \epsilon \) for all \( \epsilon > 0 \), and hence \( \lim_{c \to \infty} \zeta(c) = 0 \).

\[\square\]

**Lemma E.20.** Suppose \( \hat{Y}(x) \sim \mathcal{N}\left( \tilde{\mu}(x), \tilde{\Sigma} \right) \) for some \( \tilde{\mu}(x) \) such that \( \max_{\gamma \in V(\tilde{\Sigma})} \gamma' \tilde{\mu}(x) \geq x > 0 \), where \( V(\tilde{\Sigma}) \) is the set of vertices of the dual feasible set, \( F = \{ \gamma : \gamma \geq 0, \gamma' \tilde{\Lambda}_{(-1,-1)} = 0, \gamma' \tilde{\sigma} = 1 \} \), and \( \tilde{\sigma} \) contains the square root of the diagonal elements of \( \tilde{\Sigma} \). Then there exists a function \( \rho(x, \tilde{\Sigma}) \), not depending on \( \tilde{\mu}(x) \), such that \( \mathbb{E}\left[ \psi^c_{-FLCI}(\hat{Y}, \tilde{\Sigma}) \right] \geq \rho(x, \tilde{\Sigma}) \) and \( \lim_{n \to \infty} \rho(x, \tilde{\Sigma}) = 1 \).

**Proof.** The proof is nearly identical to that of Lemma E.19. In particular, by analogous argument we can show that the test \( \psi^c_{-FLCI} = 0 \) only if

S-42
A) \( \hat{\eta} \leq \bar{\sigma}c_{\bar{a}} \), OR

B) \( \exists \gamma_* \in \hat{V} \) such that \( \sigma_{\gamma_*} > 0 \) and \( 0 \leq \gamma'_* \hat{\gamma}/\sigma_{\gamma_*} - z_{C-\text{FLCI},\gamma_*}^0 \leq \zeta(c_{\bar{a}}) \),

where \( z_{C-\text{FLCI},\gamma_*}^0 = v_{C-\text{FLCI},\gamma_*}^0/\sigma_{\gamma_*} \), and \( c_{\bar{a}} \) solves

\[
\frac{\Phi(c_{\bar{a}}) - \Phi(c_{\bar{a}} - \zeta(c_{\bar{a}}))}{1 - \Phi(c_{\bar{a}} - \zeta(c_{\bar{a}}))} = 1 - \tilde{\alpha}.
\]

Noting that \( z_{C-\text{FLCI},\gamma_*}^0 \) is independent of \( \gamma'_* \hat{\gamma} \), we can then obtain upper bounds on the probability that conditions A) or B) hold by an analogous argument to that in the proof to Lemma E.19.

\( \square \)

Lemma E.21. Let \( \Delta = \{ \delta : A\delta \leq d \} \). Then there exists a function \( \rho_{LB}(\cdot, \cdot) \) such that for any \( \delta \in \Delta, \tau, \Sigma^*, \) and \( x > 0 \),

\[
\mathbb{E}_{(\delta, \tau, \Sigma^*)} \left[ \psi^C_{\kappa, \alpha}(\hat{\beta}_n; A, d, \theta^{ub}(\Delta, \delta + \tau) + x, \Sigma^*) \right] \geq \rho_{LB}(x, \Sigma^*),
\]

and for any \( \Sigma^* \) fixed, \( \rho_{LB}(x, \Sigma^*) \rightarrow 1 \) as \( x \rightarrow \infty \). Analogously, there exists a function \( \tilde{\rho}_{LB}(\cdot, \cdot) \) such that for any \( \delta \in \Delta, \tau, \Sigma^*, \) and \( x > 0 \),

\[
\mathbb{E}_{(\delta, \tau, \Sigma^*)} \left[ \psi^{C-\text{FLCI}}_{\kappa, \alpha}(\hat{\beta}_n; A, d, \theta^{ub}(\Delta, \delta + \tau) + x, \Sigma^*) \right] \geq \tilde{\rho}_{LB}(x, \Sigma^*),
\]

and for any \( \Sigma^* \) fixed, \( \tilde{\rho}_{LB}(x, \Sigma^*) \rightarrow 1 \) as \( x \rightarrow \infty \).

Proof. Lemma E.14 implies that there exists a scalar \( c(\Sigma^*, A) > 0 \) such that

\[
c(\Sigma^*, A) \cdot x \leq \eta^*(x; \delta, \tau, \Sigma^*) := \min_{\eta, \tilde{\eta}} \eta \text{ s.t. } A\beta - d - \tilde{A}_{(\cdot, -1)}(\theta^{ub} + x) - \tilde{A}_{(\cdot, -1)}\tilde{\eta} \leq \eta\tilde{\eta},
\]

where \( \beta = \delta + \tau \). Reformulating the minimization above in terms of its dual, we have that \( c(\Sigma^*, A) \cdot x \leq \max_{\gamma \in V(\hat{\Sigma}^*)} \gamma'\mu \), where \( V(\hat{\Sigma}^*) \) is the set of vertices of \( F = \{ \gamma'\hat{A}_{(\cdot, -1)} = 0, \gamma'\tilde{\sigma} = 0, \gamma \geq 0 \} \).

Next, recall that by definition, \( \psi^C_{\alpha}(\hat{\beta}; A, d, \hat{\bar{\theta}}, \Sigma^*) = \psi^C_{\alpha}(\hat{Y}(\hat{\beta}, A, d, \hat{\bar{\theta}}), A\Sigma^* A') \), where \( \hat{Y}(\hat{\beta}, A, d, \hat{\bar{\theta}}) = A\beta - d - \tilde{A}_{(\cdot, -1)}\hat{\theta} \).

Observe that \( \mathbb{E}_{(\delta, \tau, \Sigma^*)} \left[ \hat{Y}(\hat{\beta}, A, d, \hat{\bar{\theta}}) \right] = \bar{\mu} \).

Lemma E.19 then implies that there exists a function \( \rho(\cdot, \cdot) \) such that

\[
\mathbb{E}_{(\delta, \tau, \Sigma^*)} \left[ \psi^C_{\alpha}(\hat{\beta}_n; A, d, \theta^{ub}(\Delta, \delta + \tau) + x, \Sigma^*) \right] \geq \rho(c(\Sigma^*, A) \cdot x, A\Sigma^* A'),
\]

and \( \rho(\hat{x}, A\Sigma^* A') \rightarrow 1 \) as \( \hat{x} \rightarrow \infty \). The first desired result then follows by defining \( \rho_{LB}(x, \Sigma^*) := \rho(c(\Sigma^*, A) \cdot x, A\Sigma^* A'). \) The second desired result follows from an analogous argument, appealing to Lemma E.20 instead of Lemma E.19.

\( \square \)

Lemma E.22. For any \( (a, v) \), \( \bar{b}(a, v) \geq \frac{1}{2} \sup_{\delta_{pre} \in \Delta_{pre}} LID(\Delta, \delta_{pre}). \)
Proof. Since $\beta = \delta + \tau$, we can write the bias of the affine estimator $a + v'\hat{\beta}$ as $b = a + v'\delta + (v_{post} - l)\tau_{post}$. Since $\tau_{post}$ is unrestricted in the maximization in (22), we see that the worst-case bias will be infinite if $v_{post} \neq l$ and the lemma holds trivially. We can thus restrict attention to affine estimators with $v_{post} = l$, in which case the worst-case bias reduces to

$$\bar{b}(a, v) = \sup_{\delta \in \Delta} |a + v'\delta| = \sup_{\delta \in \Delta} |a + v'_{pre}\delta_{pre} + l'\delta_{post}|.$$ \hfill (56)

Now, pick any $\delta^*_{pre} \in \Delta_{pre}$. First, suppose that the minimum $\left(\min_{\delta} l'\delta_{post}, \text{ s.t. } \delta \in \Delta, \delta_{pre} = \delta^*_{pre}\right)$ and the equivalent maximum $\left(\max_{\delta} l'\delta_{post}, \text{ s.t. } \delta \in \Delta, \delta_{pre} = \delta^*_{pre}\right)$ are finite. Let $\delta_{min}^*_{pre}$ and $\delta_{max}^*_{pre}$ be the associated solutions. By construction, $\delta_{max}^*_{pre} = \delta_{pre}^* = \delta_{pre}$. For any $v_{pre}$, we apply the triangle inequality to show that

$$|a + v'_{pre}\delta_{max}^*_{pre} + l'\delta_{post}^*| = |a + v'_{pre}\delta_{min}^*_{pre} + l'\delta_{post}^*| = |l'\delta_{max}^*_{post} - l'\delta_{min}^*_{post}| = LID(\Delta, \delta^*_{pre}).$$

Note that for any $x_1, x_2 \geq 0$, $\max\{x_1, x_2\} \geq \frac{1}{2}(x_1 + x_2)$. It then follows from the previous display that

$$\max\{|a + v'_{pre}\delta_{max}^*_{pre} + l'\delta_{post}^*|, |a + v'_{pre}\delta_{min}^*_{pre} + l'\delta_{post}^*|\} \geq \frac{1}{2}LID(\Delta, \delta^*_{pre}).$$

Since $\delta_{max}^*_{pre}$ and $\delta_{min}^*_{pre}$ are feasible in the maximization (56), we see that $\bar{b} \geq \frac{1}{2}LID(\Delta, \delta^*_{pre})$, as needed. To complete the proof, now suppose without loss of generality that

$$\left(\max_{\delta} l'\delta_{post}, \text{ s.t. } \delta \in \Delta, \delta_{pre} = \delta^*_{pre}\right) = \infty.$$

Then, we can replay the argument above replacing $\delta_{max}^*_{pre}$ with a sequence of values $\{\delta_j\}$ such that $l'\delta_j$ diverges, which gives that $\bar{b}$ is infinite and the result follows. \hfill $\blacksquare$

**Lemma E.23.** Suppose $\Delta$ is convex. Suppose there exists $\delta \in \Delta$ such that $LID(\Delta, \delta_{pre}) = \sup_{\delta_{pre} \in \Delta_{pre}} LID(\Delta, \delta_{pre}) < \infty$. Then there exists $(a, v)$ such that $\bar{b}(a, v) = \frac{1}{2}\sup_{\delta_{pre} \in \Delta_{pre}} LID(\Delta, \delta_{pre})$. Additionally, for any $\tau$ and $\Sigma_n$, $\mathbb{E}(\delta, \tau, \Sigma_n)[a + v'\hat{\beta}_n] = \frac{1}{2}(\theta^u b + \theta^l b)$, where $\theta^u b$ and $\theta^l b$ are the upper and lower bounds of the identified set $S(\Delta, \delta + \tau)$.

**Proof.** Let $b_{max}^*(\delta_{pre}) := \left(\max_{\delta} l'\delta_{post}, \text{ s.t. } \delta \in \Delta, \delta_{pre} = \delta_{pre}^*\right)$, where we define $b_{max}^* = -\infty$ if $\delta_{pre}^* \notin \Delta_{pre}$. Likewise, define $b_{min}^*(\delta_{pre}) := \left(\min_{\delta} l'\delta_{post}, \text{ s.t. } \delta \in \Delta, \delta_{pre} = \delta_{pre}^*\right)$, where we define $b_{min}^* = \infty$ if $\delta_{pre}^* \notin \Delta_{pre}$. Note that $\Delta$ convex implies that $b_{max}^*$ is concave and $b_{min}^*$ is convex. Thus, $-LID(\delta_{pre}) = b_{min}^*(\delta_{pre}) - b_{max}^*(\delta_{pre})$ is convex (where we define $LID(\delta_{pre}) = -\infty$ if $\delta_{pre} \notin \Delta_{pre}$). The domain of $-LID(\delta_{pre})$ (i.e. the set of values for which it is finite) is $\Delta_{pre}$, since it is infinite for $\delta_{pre} \notin \Delta_{pre}$ by construction, and by assumption, $LID(\delta_{pre})$ is finite for all $\delta_{pre} \in \Delta_{pre}$. Since $\Delta$ is assumed to be convex, it is easy to verify that $\Delta_{pre}$ is a non-empty convex set, and thus
has non-empty relative interior, so the relative interior of the domain of \(-LID\) is non-empty.\(^{43}\) It follows from standard results in convex analysis (see Theorem 8.2 in Mau Nam (2019)) that \(\partial(-LID) = \partial(-b^{\text{max}}) + \partial(b^{\text{min}})\) where for a convex function \(f\), \(\partial f\) is the subdifferential

\[
\partial f(\bar{x}) := \{ v : f(\bar{x}) + v'(x - \bar{x}) \leq f(x), \forall x \}
\]

and \(\partial(-b^{\text{max}}) + \partial(b^{\text{min}})\) is the Minkowski sum of the two subdifferentials.

Additionally, if \(LID(\delta_{\text{pre}}) = \sup_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} LID(\delta_{\text{pre}})\), then \(-LID(\delta_{\text{pre}}) = \inf_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} -LID(\delta_{\text{pre}})\). Thus, standard results in convex analysis (see, e.g., Theorem 16.2 in Mau Nam (2019)) give that

\[
b^{\text{min}}(\delta_{\text{pre}}) + \bar{v}^{\text{min}}(\delta_{\text{pre}} - \delta_{\text{pre}}) \leq b^{\text{min}}(\delta_{\text{pre}})
\]

and

\[
b^{\text{max}}(\delta_{\text{pre}}) + \bar{v}^{\text{max}}(\delta_{\text{pre}} - \delta_{\text{pre}}) \leq b^{\text{max}}(\delta_{\text{pre}})
\]

The inequalities (58) and (59) together imply that for all \(\tilde{\delta}_{\text{pre}} \in \Delta_{\text{pre}},\)

\[
b^{\text{max}}(\delta_{\text{pre}}) + \bar{v}^{\text{min}}(\delta_{\text{pre}} - \delta_{\text{pre}}) \leq b^{\text{max}}(\delta_{\text{pre}}).
\]

Now, let \(v\) be the vector such that \(v_{\text{post}} = l\) and \(v_{\text{pre}} = -\bar{v}^{\text{min}}\). Observe that

\[
\max_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \tilde{\delta}_{\text{pre}} + l' \tilde{\delta}_{\text{post}} = \max_{\tilde{\delta}_{\text{pre}} \in \Delta_{\text{pre}}} \left( a + v'_{\text{pre}} \tilde{\delta}_{\text{pre}} + b^{\text{max}}(\tilde{\delta}_{\text{pre}}) \right)
\]

where the first equality nests the maximization, the second equality uses the definition of \(b^{\text{max}}\), and the inequality follows from (60). An analogous argument using (57) yields that

\[
\min_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \tilde{\delta}_{\text{pre}} + l' \tilde{\delta}_{\text{post}} = \min_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \tilde{\delta}_{\text{pre}} + b^{\text{min}}(\delta_{\text{pre}})
\]

Now, it is apparent from the formulation of the identified set in (56) that

\[
\bar{b}(a, v) = \max \left\{ \left| \max_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \tilde{\delta}_{\text{pre}} + l' \tilde{\delta}_{\text{post}} \right|, \left| \min_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \tilde{\delta}_{\text{pre}} + l' \tilde{\delta}_{\text{post}} \right| \right\},
\]

\(^{43}\)The relative interior of a set is the interior of the set relative to its affine hull. See, e.g., Mau Nam (2019), Chapter 5.
which from the results above is bounded above by

$$
\max \left\{ |a + v'_p \delta_{pre} + b^{\max}(\delta_{pre})|, |a + v'_p \delta_{pre} + b^{\min}(\delta_{pre})| \right\}.
$$

Setting $a = -v'_p \delta_{pre} - \frac{1}{2} (b^{\max}(\delta_{pre}) + b^{\min}(\delta_{pre}))$, the upper bound in the previous display reduces to $\frac{1}{2} (b^{\max}(\delta_{pre}) - b^{\min}(\delta_{pre}))$. Since $\text{LID}(\Delta, \delta_{pre}) = b^{\max}(\delta_{pre}) - b^{\min}(\delta_{pre})$ and $\text{LID}(\Delta, \delta_{pre}) = \sup_{\delta_{pre} \in \Delta_{pre}} \text{LID}(\Delta, \delta_{pre})$ by assumption, it is then immediate that $\bar{b} \leq \frac{1}{2} \sup_{\delta_{pre} \in \Delta_{pre}} \text{LID}(\Delta, \delta_{pre})$.

The inequality in the opposite direction follows from Lemma E.22.

Finally, substituting in the definition of $a$ and $v$ above and simplifying, we see that

$$
\mathbb{E}_{(\delta, \tau)} \left[ a + v' \beta_n \right] = l' \beta_{post} - \frac{1}{2} (b^{\max}(\delta_{pre}) + b^{\min}(\delta_{pre})),
$$

which from (9) and (10) we see is the midpoint of the identified set. \hfill \Box

**Lemma E.24.** Let $\chi_\alpha$ be the $1 - \alpha$ quantile of the $|N(b, \sigma^2)|$ distribution. Then $b + \sigma z_{1-\alpha} \leq \chi_\alpha \leq b + \sigma z_{1-\alpha}/2$.

**Proof.** Since $|\xi| \leq |\xi|$, we have that $q_{1-\alpha}(|\xi| | \xi \sim N(b, \sigma^2)) \geq q_{1-\alpha}(|\xi| | \xi \sim N(0, \sigma^2)) = b + \sigma z_{1-\alpha}$, which yields the first inequality. For the second inequality, observe that

$$
q_{1-\alpha}(|\xi| | \xi \sim N(b, \sigma^2)) = q_{1-\alpha}(|\xi + b| | \xi \sim N(0, \sigma^2)) \leq b + q_{1-\alpha}(|\xi| | \xi \sim N(0, \sigma^2)) = b + \sigma z_{1-\alpha}/2
$$

where the first inequality uses the triangle inequality, and the final equality uses the fact that a mean-zero normal distribution is symmetric about 0. \hfill \Box

**Lemma E.25.** Suppose the conditions of Proposition 6.2 holds. Then there is a unique pair $(\bar{a}, \bar{v})$ such that $\bar{b}(\bar{a}, \bar{v}) = \frac{1}{2} \sup_{\delta_{pre} \in \Delta_{pre}} \text{LID}(\Delta, \delta_{pre}) =: \hat{b}_{\min}$. Additionally, $\sqrt{\bar{v}A\Sigma^*A^T} \bar{v} = 1/c^*$, for $c^*$ defined in Proposition 5.2.

**Proof.** Let $b^{\min}(\delta_{pre,A}) := (\min_\delta l' \delta_{post}, \text{ s.t. } \delta \in \Delta, \delta_{pre} = \delta_{pre,A})$, and define $b^{\max}$ equivalently. In the proof to Lemma E.8, we showed that $b^{\min}$ is equivalent to the problem (48). We also showed that Assumption 3 implies that there is a solution $\delta^{**}$ such $A_{(B, post)}$ has rank $|B|$, where $|B|$ indexes the binding moments. The solution $\delta^{**}$ is thus non-degenerate. It follows that in a neighborhood of $\delta_{pre,A}$, $b^{\min}(\delta_{pre}) = b^{\min}(\delta_{A,pre}) + \gamma' A_{(\cdot, pre)}(\delta_{pre} - \delta_{A,pre})$, where $\gamma$ is a solution to the dual problem (see, e.g., Section 10.4 of Schrijver (1986)). By the complementary slackness conditions, $\gamma_{\cdot B} = 0$. Moreover, we showed in the proof to Lemma E.8 that $\gamma_{B}$ is the unique vector that satisfies $\gamma'_B \hat{A}_{(B, -)} = 0; \gamma_B \hat{A}_{(B, 1)} = 1$. By an analogous argument, we can show that $b^{\max}(\delta_{pre})$ is also linear in a neighborhood of $\delta_{A,pre}$. Moreover, the local slope (gradient) for $b^{\max}$ must be the same as for $b^{\min}$, since $\text{LID}(\Delta, \delta_{pre}) = b^{\max}(\delta_{pre}) - b^{\min}(\delta_{pre})$ and $\text{LID}$ is maximized at $\delta_{pre,A}$ and thus $\text{LID}$ must have 0 derivative at $\delta_{A,pre}$.
Next, combining the expression for $\tilde{b}$ in (63) along with the equalities in (61) and (62) in the proof to Lemma E.23, we see that for any $(a,v)$,

$$
\tilde{b}(a, v) = \max \left\{ \max_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \delta_{\text{pre}} + b^{\max}(\delta_{\text{pre}}), \min_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} a + v'_{\text{pre}} \delta_{\text{pre}} + b^{\min}(\delta_{\text{pre}}) \right\}.
$$

(64)

This implies that if $(\tilde{a}, \tilde{v})$ are such that $\tilde{b}(\tilde{a}, \tilde{v}) = \tilde{b}_{\text{min}}$, then for all $\delta_{\text{pre}} \in \Delta_{\text{pre}},$

$$
\tilde{b}_{\text{min}} \geq \max \left\{ |\tilde{a} + \tilde{v}'_{\text{pre}} \delta_{\text{pre}} + b^{\max}(\delta_{\text{pre}})|, |\tilde{a} + \tilde{v}'_{\text{pre}} \delta_{\text{pre}} + b^{\min}(\delta_{\text{pre}})| \right\}.
$$

(65)

Now, note that by the triangle inequality, for any scalars $x_1, x_2, x_3$ with $x_2 \geq x_3$, max$\{|x_1 + x_2|, |x_1 + x_3|\} \geq \frac{1}{2} |x_2 - x_3|$, with equality if and only if $x_1 + x_3 = -(x_1 + x_2)$. Further, recall that $b^{\max}(\delta_{A, \text{pre}}) - b^{\min}(\delta_{A, \text{pre}}) = LID(\Delta, \delta_{A, \text{pre}}) = 2\tilde{b}_{\text{min}}$. It follows from these two facts along with the expression in the previous display that

$$
\tilde{b}_{\text{min}} = \tilde{a} + \tilde{v}'_{\text{pre}} \delta_{A, \text{pre}} + b^{\max}(\delta_{A, \text{pre}}) = -\left( \tilde{a} + \tilde{v}'_{\text{pre}} \delta_{A, \text{pre}} + b^{\min}(\delta_{A, \text{pre}}) \right).
$$

(66)

Displays (65) and (66) together imply that

$$
\nabla_{\delta_{\text{pre}} = \delta_{A, \text{pre}}} (\tilde{a} + \tilde{v}'_{\text{pre}} \delta_{\text{pre}} + b^{\max}(\delta_{\text{pre}})) \leq 0,
$$

where $\nabla_{\delta_{\text{pre}} = \delta_{A, \text{pre}}} f$ represents the gradient of the function $f(\delta_{\text{pre}})$ evaluated at $\delta_{\text{pre}} = \delta_{A, \text{pre}}$, since otherwise there is a local perturbation to $\delta_{A, \text{pre}}$ that violates (65). Using the local linearity derived early, the previous display implies that $\tilde{v}'_{\text{pre}} + \tilde{A}'_{\text{pre}} \leq 0$. By an analogous argument, (65) and (66) imply that

$$
\nabla_{\delta_{\text{pre}} = \delta_{A, \text{pre}}} (\tilde{a} + \tilde{v}'_{\text{pre}} \delta_{\text{pre}} + b^{\min}(\delta_{\text{pre}})) \geq 0,
$$

which implies that $\tilde{v}'_{\text{pre}} + \tilde{A}'_{\text{pre}} \geq 0$. Hence, we see that $\tilde{v}'_{\text{pre}} = -\tilde{A}'_{\text{pre}}$. We argued in the proof to Lemma E.22 that $\tilde{v}_{\text{post}}$ must equal $l$, so we have shown that there is a unique value of $\tilde{v}$. Further, (65) uniquely pins downs $\tilde{a}$ in terms of $\tilde{v}$, and so the pair $(\tilde{a}, \tilde{v})$ is unique, as claimed.

Finally, recall from the proof to Lemma E.8 that $-\gamma'_{\text{pre}} A_{\text{(pre)}} = l'$. Hence $\tilde{v}' = (-\gamma'_{\text{pre}} A_{\text{(pre)}}, -\gamma'_{\text{post}} A_{\text{(post)}}) = -\gamma' A$ and thus $\tilde{v}' \Sigma^* \tilde{v} = \gamma' A \Sigma^* A' \tilde{v}$. Since $\gamma_{\text{-B}} = 0$ and $\gamma_{\text{B}} A_{\text{(B,1)}} = 1$, we see that $1/\sqrt{\tilde{v}' \Sigma^* \tilde{v}}$ corresponds with the formula for $e^*$ given in the footnote to Proposition 5.2.

\[\square\]

**Lemma E.26.** Suppose the conditions of Proposition 6.2 hold. Then $\frac{\sigma_{v_n, n}}{\sigma_{\tilde{v}, n}} \to 1$, where the optimal FLCI is based on the affine estimator $a_n + v'_n \delta_{\text{pre}}$ and $\tilde{v}$ is the unique value such that $b(\tilde{a}, \tilde{v}) = \tilde{b}_{\text{min}}$.

**Proof.** It suffices to show that $\sqrt{n} \sigma_{v_n, n} / \sqrt{n} \sigma_{\tilde{v}, n} \to 1$. Note that $\sqrt{n} \sigma_{v_n, n} = \sigma_{\tilde{v}, 1} = \sqrt{\tilde{v}' \Sigma^* \tilde{v}}$. By assumption, $\Sigma^*$ is positive definite, and we showed in the proof to Lemma E.22 that $\tilde{v}_{\text{post}} = l$, so $\tilde{v} \neq 0$. Hence $\sigma_{\tilde{v}, 1} > 0$. Next, observe that $\sqrt{n} \sigma_{v_n, n} = \sqrt{\tilde{v}' \Sigma^* v_n}$. It thus suffices to show that $v_n \to \tilde{v}$, since
then both the numerator and denominator converge to the same non-zero limit. To do this, we will show that every subsequence of $v_n$ has a convergent subsequence. Consider a subsequence $v_{n_m}$. We argued in the proof to Proposition 6.1 that $\sigma_{v_{n_m}} \leq \sigma_{E,n}$, which implies that $\sqrt{v_n' \Sigma^* v_n} \leq \sqrt{v(v' \Sigma^* v)}$. Thus, $v_n$ is bounded in the Mahalanobis norm using $\Sigma^*$, which implies that $v_n$ is bounded in the standard euclidean norm since $\Sigma^*$ positive definite. It follows that $v_{n_m}$ has a convergent subsequence, $v_{n_m,1} \to v^\ast$. We argued in the proof to Proposition 6.1 that $\tilde{b}(a_n, v_n) \to \tilde{b}_{\min}$. This implies, however, that $a_{n_m,1}$ is bounded. To see why this is the case, note that if there is a divergent subsequence $a_{n_m,2}$, then for any $\tilde{\delta}_{\pre} \in \Delta_{\pre}$, $|a_{n_m,2} + v_{n_m,2,\pre} \tilde{\delta}_{\pre} + b_{\max}(\tilde{\delta}_{\pre})|$ diverges since $v_{n_m,2,\pre} \to v_{\ast,\pre}$. Equation (64) then implies that $\tilde{b}(a_{n_m,2}, v_{n_m,2})$ diverges, which is a contradiction. Thus $a_{n_m,1}$ is bounded, and so we can extract a further subsequence such that $(a_{n_m,2}, v_{n_m,2}) \to (a^\ast, v^\ast)$. To complete the proof, we argue that $\tilde{b}(a, v)$ is continuous in $(a, v)$ on the set of values $(a, v)$ such that $\tilde{b}(a, v)$ is finite. This suffices for the desired result, since the continuity of $\tilde{b}$ along with the convergences shown above imply that $\tilde{b}(a^\ast, v^\ast) = \tilde{b}_{\min}$, and Lemma E.25 then implies that $v^\ast = \hat{v}$.

To show the continuity of $\tilde{b}(a, v)$ on the set where it is finite, note that (56) implies that $\tilde{b}(a, v) = \max\{\max_{\delta} a + v' \delta \text{ s.t. } A \delta \leq d\}, (\min_{\delta} -a - v' \delta \text{ s.t. } A \delta \leq d\}$). However, the results in Section 10.4 of Schrijver (1986) imply that these functions are continuous in $v$ for values of $v$ for which they are finite.

\begin{lemma}
\lim_{x \to \infty} (cv_\alpha(x) - (z_{1-\alpha} + x)) = 0.
\end{lemma}

\begin{proof}
$cv_\alpha(x)$ solves
\begin{equation}
\Phi(cv_\alpha(x)) - x = \Phi(-cv_\alpha(x)) - x = 1 - \alpha.
\end{equation}
By Lemma E.24, $cv_\alpha(x) \geq x + z_{1-\alpha}$, which diverges as $x \to \infty$. Thus, the second normal CDF term in the previous display converges to 0. Hence, $\Phi(cv_\alpha(x) - x) \to 1 - \alpha$, and thus $cv_\alpha(x) - x \to z_{1-\alpha}$, which gives the desired result.
\end{proof}

\begin{lemma}
Suppose that Assumption 3 holds at $\delta_{A,\pre}$. Then LID($\Delta, \delta_{A,\pre}) > 0$.
\end{lemma}

\begin{proof}
From (9) and (10), we see that that LID($\Delta, \delta_{A,\pre} = 0$ if and only if $b_{\max}(\delta_{A,\pre}) = b_{\min}(\delta_{A,\pre})$, where $b_{\min}(\delta_{A,\pre}) := (\min_\delta l' \delta_{\post} \text{ s.t. } \delta \in \Delta, \delta_{\pre} = \delta_{\pre,\pre}, A)$, and $b_{\max}$ is defined analogously. In the proof to Lemma E.8, we showed that $b_{\min}$ is equivalent to the problem (48). Assumption 3 implies that there is a solution $\delta_{\ast,\post}$ such that
\begin{align*}
A_{(B,\post)} \delta_{\ast,\post} &= d_B - A_{(B,\pre)} \delta_{A,\pre} \\
A_{(-B,\post)} \delta_{\ast,\post} &= d_{-B} - A_{(-B,\pre)} \delta_{A,\pre},
\end{align*}
where $A_{B,\post}$ has rank $|B|$. Observe that if $b_{\min}(\delta_{A,\pre}) = b_{\max}(\delta_{A,\pre})$, then it must be that $l' \delta_{\pre} = l' \tilde{\delta}_{\pre}$ for any $\delta_{\pre}$ that is feasible in the problem (48). It thus suffices to construct a
feasible value \( \tilde{\delta}_{\text{pre}} \) such that \( l'\tilde{\delta}_{\text{pre}} \neq l'^*\delta_{\text{pre}}^* \). Since \( A_{(B,\text{post})} \) has rank \(|B|\), its image is \( \mathbb{R}^{|B|} \), so there exists \( \tilde{\delta}_{\text{post}} \) such that \( A_{(B,\text{post})}\tilde{\delta}_{\text{post}} = -t \), for \( t \) the vector of ones. Thus, for any \( \epsilon_1 > 0 \), we have that \( A_{(B,\text{post})}(\delta_{\text{pre}}^* + \epsilon_1 \tilde{\delta}_{\text{post}}) < d_B - A_{(B,\text{pre})}\delta_{\text{pre}} \). However, since the moments \(-B\) are slack at \( \delta_{A,\text{pre}} \), for \( \epsilon_1 \) sufficiently small, we also have \( A_{(-B,\text{pre})}(\delta_{\text{post}}^* + \epsilon_1 \tilde{\delta}_{\text{post}}) < d_{-B} - A_{(-B,\text{pre})}\delta_{A,\text{pre}} \). If \( l'\tilde{\delta}_{\text{post}} \neq 0 \), then we are done. If \( l'\tilde{\delta}_{\text{post}} = 0 \), then since all of the moments are slack at \( \delta_{\text{post}}^* + \epsilon_1 \tilde{\delta}_{\text{post}} \), for \( \epsilon_2 > 0 \) sufficiently small, \( \tilde{\delta}_{\text{post}} = \delta_{\text{post}}^* + \epsilon_1 \tilde{\delta}_{\text{post}} + \epsilon_2 l \) is also feasible, and by construction \( l'\tilde{\delta}_{\text{post}} - \delta_{\text{post}}^* = \epsilon_2 l'l > 0 \).

**Lemma E.29.** Suppose \( \Delta \) is convex and centrosymmetric, and \( \delta_A \) is such that \( \delta \in \Delta \) implies \( \delta - \delta_A \in \Delta \). Then \( \delta_A \) satisfies Assumption 4.

**Proof.** Recall from the proof to Lemma E.23 that for any \( \delta_{\text{pre}}^* \in \Delta_{\text{pre}} \), \( \text{LID}(\Delta, \delta_{\text{pre}}^*) = b_{\text{max}}(\delta_{\text{pre}}^*) - b_{\text{min}}(\delta_{\text{pre}}^*) \), where the functions \( b_{\text{min}} \) and \( -b_{\text{max}} \) are convex. Observe that

\[
b_{\text{min}}(\delta_{\text{pre}}^*) = \left( \min_{\delta} l'\delta_{\text{post}}, \text{ s.t. } \delta \in \Delta, \delta_{\text{pre}} = \delta_{\text{pre}}^* \right) = -\left( \max_{\delta} l'(-\delta_{\text{post}}), \text{ s.t. } \delta \in \Delta, \delta_{\text{pre}} = \delta_{\text{pre}}^* \right) = -\left( \max_{\delta} l'\delta_{\text{post}}, \text{ s.t. } \delta \in \Delta, \delta_{\text{pre}} = -\delta_{\text{pre}}^* \right) = -b_{\text{max}}(-\delta_{\text{pre}}^*),
\]

where the third equality uses the fact that \( \Delta \) is centrosymmetric. Hence, \( -\text{LID}(\Delta, \delta_{\text{pre}}^*) = -b_{\text{max}}(\delta_{\text{pre}}^*) - b_{\text{max}}(-\delta_{\text{pre}}^*) \). It follows from the subdifferential sum and claim rules for convex functions (e.g., Theorems 8.2 and 9.3 in Mau Nam (2019)) that

\[
\partial - \text{LID}(\Delta, \delta_{\text{pre}}^*) = \partial(-b_{\text{max}})(\delta_{\text{pre}}^*) + (-\partial(-b_{\text{max}})(\delta_{\text{pre}}^*)),
\]

for the Minkowski sum. It is then immediate that \( 0 \in \partial(-\text{LID}(\Delta, 0)) \), and hence \( 0 \in \arg\min_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} -\text{LID}(\Delta, \delta_{\text{pre}}) \). This implies that \( \text{LID}(\Delta, 0) = \sup_{\delta_{\text{pre}} \in \Delta_{\text{pre}}} \text{LID}(\Delta, \delta_{\text{pre}}) \).

To complete the proof, we show that \( \text{LID}(\Delta, \delta_{A,\text{pre}}) \geq \text{LID}(\Delta, 0) \). We first claim that for any \( \delta \in \Delta \), \( \delta + \delta_A \in \Delta \). Indeed, by centrosymmetry, \( -\delta \in \Delta \). By assumption, this implies that \( -\delta - \delta_A \in \Delta \). Applying centrosymmetry again, we see that \( \delta + \delta_A \in \Delta \), as desired. Next, suppose that \( \delta_{\text{max}} \) is optimal in the maximization \( b_{\text{max}}(0) = (\max_{\delta} l'\delta_{\text{post}}, \text{ s.t. } \delta \in \Delta, \delta_{\text{pre}} = 0) \). Then \( \delta_{\text{max}} + \delta_A \) is feasible in the optimization \( (\max_{\delta} l'\delta_{\text{post}}, \text{ s.t. } \delta \in \Delta, \delta_{\text{pre}} = \delta_{A,\text{pre}}) \), and thus \( b_{\text{max}}(\delta_{A,\text{pre}}) \geq b_{\text{max}}(0) + l'\delta_{A,\text{post}} \). By analogous argument, we can obtain that \( b_{\text{min}}(\delta_{A,\text{pre}}) \leq b_{\text{min}}(0) + l'\delta_{A,\text{post}} \). It follows that \( \text{LID}(\Delta, \delta_{A,\text{pre}}) = b_{\text{max}}(\delta_{A,\text{pre}}) - b_{\text{min}}(\delta_{A,\text{pre}}) \geq b_{\text{max}}(0) - b_{\text{min}}(0) = \text{LID}(\Delta, 0) \), as needed.

**Lemma E.30.** Fix \( \Sigma^* \) positive definite, \( \delta_A \in \Delta \), and \( \tau_A \). Suppose Assumption 3 holds at \( \delta_A \), and let \( B = B(\delta^*) \). Let \( \hat{V}_n \) denote the set of optimal vertices used in \( \psi^C(\hat{\beta}_n; A, \sqrt{n}d, \theta_n^{ab} + x, \Sigma^*) \), where \( \theta_n^{ab} = \sup S(\Delta_n, \sqrt{n}(\delta_A + \tau_A)) \), \( \Delta_n = \sqrt{n}\Delta \). Then

\[
\lim_{n \to \infty} \mathbb{P}_{(\sqrt{n}\delta_A, \sqrt{n}\tau_A, \Sigma^*)} \left( \hat{V}_n = \{c_{\gamma}^*\} \right) = 1,
\]
where $c > 0$ and $\bar{\gamma}$ is the vector such that $\bar{\gamma}_B = 0$ and $\bar{\gamma}_B$ is the unique vector such that $\bar{\gamma}^T \hat{A}_{(-1)} = 0$, $\bar{\gamma} \geq 0$, $||\bar{\gamma}|| = 1$.

**Proof.** Observe that

$$\hat{V}_n = \arg \min_{\gamma \in V(\Sigma^*)} \gamma' \hat{Y}_n,$$

where $\hat{Y}_n = A \hat{\gamma}_n - \sqrt{n} d - \hat{A}_{(-1)}(\theta_n^{\nu B} + x)$. Since all vertices $\gamma \in V(\Sigma^*)$ satisfy $\gamma^T \hat{A}_{(-1)} = 0$ by definition, we have that

$$\hat{V}_n = \arg \min_{\gamma \in V(\Sigma^*)} \gamma' \hat{Y}_n,$$

for $\hat{Y}_n = \hat{Y}_n - \hat{A}_{(-1)}(\sqrt{n} \bar{\gamma}^{\nu B})$ and $\bar{\gamma}^{\nu B}$ the vector constructed in the proof to Lemma E.9. However, we showed in the proof to Lemma E.9 that $\mathbb{E}(\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*) [\hat{Y}_{n,B}] = -\hat{A}_{(B,1)} x$ and $\mathbb{E}(\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*) \left[ \hat{Y}_{n,B} \right] \to -\infty$ as $n \to \infty$.

Lemmas F.1 and G.7 together imply that there is a unique vector $\gamma^* \in V(\Sigma^*)$ such that $\gamma^*_B = 0$, which satisfies $\gamma^*_B = c \bar{\gamma}$ for $c > 0$. By definition, $\gamma \geq 0$ for all $\gamma \in V(\Sigma^*)$, and thus $\gamma_B$ has at least one strictly positive element for all $\gamma \in V(\Sigma^*) \setminus \{\gamma^*\}$. It follows that

$$\lim_{n \to \infty} \mathbb{E}(\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*) \left[ \gamma^* \hat{Y}_n \right] = -\gamma^*_B \hat{A}_{(B,1)} x$$

$$\lim_{n \to \infty} \mathbb{E}(\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*) \left[ \gamma^T \hat{Y}_n \right] = -\infty, \forall \gamma \in V(\Sigma^*) \setminus \{\gamma^*\}.$$

Let $P_n$ denote the sequence of data-generating processes characterized by $(\sqrt{n} \delta_A, \sqrt{n} \tau_A, \Sigma^*)$. Note that for all $n$, $(\gamma^* - \gamma) \hat{Y}_n$ is normally distributed with covariance $(\gamma^* - \gamma)' \Sigma^* (\gamma^* - \gamma)$ under $P_n$. This combined with the results in the previous display imply that $\gamma^* \hat{Y}_n - \gamma^T \hat{Y}_n \xrightarrow{P} \infty$ for all $\gamma \in V(\Sigma^*) \setminus \{\gamma^*\}$. Since $V(\Sigma^*) \setminus \{\gamma^*\}$ is finite, these convergences hold uniformly over $\gamma \in V(\Sigma^*) \setminus \{\gamma^*\}$, from which we see that

$$\gamma^* \hat{Y}_n = \max_{\gamma \in V(\Sigma^*)} \gamma^T \hat{Y}_n$$

with probability approaching 1 under $P_n$, which gives the desired result. \qed
Uniform asymptotic results

The main text of the paper considers the finite sample normal model introduced in Section 3.1, which we motivated as an asymptotic approximation. In this section, we show that our main results translate to uniform asymptotic results for a large class of data-generating processes.

F.1 Assumptions

Throughout this section, we fix $\Delta = \{A \delta \leq d\}$ for some $A$ with all non-zero rows, and assume that $\Delta$ is non-empty. We consider a class of data-generating processes, indexed by $P \in \mathcal{P}$, under which $\sqrt{n}(\hat{\beta}_n - \beta_P)$ is asymptotically normal, where the asymptotic mean $\beta_P$ can be decomposed as the sum of $\delta_P \in \Delta$ and $M_{\text{post}} \tau_P$ with $\tau_P \in \mathbb{R}^T$. The parameter of interest is $\theta_P := l' \tau_P$, for some fixed $l \neq 0$.

**Assumption 6.** Let $BL_1$ denote the set of Lipschitz functions which are bounded by 1 in absolute value and have Lipschitz constant bounded by 1. We assume

$$\lim_{n \to \infty} \sup_{P \in \mathcal{P}} \sup_{f \in BL_1} \left\| \mathbb{E}_P \left[ f(\sqrt{n}(\hat{\beta}_n - \beta_P)) \right] - \mathbb{E} \left[ f(\xi_P) \right] \right\| = 0,$$

where $\xi_P \sim \mathcal{N}(0, \Sigma_P)$, and $\beta_P = \delta_P + M_{\text{post}} \tau_P$ for $\delta_P \in \Delta$ and $\tau_P \in \mathbb{R}^T$.

Convergence in distribution is equivalent to convergence in bounded Lipschitz metric (see Theorem 1.12.4 in van der Vaart and Wellner (1996)), so Assumption 6 formalizes the notion of uniform convergence in distribution of $\sqrt{n}(\hat{\beta}_n - \beta_P)$ to a $\mathcal{N}(0, \Sigma_P)$ variable under $P$.

Our next assumption requires that the eigenvalues of the asymptotic variance of the event-study coefficients be bounded above and away from zero.

**Assumption 7.** Let $S$ denote the set of matrices with eigenvalues bounded below by $\lambda > 0$ and above by $\bar{\lambda} \geq \lambda$. For all $P \in \mathcal{P}$, $\Sigma_P \in S$.

Next, we assume that there is a uniformly consistent estimator of the variance of $\hat{\beta}$.

**Assumption 8.** We have an estimator $\hat{\Sigma}$ that is uniformly consistent for $\Sigma_P$,

$$\lim_{n \to \infty} \sup_{P \in \mathcal{P}} \mathbb{P}_P \left( \| \hat{\Sigma}_n - \Sigma_P \| > \epsilon \right) = 0,$$

for all $\epsilon > 0$.

In order to more clearly articulate our next assumption, it is useful to first present the following result, which characterizes the set of dual vertices under Assumption 7.

---

44To avoid notational clutter, we drop the additional subscript “post” on $\tau$ and simply index $\tau$ by the underlying data generating process $P$. 

---

S-51
Lemma F.1. Let \( F(\Sigma) := \{ \gamma : \hat{A}_i(\gamma) = 0, \hat{\sigma}(\Sigma)'\gamma = 1, \gamma \geq 0 \} \) be the feasible set of the dual problem, where \( \hat{\sigma}(\Sigma) \) is the vector containing the square-roots of the diagonal elements of \( \Sigma A' \). Let \( V(\Sigma) \) denote the set of vertices of \( F(\Sigma) \). Then there exists a finite set of distinct, non-zero vectors \( \tilde{\gamma}_1, \ldots, \tilde{\gamma}_J \) such that \( ||\tilde{\gamma}_j|| = 1 \) and \( \gamma_j \geq 0 \) for all \( j \), and for any \( \Sigma \) positive definite

\[
V(\Sigma) = \{ c_1(\Sigma)\tilde{\gamma}_1, \ldots, c_J(\Sigma)\tilde{\gamma}_J \},
\]

where \( c_j(\Sigma) = (\tilde{\gamma}_j'\hat{\sigma}(\Sigma))^{-1} \).

For ease of notation, we define \( \gamma_j(\Sigma) := c_j(\Sigma)\tilde{\gamma}_j \). With this notation in hand, we can then state our next assumption.

Assumption 9. Suppose \( \tilde{\gamma}_j'A \neq 0 \). Then for all \( i \neq j \) and all \( P \in \mathcal{P} \),

\[
(\gamma_j(\Sigma_P) - \gamma_i(\Sigma_P))'A\Sigma_P A'(\gamma_i(\Sigma_P) - \gamma_j(\Sigma_P)) > c,
\]

for some constant \( c > 0 \).

Assumption 9 guarantees that there are not two vertices of the feasible set that produce non-degenerate objective values in the dual problem (16) and are perfectly correlated asymptotically. Assumption 9 holds trivially if the minimal eigenvalue of \( \Sigma A' \) is bounded from below. Note that under Assumption 6, \( \Sigma A' \) is the asymptotic variance of \( \sqrt{n}\hat{\beta}_n \), and thus corresponds with the asymptotic variance of \( \sqrt{n}\hat{Y}_n(\tilde{\theta}) \), the moments used in the conditional and hybrid tests scaled by \( \sqrt{n} \). Assumption 9 can be dispensed with if we use a modified version of the conditional and hybrid tests that adds full-rank normal noise to \( \hat{Y}_n \) such that the asymptotic covariance of the scaled moments is positive definite.

F.2 Size control

We now establish uniform asymptotic size control for the conditional and conditional-least favorable hybrid procedures. ARP establish uniform asymptotic size control under high-level conditions, whereas here we show size control in our setting under the conditions introduced above. These conditions are somewhat weaker than the higher-level conditions in ARP. For instance, we allow for the possibility that \( \hat{\eta} \) has zero variance conditional on a set of optimal multipliers, which is ruled out by assumptions in ARP but can be shown to arise in our context.

As in ARP, we show size control for a modified version of the conditional and hybrid tests that never rejects if the critical value is below a certain finite value \(-C\). That is, we consider \( \psi^C_{*,\alpha} = \psi^C_\alpha \mathbf{1}[\hat{\eta} \geq -C] \), for \( \psi^C_\alpha \) an indicator for whether the \( \alpha \)-level conditional test rejects and \( \hat{\eta} \) the solution to the linear program (15). Likewise, we consider the modified conditional least favorable hybrid \( \psi^{C,-LF}_{*,\alpha} \) defined analogously. We do this for technical reasons to avoid complications related to sequences where both \( \hat{\eta} \) and the critical values diverge to \(-\infty \). However, this modification
is reasonable on substantive grounds, since when \( \hat{\eta} \) is very small all of the moments are satisfied in the data, and the conditional test (potentially) rejects only due to extreme realizations of the critical values. Moreover, we show in Section F.4 below that the modified tests retain desirable asymptotic power properties.

Under the assumptions stated in the previous section, the modified conditional test uniformly controls size.

**Proposition F.1.** Suppose Assumptions 6 to 9 hold. Then

\[
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \mathbb{E} \left[ \psi^C_{\hat{\beta}_n, A, d, \theta_P, \frac{1}{n} \hat{\Sigma}_n} \right] \leq \alpha.
\]

Likewise, the modified conditional least-favorable hybrid test controls size as well.

**Corollary F.1.** Suppose Assumptions 6 to 9 hold. Then, for any \( \alpha \in (0, 1) \) and \( \kappa \in (0, \min\{\alpha, 0.5\}) \),

\[
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \mathbb{E} \left[ \psi^{C-LF}_{\hat{\beta}_n, A, d, \theta_P, \frac{1}{n} \hat{\Sigma}_n} \right] \leq \alpha.
\]

**F.3 Consistency**

We now provide conditions under which the conditional and conditional-least favorable hybrid tests are uniformly consistent. The results in this section are uniform asymptotic versions of the consistency results in Section 5 in the context of the finite sample normal model.

For the least-favorable hybrid test, we obtain uniform consistency under Assumptions 6 to 8. (Note Assumption 9 is not necessary for the hybrid to be consistent).

**Proposition F.2.** Suppose Assumptions 6 to 8 hold. Let \( \theta_{ub} = \sup_S \Delta, \beta \) and \( \mathcal{P} \). Then, for any \( x > 0 \),

\[
\liminf_{n \to \infty} \inf_{P \in \mathcal{P}} \mathbb{E} \left[ \psi^{C-LF}_{\hat{\beta}_n, A, d, \theta_{ub}, \frac{1}{n} \hat{\Sigma}_n} \right] = 1.
\]

To show uniform consistency for the conditional test, we require some additional assumptions on the asymptotic distribution of the estimated covariance matrix \( \hat{\Sigma} \).

**Assumption 10.** Let \( W_n = ((\hat{\beta}_n - \beta_P)', (\text{vec}(\hat{\Sigma}_n) - \text{vec}(\Sigma_P))')', \) where \( \text{vec}(\Sigma) \) is the vector of the elements of the matrix \( \Sigma \). We assume

\[
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \sup_{f \in BL_1} \| \mathbb{E} \left[ f(\sqrt{n}W_n) \right] - \mathbb{E} \left[ f(\xi_P^+) \right] \| = 0,
\]

where \( \xi_P^+ \sim \mathcal{N}(0, V_P) \), \( V_P = \begin{pmatrix} \Sigma_P & V_P,\Sigma_P \\ V_P,\Sigma_P & V_P,\Sigma \end{pmatrix} \) and \( \beta_P = \delta_P + M_{post}\tau_P \) for \( \delta_P \in \Delta \) and \( \tau_P \in \mathbb{R}^T \).

\footnote{The assumption that \( \kappa < 0.5 \) can be relaxed if either \( -C > 0 \) or if the hybrid test is modified not to reject whenever \( \sigma_P = 0 \) and \( \hat{\eta} \leq 0 \).}
Assumption 11. For all $P \in \mathcal{P}$, the matrix $V_P$ defined in Assumption 10 lies in a compact set $V$. Additionally, $\Sigma_P$ has eigenvalues bounded between $\lambda > 0$ and $\bar{\lambda}$, and $(\Sigma_P - V_P \Sigma V_P^{-1} V_P \Sigma \beta)$ has eigenvalues bounded below by $\tilde{\lambda} > 0$.

Assumption 10 strengthens Assumption 6 to require that the pair $(\hat{\beta}, \hat{\Sigma})$ converge uniformly to a joint normal distribution centered at their respective means. Although somewhat more restrictive, we note that event-study estimates are often estimated via OLS, and standard covariance estimators for OLS, including cluster-robust variance estimators, produce asymptotically normal estimates as the number of clusters grows large (Hansen, 2007; Stock and Watson, 2008; Hansen and Lee, 2019).

Note that we do not impose that the asymptotic distributions of $\hat{\beta}$ and $\hat{\Sigma}$ are independent, as would occur in linear models if the linear model is properly specified. Likewise, Assumption 11 strengthens Assumption 7 to require that the asymptotic variance matrix of the pair $(\hat{\beta}, \hat{\Sigma})$ lies in a compact set, and that the error in $\hat{\beta}$ is not perfectly colinear with the error in $\hat{\Sigma}$. The latter condition can be ensured to hold by adding full-rank noise to $\hat{\beta}$. With these added conditions, we obtain asymptotic consistency for the (modified) conditional test.

Proposition F.3. Suppose Assumptions 8 to 11 hold. Then for any $x > 0$,

$$\lim_{n \to \infty} \inf_{P \in \mathcal{P}} \mathbb{E}_P \left[ \psi_{*,\alpha}^C (\hat{\beta}_n, A, d, \theta_P^{ub} + x, \frac{1}{n} \hat{\Sigma}_n) \right] = 1.$$ 

F.4 Local Asymptotic Power

We now establish conditions under which the power of the conditional test converges uniformly to the power envelope. The results in this section are uniform asymptotic versions of the results in Section 5 for the finite sample normal model.

Recall that in the finite sample normal model, we showed that the local power of the conditional test converged to the power envelope under Assumption 3, which intuitively guaranteed that the “right” number of moments bind at the edge of the identified set. We define $\mathcal{P}_\epsilon$ to be the set of distributions for which this condition holds and the non-binding moments are slack by at least $\epsilon$.

**Definition 2.** For $\epsilon > 0$, let $\mathcal{P}_\epsilon$ denote the set of distributions $P \in \mathcal{P}$ such that Assumption 3 holds when setting $\delta_A = \delta_P$, and for which all elements of the vectors $\epsilon_{B(\beta^*)}$ and $\epsilon_{B(\beta^{**})}$ as defined in Assumption 3 are bounded below by $\epsilon$.

Recall from Appendix A that our Assumption 3 is implied by linear independence constraint qualification (LICQ). Assuming that $P \in \mathcal{P}_\epsilon$ is thus similar to a uniform LICQ assumption, as in e.g., Gafarov (2019) and Cho and Russell (2018). We note, however, that we require this assumption only for our uniform local asymptotic power results, and not for uniform asymptotic size control.

Our next result states that the local power of the conditional test converges to the power envelope in the limiting model uniformly over $\mathcal{P}_\epsilon$. This can be viewed as an asymptotic version of Proposition 5.2.
Proposition F.4. Suppose Assumptions 6 to 8 hold. Let $\theta_P^{ub} = \sup S(\Delta, \beta_P)$. Then for any $\epsilon > 0$ and $x > 0$,

$$
\lim_{n \to \infty} \sup_{P \in \mathcal{P}_n} \mathbb{E}_P \left[ q^{C*}_{s,\alpha}(\beta, A, d, \theta_P^{ub} + \frac{1}{\sqrt{n}} x, \frac{1}{n} \Sigma) - \rho^*(P, x) \right] = 0,
$$

where

$$
\rho^*(P, x) = \lim_{n \to \infty} \sup_{C \subset \mathcal{P}, \tau \in \mathbb{R}^+} \mathbb{P}(\delta P, \tau P, \frac{1}{n} \Sigma) \left( (\theta_P^{ub} + \frac{1}{\sqrt{n}} x, \frac{1}{n} \Sigma) \notin C \right)
$$

is the optimal limiting power of a size-$\alpha$ test in the finite sample normal model using $(\delta_A, \tau_A, \Sigma^*) = (\delta_P, \tau_P, \Sigma_P)$, provided that $-C$, the threshold for the modified conditional test, is set sufficiently small.

One can show that if $\alpha \in (0, 5]$, then $C = 0$ is sufficient for the conclusion of Proposition F.4 to hold.

Proposition F.4 shows that the power of the conditional test converges to the power of the optimal test in the limit of the finite sample normal model as $n \to \infty$. Using results from Müller (2011), we next show that the power bound $\rho^*(P, x)$ from the limiting model is an upper bound on the asymptotic power of a large class of confidence sets that control size asymptotically.

In particular, we consider the set of confidence sets that i) can be written as functions of $\sqrt{n} \hat{\beta}_n$ and $\hat{\Sigma}_n$, ii) control size asymptotically over all sequences of distributions that induce a normal limit, and iii) are invariant to transformations that preserve the identified set for all values of $\beta$. To formalize iii), let $A^\perp = \{v : Av = 0\}$ denote the null space of $A$ and let $G$ be the group of transformations of the form $g_v : \beta \mapsto \beta + v$ for $v \in A^\perp$. It is then immediate from the definition of the identified set, $S(\Delta, \beta) = \{\theta : \exists \delta \in \Delta, \tau_{\text{post}} \text{ s.t. } \beta = \delta + M_{\text{post}} \tau_{\text{post}}, \ell' \tau_{\text{post}} = \theta\}$, that $S(\Delta, \beta) = S(\Delta, g_v \beta)$ for any $\beta$ and $g_v \in G$. By iii) we mean that we will consider the class of confidence sets such that $C(\sqrt{n} \hat{\beta}, \hat{\Sigma}) = C(g_v(\sqrt{n} \hat{\beta}, \hat{\Sigma}))$ for all $g_v \in G$ and all $\hat{\beta}$.

Proposition F.5. Suppose that $C(\hat{\beta}, \hat{\Sigma})$ is such that

$$
\lim_{n \to \infty} \sup_{P_n} \mathbb{P}_{P_n} \left( \theta_{P_n} \notin C(\sqrt{n} \hat{\beta}_n, \hat{\Sigma}_n) \right) \leq \alpha
$$

for any sequence of distributions $P_n$ such that $\sqrt{n} (\hat{\beta}_n - \beta_{P_n}) \to_d N(0, \Sigma^*)$ and $\hat{\Sigma}_n \to_p \Sigma^*$ under $P_n$, and any value $\theta_{P_n}$ for which there exists $\delta_{P_n} \in \Delta$ and $\tau_{P_n,\text{post}}$ such that $\beta_{P_n} = \delta_{P_n} + M_{\text{post}} \tau_{P_n,\text{post}}$ and $\theta_{P_n} = \ell' \tau_{P_n,\text{post}}$. Assume also that $C(\sqrt{n} \hat{\beta}, \hat{\Sigma}) = C(g_v(\sqrt{n} \hat{\beta}, \hat{\Sigma}))$ for all $g_v \in G$ and all $\hat{\beta}$.

Suppose that under $P^*$, $\sqrt{n} (\hat{\beta}_n - \beta_{P^*}) \to_d N(0, \Sigma^*)$ and $\hat{\Sigma}_n \to_p \Sigma^*$, where $\beta_{P^*} = \delta_{P^*} +

$$
\begin{pmatrix}
0 \\
\tau_{P^*,\text{post}}
\end{pmatrix} \quad \text{for } \delta_{P^*} \in \Delta
$$

satisfying Assumption 3. Let $\theta_{P^*}^{ub} = \sup S(\Delta, \beta_{P^*})$ be the upper bound of the identified set given $\beta_{P^*}$. Then, for any $x > 0$,
\[
\limsup_{n \to \infty} P^* \left( \theta_P^{ub} + \frac{1}{\sqrt{n}} x \notin C(\sqrt{n} \hat{\beta}_n, \hat{\Sigma}_n) \right) \leq \rho^*(P^*, x),
\]
where \( \rho^*(P^*, x) \) is defined in Proposition F.4.

As in the finite sample normal model, we also obtain that the conditional-least favorable hybrid test that uses a size \( \kappa \) first stage does no worse than the best size \( \tilde{\alpha} = \frac{\alpha - \kappa}{1 - \kappa} \) test.

**Corollary F.2.** Suppose Assumptions 6 to 8 hold. Let \( \theta_P^{ub} = \sup S(\Delta, \delta_P, \tau_P) \). Then, for any \( \epsilon > 0 \) and \( x > 0 \),

\[
\limsup_{n \to \infty} \sup_{P \in P_X} \left[ \rho^*(P, x; \tilde{\alpha}) - \mathbb{E}_P \left[ \psi^C_{\kappa, \alpha} \left( \hat{\beta}_n, A, d, \theta_P^{ub} + \frac{1}{\sqrt{n}} x, \frac{1}{n} \hat{\Sigma}_n \right) \right] \right] \leq 0,
\]

where

\[
\rho^*(P, x, \tilde{\alpha}) = \lim_{n \to \infty} \sup_{C_{\tilde{\alpha}, n} \in C_{\tilde{\alpha}}} \mathbb{P}_{(\delta_P, \tau_P, \frac{1}{n} \Sigma_P)} \left( (\theta_P^{ub} + \frac{1}{\sqrt{n}} x) \notin C_{\tilde{\alpha}, n} \right)
\]

is the optimal limiting power of a size \( \tilde{\alpha} \) test in the finite sample normal model using \((\delta_A, \tau_A, \Sigma^*) = (\delta_P, \tau_P, \Sigma_P)\), provided that \(-C\), the threshold for the modified hybrid test, is set sufficiently small.

**G Proofs of uniform asymptotic results**

**G.1 Proofs and Auxiliary Lemmas for Uniform Size Control**

**Proof of Lemma F.1**

**Proof.** Recall from Section 8.5 of Schrijver (1986) that \( v \) is a vertex of the polyhedron \( P = \{ x \in \mathbb{R}^K : Wx \leq b \} \) iff \( v \in P \) and \( W_{(J, \cdot)} x = b_J \) for \( J \) a set of indices such that \( W_{(J, \cdot)} \) has \( K \) independent rows. It follows that \( v \in V(\Sigma) \) iff \( v \geq 0 \) and there exists \( J \) such that

\[
W_{J, \cdot} := \begin{pmatrix} \tilde{A}'_{(\cdot,-1)} \\ I_{(J, \cdot)} \\ \hat{\sigma}' \end{pmatrix}
\]

has row rank equal to \( K \), and \( W_{J, \cdot} v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \).

Now, let \( J \) be the set of indices \( J \) such that \( \tilde{W}_J := \begin{pmatrix} \tilde{A}'_{(\cdot,-1)} \\ I_{(J, \cdot)} \end{pmatrix} \) has \( K - 1 \) linearly independent rows and there exists a vector \( v_J \neq 0 \) such that \( \tilde{W}_J v = 0 \) and \( v_J \geq 0 \). Since by construction \( \tilde{W}_J \) has rank \( K - 1 \) and \( K \) columns, its nullspace is 1-dimensional. It is then immediate that for
each $\mathcal{J} \in \mathcal{J}$, there is a unique vector $\tilde{v}_\mathcal{J} \geq 0$ such that $||\tilde{v}_\mathcal{J}|| = 1$ and $\tilde{W}_\mathcal{J} \tilde{v}_\mathcal{J} = 0$. Moreover, $\mathcal{J}$ is finite, since there are a finite number of possible subindices of $I$, and thus we can write \{\tilde{v}_\mathcal{J} : \mathcal{J} \in \mathcal{J}\} = \{\tilde{v}_1, ..., \tilde{v}_\mathcal{J}\}$ for distinct vectors $\tilde{v}_1, ..., \tilde{v}_\mathcal{J}$.

It now remains to show that $V(\Sigma) = \{c_1(\Sigma)\tilde{v}_1, ..., c_J(\Sigma)\tilde{v}_J\}$, for $c_j$ as defined above. First, suppose that $v = c_J(\Sigma)\tilde{v}_J$. By construction, $\tilde{A}'(-1)v = 0$, $v \geq 0$, and $\tilde{\sigma}'v = (\tilde{\sigma}'v_j)^{-1}(\tilde{\sigma}'v_j) = 1$, and so $v \in F$. Additionally, there exists $\mathcal{J}$ such that $\tilde{W}_\mathcal{J} = \begin{pmatrix} \tilde{A}'(-1) \\ I(\mathcal{J}, \cdot) \end{pmatrix}$ has rank $K-1$ and $\tilde{W}_\mathcal{J}v = 0$. From the fact that $\tilde{W}_\mathcal{J}v = 0$, whereas $\tilde{\sigma}'v = 1$, we see that $\tilde{\sigma}'$ must be linearly independent from the rows of $\tilde{W}_\mathcal{J}$, and thus $W_\mathcal{J} = \begin{pmatrix} \tilde{W}_\mathcal{J} \\ \tilde{\sigma}' \end{pmatrix}$ has rank $K$. It follows that $v \in V(\Sigma)$.

Next, suppose that $v \in V(\Sigma)$. Then $v \geq 0$, and there exists $\mathcal{J}$ such that

$$W_\mathcal{J} := \begin{pmatrix} \tilde{A}'(-1) \\ I(\mathcal{J}, \cdot) \end{pmatrix}$$

has row rank equal to $K$, and $W_\mathcal{J}v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Let $\tilde{W}_\mathcal{J} = \begin{pmatrix} \tilde{A}'(-1) \\ I(\mathcal{J}, \cdot) \end{pmatrix}$. Note that since $\tilde{W}_\mathcal{J}'v = 0$, whereas $\tilde{\sigma}'v = 1$, $\tilde{\sigma}'$ must be linearly independent of the other rows of $W_\mathcal{J}$, from which it follows that $\tilde{W}$ has row rank $K-1$. Thus, $\mathcal{J} \in \mathcal{J}$, and so $v = c\tilde{v}_j$ for some $j$ and $c > 0$. Since $\tilde{\sigma}'v = 1$, we have $c\tilde{\sigma}'\tilde{v}_j = 1$, which implies $c = (\tilde{\sigma}'\tilde{v}_j)^{-1}$, which gives the desired result.

**Proof of Proposition F.1**

**Proof.** First, note that by Lemma E.2, $\psi_C(\hat{\beta}_n, A, d, \theta P, \frac{1}{n}\hat{\sum}_n) = \psi_C(\sqrt{n}\hat{\beta}_n, A, \sqrt{n}d, \sqrt{n}\theta P, \hat{\sum}_n)$. Additionally, we show in the proof to Lemma E.2 that the values of $\hat{\eta}$ for these two problems are the same, from which it follows that the modified tests are tests are equivalent as well, $\psi^C(\hat{\beta}_n, A, d, \theta P, \frac{1}{n}\hat{\sum}_n) = \psi^C(\sqrt{n}\hat{\beta}_n, A, \sqrt{n}d, \sqrt{n}\theta P, \hat{\sum}_n)$. It thus suffices to show that

$$\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \mathbb{E}[\psi^C(\sqrt{n}\hat{\beta}_n, A, \sqrt{n}d, \sqrt{n}\theta P, \hat{\sum}_n)] \leq \alpha.$$ 

Towards contradiction, suppose the proposition is false. Then, following Andrews et al. (Forthcoming), there exists a sequence of distributions $P_m$ and an increasing sequence of sample sizes $n_m$ such that

$$\liminf_{m \to \infty} \mathbb{E}_{P_m} \left[ \psi^C(\sqrt{n_m}\hat{\beta}_{n_m}, A, \sqrt{n_m}d, \sqrt{n_m}\theta P, \hat{\sum}_{n_m}) \right] \geq \alpha + \omega, \tag{67}$$

for some $\omega > 0$.

Define $Y_m := \sqrt{n_m}\left(A\hat{\beta}_{n_m} - d - \tilde{A}'(-1)\theta P_m\right)$ and $X := \tilde{A}'(-1)$. Then,
Further, define \( \tilde{Y}_m := Y_m - \tilde{A}_{(\cdot,1)} \Gamma_{(\cdot,1)(\cdot)} \) for notational convenience, let \( \Sigma_m := \Sigma_{P_m} \) and \( \bar{\Sigma}_m := \hat{\Sigma}_{nm} \). By Lemma 16 in ARP, \( \psi_{*,\alpha}^{C}(Y_m, X, A\Sigma_{m}A^t) = \psi_{*,\alpha}^{C}(\tilde{Y}_m, X, A\bar{\Sigma}_mA^t) \). Additionally, observe that

\[
\tilde{Y}_m = \sqrt{m} \left( A\beta_{nm} - d - \tilde{A}_{(\cdot,1)} \theta_{P_m} - \tilde{A}_{(\cdot,1)} \Gamma_{(\cdot,1)(\cdot)} \bar{\tau}_{P_m} \right)
\]

where the first equality uses the definition of \( \theta_{P_m} = t' \bar{\tau}_{P_m} \) and the second equality follows from Lemma G.5. This implies that

\[
\tilde{Y}_m = A \sqrt{m} \left( \beta_{nm} - \delta_{P_m} - \left( \begin{array}{c} 0 \\ \bar{\tau}_{P_m} \end{array} \right) \right) + \sqrt{m} \left( A\delta_{P_m} - d \right).
\]

Next, observe that by Assumption 6, \( \delta_{P} \in \Delta = \{ \delta : A\delta \leq d \} \) for all \( P \), and so \( \sqrt{m} \left( A\delta_{P_m} - d \right) \leq 0 \). We can therefore extract a subsequence \( m_1 \) such that

\[
\sqrt{m_1} \left( A\delta_{P_m} - d \right) \rightarrow \mu^*_1 \in R \cup \{-\infty\}.
\]

Passing to further subsequences, we can extract a subsequence \( m_K \) (for \( K \) the number of rows of \( A \)) along which

\[
\sqrt{m_K} \left( A\delta_{P_{m_K}} - d \right) \rightarrow \mu^* \in [R \cup \{-\infty\}]^K.
\]

Additionally, by Assumption 7, \( \Sigma_{P_m} \) is contained within a compact set, and so we can extract a further subsequence \( m_{K+1} \) along which \( \Sigma_{m_{K+1}} \rightarrow \Sigma^* \) for some \( \Sigma^* \in S \). For notational ease, we will assume that these convergences hold for the original sequence \( (m, n_m) \) for the remainder of the proof.

Now, equation (68) along with Assumptions 6 and 8 and the continuous mapping theorem imply that

\[
(\tilde{Y}_m, \hat{\Sigma}_m) \overset{d}{\rightarrow} (\xi + \mu^*, \Sigma^*),
\]

for \( \xi \sim \mathcal{N}(0, A\Sigma^*A^t) \). Observe from (68) that for all \( m \), \( \tilde{Y} \in \text{col}(A) + \{-a \cdot d : a > 0\} \), where \( \text{col}(A) \) is the column space of \( A \) and \( + \) represents the Minkowski sum. Likewise, if \( \xi \sim \mathcal{N}(0, A\Sigma^*A^t) \), then
\( \xi = A \xi \) for \( \xi \sim N(0, \Sigma) \), and so \( \xi \) is supported on \( \text{col}(A) \). Thus, \( \xi + \mu^* \) is supported on \( \text{col}(A) + \mu^* \). We then see that both \( \hat{Y}_m \) and \( \xi + \mu^* \) are supported on \( \Omega := \text{col}(A) + \{ -a \cdot d : a \in \mathbb{R} \} \cup \{ \mu^* \} \).

Suppose first that \( \max_{\gamma \in \mathbb{V}(\Sigma^*)} \gamma' \mu^* = -\infty \). Note that \( \hat{\eta}_m = \max_{\gamma \in \mathbb{V}(\Sigma_m)} \gamma' \hat{Y}_m \). From Lemma F.1, \( \mathbb{V}(\Sigma) = \{ c_1(\Sigma)\gamma_1, \ldots, c_J(\Sigma)\gamma_J \} \), where the functions \( c_j(\Sigma) \) are continuous and by Lemma G.3, \( c_j(\Sigma^*) \geq -C > 0 \) for all \( j \). Since \( \max_{\gamma \in \mathbb{V}(\Sigma^*)} \gamma' \mu^* = -\infty \), for all \( j \), we have \( c_j(\Sigma^*)\gamma_j^* \mu^* = -\infty \). But the continuous mapping theorem then implies that for all \( j \), \( c_j(\hat{\Sigma}_m)\gamma_j^* \hat{Y}_m \to_d c_j(\Sigma^*)\gamma_j^* (\xi + \mu^*) = -\infty \), and hence \( \hat{\eta}_m \to_p -\infty \). Thus, \( \mathbb{P}(\hat{\eta}_m < -C) \to 1 \), and so our tests never reject asymptotically, which contradicts size control failing. For the remainder of the proof, we assume that \( \max_{\gamma \in \mathbb{V}(\Sigma^*)} \gamma' \mu^* \) is finite. (Note that since \( \gamma_j^* \geq 0 \) and \( \mu^* \leq 0 \), we cannot have \( \max_{\gamma \in \mathbb{V}(\Sigma^*)} \gamma' \mu^* = -\infty \).)

Next, note that it follows readily from the construction of the (unmodified) conditional test in Section 4 that the unmodified conditional test rejects iff

\[
p(Y, \Sigma) := \mathbb{P} \left( \zeta < \hat{\eta}(Y, \Sigma) \mid \zeta \in [v_{lo}(Y, \Sigma), v_{up}(Y, \Sigma)], \zeta \sim \mathcal{N}(0, \sigma^2_\eta(Y, \Sigma)) \right) > 1 - \alpha,
\]

where the functions \( \hat{\eta} \), \( \sigma^2_\eta \), \( v_{lo} \) and \( v_{up} \) are defined as follows. We define \( \hat{\eta}(Y, \Sigma) \) to be the conditional test statistic using \( Y \) and \( \Sigma \),

\[
\hat{\eta}(Y, \Sigma) := \max_{\gamma \in \mathbb{V}(\Sigma)} \gamma' Y,
\]

We define \( \sigma^2_\eta(Y, \Sigma) \) to be the estimated variance of \( \gamma_*' Y \) for \( \gamma_* \in \arg \max_{\gamma \in \mathbb{V}(\Sigma)} \gamma' Y \). That is,

\[
\sigma^2_\eta(Y, \Sigma) = \gamma_*' \Sigma \Sigma' \gamma_*.
\]

Note that \( \sigma^2_\eta(Y, \Sigma) \) is only well-defined if \( \gamma_*' \Sigma \Sigma' \gamma_* \) is the same for all \( \gamma_* \in \arg \max_{\gamma \in \mathbb{V}(\Sigma)} \gamma' Y \). We will show below, however, that this occurs with probability 1 in the limiting model.

If \( \sigma^2_\eta(Y, \Sigma) > 0 \), then we define \( v_{lo}(Y, \Sigma) \) and \( v_{up}(Y, \Sigma) \) to be the minimum and maximum of the set

\[
C = \{ c : \max_{\gamma \in \mathbb{V}(\Sigma)} \gamma' \left( S_{\gamma_*} + \frac{\hat{\Sigma}_{\gamma_*}}{\gamma_*' \hat{\Sigma}_{\gamma_*}} c \right) \},
\]

where as before \( \gamma_* \) is an element of \( \arg \max_{\gamma \in \mathbb{V}(\Sigma)} \gamma' Y \) and we define

\[
S_{\gamma_*} = \left( I - \frac{\hat{\Sigma}_{\gamma_*} \gamma_*'}{\gamma_*' \hat{\Sigma}_{\gamma_*}} \right) Y.
\]

On the other hand, if \( \sigma^2_\eta(Y, \Sigma) = 0 \), then we define \( v_{lo} = -\infty \) and \( v_{up} = \infty \). This is a notational convenience that allows us to capture the fact that when \( \sigma^2_\eta = 0 \), the unmodified conditional test
rejects iff \( \hat{\eta}(Y, \Sigma) > 0 \), since \( \mathbb{P}(\zeta < \hat{\eta} | \zeta \sim \mathcal{N}(0, 0)) = 1[\hat{\eta} > 0] \).

Since the modified conditional test rejects only if the unmodified conditional test rejects, \((67)\) thus implies that

\[
\liminf_{m \to \infty} \mathbb{P}_{P_m} \left( p(Y_m, \hat{\Sigma}) > 1 - \alpha \right) \geq \alpha + \omega.
\]

(69)

Lemma G.1 shows that the function \( p(X, Y) \) is continuous at \((\xi + \mu^*, \Sigma^*)\) for almost every \( \xi \sim \mathcal{N}(0, A\Sigma^* A') \). The continuous mapping theorem then implies that

\[
p(Y_m, \hat{\Sigma}) \xrightarrow{d} p(\xi + \mu^*, \Sigma^*).
\]

Moreover, Lemma G.2 implies that the distribution of \( p(\xi + \mu^*, \Sigma^*) \) does not have a mass point at \( 1 - \alpha \), and hence

\[
\mathbb{P}_{P_m} \left( p(Y_m, \hat{\Sigma}) > 1 - \alpha \right) \rightarrow \mathbb{P} \left( p(\xi + \mu^*, \Sigma^*) > 1 - \alpha \right).
\]

However, since the conditional test controls size in the finite-sample normal model,

\[
\mathbb{P}_{\xi} \left( p(\xi + \mu^*, \Sigma^*) > 1 - \alpha \right) \leq \alpha,
\]

and thus

\[
\liminf_{m \to \infty} \mathbb{P}_{P_m} \left( p(Y_m, \hat{\Sigma}) > 1 - \alpha \right) \leq \alpha,
\]

which contradicts \((69)\).

Proof of Corollary F.1

Proof. Observe that the conditional least-favorable hybrid test is of nearly the same form as the conditional test, and we can thus obtain the result via a slight modification to the proof of Proposition F.1. Recall from the proof to Proposition F.1, the (unmodified) conditional test rejects iff \( p(Y, \Sigma) > 1 - \alpha \), for

\[
p(Y, \Sigma) := \mathbb{P} \left( \zeta < \hat{\eta}(Y, \Sigma) | \zeta \in [v^{lo}(Y, \Sigma), v^{up}(Y, \Sigma)], \zeta \sim \mathcal{N}(0, \sigma_\eta^2(Y, \Sigma)) \right),
\]

and showed that \( p(\xi + \mu^*, \Sigma^*) \) is continuous for almost every \( \xi \sim \mathcal{N}(0, A\Sigma^* A') \). Size control then followed from the continuous mapping theorem combined with size control in the finite sample normal model. Similarly, observe that the conditional least-favorable hybrid test rejects iff \( p_{C-LF}(Y, \Sigma) > 1 - \tilde{\alpha} \), for
\[ p_{\text{C-LF}}(Y, \Sigma) = 1[\hat{\eta}(Y, \Sigma) > c^L_\kappa(\Sigma)] + 1[\hat{\eta}(Y, \Sigma) \leq c^L_\kappa(\Sigma)] \hat{p}_{\text{C-LF}}(Y, \Sigma), \]

where

\[ \hat{p}_{\text{C-LF}}(Y, \Sigma) := P\left( \zeta < \hat{\eta}(Y, \Sigma) \mid \zeta \in [v^{lo}(Y, \Sigma), v^{up}_{\text{C-LF}}(Y, \Sigma)], \zeta \sim \mathcal{N}(0, \sigma^2_\eta(Y, \Sigma)) \right) \]

is nearly identical to \( p(Y, \Sigma) \) except that it replaces \( v^{up}(Y, \Sigma) \) with \( v^{up}_{\text{C-LF}}(Y, \Sigma) = \min\{c^L_\kappa(\Sigma), v^{up}(Y, \Sigma)\} \).

To show that \( p_{\text{C-LF}}(Y, \Sigma) \) is almost surely continuous at \((\xi + \mu^*, \Sigma^*)\), it suffices to show that \( c^L_\kappa(\Sigma) \) and \( 1[\eta(Y, \Sigma) > c^L_\kappa(\Sigma)] \) are almost surely continuous, as the remaining parts of the proof are identical to those for \( p(Y, \Sigma) \).

We first show that the least-favorable critical value \( c^L_\kappa(\Sigma) \) is continuous in \( \Sigma \) at \( \Sigma^* \). Recall that \( c^L_\kappa(\Sigma) = q_{1-\kappa} \left( \max_{\gamma \in \mathcal{V}(\Sigma)} \gamma' \xi^* : \xi^* \sim \mathcal{N}(0, \Lambda \Sigma A') \right) \). Lemma F.1 implies that \( \max_{\gamma \in \mathcal{V}(\Sigma)} \gamma' \xi^* = \max_1 \{ c_1(\Sigma) \gamma_1^* \xi^*, \ldots, c_J(\Sigma) \gamma_J^* \xi^* \} \) for functions \( c_j(\Sigma) \) that are continuous at \( \Sigma^* \). This implies that \( \max_{\gamma \in \mathcal{V}(\Sigma)} \gamma' \xi^* \) is continuous in \( \Sigma \) for every \( \xi \). The continuity of \( c^L_\kappa(\Sigma) \) follows immediately.

Next, we show that \( 1[\hat{\eta}(Y, \Sigma) > c^L_\kappa(\Sigma)] \) is continuous at \((\xi + \mu^*, \Sigma^*)\) for almost every \( \xi \sim \mathcal{N}(0, \Lambda \Sigma A') \). We’ve already shown that \( c^L_\kappa(\Sigma) \) is continuous, and Lemma G.4 implies that \( \hat{\eta}(Y, \Sigma) \) is continuous at \((\xi + \mu^*, \Sigma^*)\) for almost every \( \xi \). It thus suffices to show that \( \hat{\eta}(\xi + \mu^*, \Sigma^*) \neq c^L_\kappa(\Sigma) \) for almost every \( \xi \). From Lemma G.4, for almost every \( \xi \), either i) \( \hat{\eta}(\xi + \mu^*, \Sigma^*) \leq 0 \), or ii) \( \sigma^2_\eta(\xi + \mu^*, \Sigma^*) > 0 \) and \( v^{lo}(\xi + \mu^*, \Sigma^*) < v^{up}(\xi + \mu^*, \Sigma^*) \). Since \( \Sigma^* \) is positive definite and \( A \) has no non-zero rows, \( \Lambda \Sigma A' \) has strictly positive diagonal entries, and thus since \( \kappa < \alpha \), \( c^L_\kappa(\Sigma^*) > 0 \). Thus, if i) holds, then \( \hat{\eta} < c^L_\kappa \). On the other hand, note that conditional on \( v^{lo}(\xi + \mu^*, \Sigma^*) < v^{up}(\xi + \mu^*, \Sigma^*) \) and \( \sigma^2_\eta(\xi + \mu^*, \Sigma^*) \) has a truncated normal distribution with truncation points \( \hat{v}^{lo}(\xi + \mu^*, \Sigma^*) \) and \( \hat{v}^{up}(\xi + \mu^*, \Sigma^*) \) and untruncated variance \( \sigma^2_\hat{\eta} \). Thus, conditional on \( v^{lo} < v^{up} \) and \( \sigma^2_\hat{\eta} > 0 \), \( \hat{\eta} = c^L_\kappa(\Sigma^*) \) with probability 0.

\[ \Box \]

**Lemma G.1.** Let \( \mu^* \), \( \Sigma^* \), and \( \Omega \) be as defined in the proof to Proposition F.1, and assume \( \max_{\gamma \in \mathcal{V}(\Sigma^*)} \gamma' \mu^* \) is finite. Let \( N(\Sigma^*) \) be an open set containing \( \Sigma^* \). Define \( p : \Omega \times N(\Sigma^*) \to [0, 1] \) by

\[ p(Y, \Sigma) := P\left( \zeta < \hat{\eta}(Y, \Sigma) \mid \zeta \in [v^{lo}(Y, \Sigma), v^{up}(Y, \Sigma)], \zeta \sim \mathcal{N}(0, \sigma^2_\eta(Y, \Sigma)) \right). \]

Then \( p(Y, \Sigma) \) is continuous in both arguments at \((\xi + \mu^*, \Sigma^*)\) for almost every \( \xi \sim \mathcal{N}(0, \Lambda \Sigma A') \) and \( \Sigma^* \in S \) non-stochastic.

**Proof.** From Lemma G.4, for almost every \( \xi \), the functions \( \hat{\eta} \), \( v^{lo}, v^{up}, \sigma^2_\eta \) are continuous at \((\xi + \mu^*, \Sigma^*)\). Additionally, for almost every \( \xi \), either

1) There is a neighborhood of \((\xi + \mu^*, \Sigma^*)\) on which \( \sigma^2_\eta(Y, \Sigma) > 0 \) and \( v^{lo}(Y, \Sigma) < v^{up}(Y, \Sigma) \).
2) There is a neighborhood of \((\xi + \mu^*, \Sigma^*)\) on which \(\tilde{\eta}(Y, \Sigma) \leq 0, \sigma_\eta^2(Y, \Sigma) = 0\) and \(v^{lo}(Y, \Sigma) = -\infty, v^{up}(Y, \Sigma) = \infty\).

First, suppose 1) holds. Note that for \(v^{lo} < v^{up}\) and \(\sigma_\eta > 0\),
\[
\mathbb{P} \left( \zeta < \tilde{\eta} \mid \zeta \in [v^{lo}, v^{up}], \zeta \sim \mathcal{N}(0, \sigma_\eta^2) \right) = \frac{\Phi(\tilde{\eta}/\sigma_\eta) - \Phi(v^{lo}/\sigma_\eta)}{\Phi(v^{up}/\sigma_\eta) - \Phi(v^{lo}/\sigma_\eta)},
\]
which is clearly continuous in \(\tilde{\eta}, v^{lo}, v^{up}\), and \(\sigma_\eta\). The continuity of \(p(Y, \Sigma)\) then follows from the continuity of \(\tilde{\eta}, v^{lo}, v^{up}\), and \(\sigma_\eta\).

Next, suppose 2) holds. Note that
\[
\mathbb{P} \left( \zeta < \tilde{\eta} \mid \zeta \in [-\infty, \infty], \zeta \sim \mathcal{N}(0, 0) \right) = 1[\tilde{\eta} > 0].
\]
It then follows that when 2) holds, \(p(Y, \Sigma) = 0\) in a neighborhood of \((\xi + \mu^*, \Sigma^*)\), and thus is continuous.

\[\begin{proof}
\end{proof}\]

**Lemma G.2.** Let \(p(Y, \Sigma)\) be as defined in Lemma G.1, and suppose \(\max_{\gamma \in V(\Sigma^*)} \gamma' \mu^*\) is finite. Let \(\xi \sim \mathcal{N}(0, A \Sigma^* A')\). Then for any \(\alpha \in (0, 1)\), \(\mathbb{P}(\xi \mu^*, \Sigma^*) = 1 - \alpha = 0\).

\[\begin{proof}
\end{proof}\]

However, (17) implies that \(\tilde{\eta}(\xi + \mu^*, \Sigma^*)\) has a truncated normal distribution conditional on \(v^{lo}(\xi + \mu^*, \Sigma^*), v^{up}(\xi + \mu^*, \Sigma^*)\) and \(\sigma_\eta^2(\xi + \mu^*, \Sigma^*)\), with truncation points \(v^{lo}(\xi + \mu^*, \Sigma^*)\) and \(v^{up}(\xi + \mu^*, \Sigma^*)\) and (untruncated) variance \(\sigma_\eta^2(\xi + \mu^*, \Sigma^*)\), and hence is continuously distributed when \(v^{lo}(\xi + \mu^*, \Sigma^*) < v^{up}(\xi + \mu^*, \Sigma^*)\) and \(\sigma_\eta^2(\xi + \mu^*, \Sigma^*) > 0\). Thus, conditional on \(v^{lo}(\xi + \mu^*, \Sigma^*) < v^{up}(\xi + \mu^*, \Sigma^*)\) and \(\sigma_\eta^2(\xi + \mu^*, \Sigma^*) > 0\), \(\tilde{\eta}(\xi + \mu^*, \Sigma^*) = c_{1-\alpha}(\xi + \mu^*, \Sigma^*)\) with probability zero.

Additionally, observe that
\[
\mathbb{P} \left( \zeta < \tilde{\eta} \mid \zeta \in [-\infty, \infty], \zeta \sim \mathcal{N}(0, 0) \right) = 1[\tilde{\eta} > 0].
\]
Hence, whenever \(\tilde{\eta}(\xi + \mu^*, \Sigma^*) \leq 0, v^{lo}(\xi + \mu^*, \Sigma^*) = -\infty, v^{up}(\xi + \mu^*, \Sigma^*) = \infty\) and \(\sigma_\eta(\xi + \mu^*, \Sigma^*) = 0\), we have \(p(\xi + \mu^*, \Sigma^*) = 0 \neq 1 - \alpha\).

However, from Lemma G.4, with probability 1 either i) \(v^{lo}(\xi + \mu^*, \Sigma^*) < v^{up}(\xi + \mu^*, \Sigma^*)\) and \(\sigma_\eta^2(\xi + \mu^*, \Sigma^*) > 0\), or ii) \(\tilde{\eta}(\xi + \mu^*, \Sigma^*) \leq 0, v^{lo}(\xi + \mu^*, \Sigma^*) = -\infty, v^{up}(\xi + \mu^*, \Sigma^*) = \infty\) and \(\sigma_\eta(\xi + \mu^*, \Sigma^*) = 0\). The desired result then follows immediately.
Lemma G.3. Suppose Assumption 7 holds. Then for any \( x \) and \( \Sigma \in \mathbf{S} \), \( \lambda x' x \leq x' \Sigma x \leq \bar{\lambda} x' x \). Additionally, there exist constants \( \underline{c} > 0 \) and \( \bar{c} \) such that for all \( \Sigma \in \mathbf{S} \) and all \( j = 1, \ldots, J \), \( \underline{c} \leq c_j(\Sigma) \leq \bar{c} \), for \( c_j(\Sigma) \) as defined in Lemma F.1.

Proof. By the singular value decomposition, we can write \( \Sigma = U \Lambda U' \), where \( U \) is a unitary matrix \( (UU' = I) \) and \( \Lambda \) is the diagonal matrix with the eigenvalues of \( \Sigma \) on the diagonal. By Assumption 7, these eigenvalues are bounded between \( \underline{\lambda} > 0 \) and \( \bar{\lambda} \geq \lambda \). Thus, for any \( x \), we have \( x' \Sigma x = (U'x)' \Lambda(U'x) = \sum_i \lambda_i (U'x)^2_i \). It follows that \( x' \Sigma x \leq \sum_i \bar{\lambda} (U'x)^2_i = \bar{\lambda} x' UU' x = \bar{\lambda} x' x \). It can be shown analogously that \( x' \Sigma x \geq \underline{\lambda} x' x \). Now, recall that \( c_j(\Sigma) = (\gamma_j' \overline{\sigma}(\Sigma))^{-1} \), where \( \overline{\sigma}^2_j = A'_{(i,j)} \Sigma A_{(i,j)} \).

Let \( \bar{m}_A = \max_i A'_{(i,i)} A_{(i,i)} \) and \( m_A = \min_i A'_{(i,i)} A_{(i,i)} \), and note that both \( \bar{m} \) and \( m \) are strictly positive since \( A \) is assumed to have no all-zero rows. It then follows from the previous discussion that \( \bar{\sigma}_i \in [\sqrt{\bar{m}_A}, \sqrt{m_A}] := [\sigma_{lb}, \sigma_{ub}] \). Moreover, since \( \bar{\gamma}_j \geq 0 \) and \( \bar{\gamma}_j \neq 0 \) for all \( j \), we have that \( \gamma_j' \bar{\sigma} \geq \max \{ \gamma_j \} \bar{\sigma}_{lb} \geq \min_j \{ \max \{ \gamma_j \} \} \bar{\sigma}_{lb} > 0 \), where the last inequality uses the fact that the set \( \{ \gamma_j \}_{j=1}^{..J} \) is finite. Likewise, for \( K \) the dimension of \( \bar{\gamma}_j \), we have \( \gamma_j' \bar{\sigma} \leq K \max \{ \gamma_j \} \bar{\sigma}_{ub} \leq \max_j \max \{ \gamma_j \} \bar{\sigma}_{ub} < \infty \). We have thus shown that \( \gamma_j' \bar{\sigma}(\Sigma) \) is bounded between two positive finite values, and thus the same is true of its inverse, which suffices for the result.

Lemma G.4. Let \( \mu^*, \Sigma^* \), and \( \Omega \) be as defined in the proof to Proposition F.1, and assume \( \max_{\gamma \in V(\Sigma^*)} \gamma' \mu^* \) is finite. Let \( N(\Sigma^*) \) be an open set containing \( \Sigma^* \). Then \( \hat{\eta}(Y, \Sigma) \), \( \sigma^2_{\eta}(Y, \Sigma) \), \( v^{lo}(Y, \Sigma) \), \( v^{up}(Y, \Sigma) \) – when viewed as functions over \( \Omega \times N(\Sigma^*) \) – are continuous in \( (Y, \Sigma) \) at \( (\xi + \mu^*, \Sigma^*) \) for almost every \( \xi \sim \mathcal{N}(0, A\Sigma^* A') \). Additionally, for almost every \( \xi \), one of the following holds:

1) There is a neighborhood of \( (\xi + \mu^*, \Sigma^*) \) on which \( \sigma^2_{\eta}(Y, \Sigma) > 0 \) and \( v^{lo}(Y, \Sigma) < v^{up}(Y, \Sigma) \).

2) There is a neighborhood of \( (\xi + \mu^*, \Sigma^*) \) on which \( \hat{\eta}(Y, \Sigma) \leq 0 \), \( \sigma^2_{\eta}(Y, \Sigma) = 0 \) and \( v^{lo}(Y, \Sigma) = -\infty \), \( v^{up}(Y, \Sigma) = \infty \).

Proof. We first show that \( \hat{\eta}(Y, \Sigma) \) is continuous. Lemma F.1 implies that

\[
\hat{\eta}(Y, \Sigma) := \max_{\gamma \in V(\Sigma)} \gamma' Y = \max \{ c_1(\Sigma) \gamma_1^1 Y, \ldots, c_J(\Sigma) \gamma_J^J Y \},
\]

where the functions \( c_j(\Sigma) \) are continuous. We claim that each of the functions in the max above are continuous in \( (Y, \Sigma) \) at \( (\xi + \mu^*, \Sigma^*) \). If \( Y \) were finite-valued, then this would hold trivially. However, since some elements of \( Y \) may be equal to \( -\infty \), we additionally need to show that there is a neighborhood of \( \Sigma^* \) such that for all \( \Sigma \) in this neighborhood and all \( j \), the elements of \( c_j(\Sigma) \gamma_j \) do not change from 0 to non-zero or vice versa. However, by Lemma G.3, \( c_j(\Sigma^*) \geq c > 0 \) for all \( j \), and so for \( \Sigma \) sufficiently close to \( \Sigma^* \), \( c_j(\Sigma) > 0 \), and thus each element of \( c_j(\Sigma) \gamma_j \) has the same sign (0 or positive) as the corresponding element of \( \gamma_j \), as we desired to show.
Next, define $\hat{V}(Y, \Sigma) := \arg \max_{\gamma \in V(\Sigma)} \gamma/Y$. We claim that with probability 1, either $\hat{V}(\xi + \mu^*, \Sigma^*)$ is unique, or $\gamma_* A = \gamma_2 \bar{\Sigma} A$. Observe that since $\xi$ is finite with probability 1 and $\max_{\gamma \in V(\Sigma^*)} \gamma^\prime \mu^*$ is finite by assumption, it follows that $\max_{\gamma \in V(\Sigma^*)} \gamma^\prime (\xi + \mu^*)$ is finite with probability 1. Let $\gamma_1, \gamma_2 \in V(\Sigma^*)$. Note that $\gamma_1, \gamma_2 \in \hat{V}(\xi, \Sigma^*)$ only if $(\gamma_1 - \gamma_2)^\prime \xi = (\gamma_2 - \gamma_1)^\prime \mu^*$. Observe further that $(\gamma_1 - \gamma_2)^\prime \xi$ is normally distributed with variance $(\gamma_1 - \gamma_2)^\prime A \Sigma^* A^\prime (\gamma_1 - \gamma_2)^\prime$. Thus, $(\gamma_1 - \gamma_2)^\prime \xi$ is equal to a constant with positive probability only if $(\gamma_1 - \gamma_2)^\prime A \Sigma^* A^\prime (\gamma_1 - \gamma_2)^\prime = 0$. If $\Sigma^*$ is positive definite, $(\gamma_1 - \gamma_2)^\prime A \Sigma^* A^\prime (\gamma_1 - \gamma_2)^\prime = 0$ if $(\gamma_1 - \gamma_2)^\prime A = 0$. However, by Assumption 9, $(\gamma_1 - \gamma_2)^\prime A = 0$ only if $\gamma_1 A = \gamma_2 A = 0$. It follows that at most one of $\gamma_1$ and $\gamma_2$ are in $\hat{V}$ with probability 1, or $\gamma_1 A = \gamma_2 A = 0$. Since the set $V(\Sigma^*)$ is finite, it follows that either $\hat{V}(\xi + \mu^*, \Sigma^*)$ is unique, or all of its elements have $\gamma_* A = 0$, as needed.

Suppose first that every $\gamma_* \in \hat{V}(\xi + \mu^*, \Sigma^*)$ satisfies $\gamma_* A = 0$. Without loss of generality, assume that $\hat{V}(\xi + \mu^*) = \{c_1(\Sigma^*) \gamma_1, ..., c_J(\Sigma^*) \gamma_J\}$, where $1 \leq J \leq J$. We first claim that there is a neighborhood of $(\xi + \mu^*, \Sigma^*)$ on which $\max_{\gamma \in V(\Sigma^*)} \gamma^\prime Y = c_j(\Sigma) \gamma_j^\prime Y$ for some $j \leq J$. This is trivial if $J = 1$. If not, let $j \leq J_1$ and $i > J_1$. Since $c_j(\Sigma^*) \gamma_j^\prime (\xi + \mu^*) \in \hat{V}(\xi + \mu^*, \Sigma^*)$ and $c_i(\Sigma^*) \gamma_i^\prime (\xi + \mu^*) \notin \hat{V}(\xi + \mu^*, \Sigma^*)$, we must have $c_j(\Sigma^*) \gamma_j^\prime (\xi + \mu^*) > c_i(\Sigma^*) \gamma_i^\prime (\xi + \mu^*)$. We showed above that the functions on both sides of the inequality are continuous in $(Y, \Sigma)$ at $(\xi + \mu^*, \Sigma^*)$, and thus there exists a neighborhood of $(\xi + \mu^*, \Sigma^*)$ on which the inequality is preserved, and hence $\max_{\gamma \in V(\Sigma^*)} \gamma^\prime Y > c_1(\Sigma) \gamma_1^\prime (\xi + \Sigma)$. Additionally, since there are finitely many $i > J_1$, we can choose a neighborhood such that this holds simultaneously for all $i > J_1$, which implies that in this neighborhood $\hat{V}(Y, \Sigma) \subseteq \{c_1(\Sigma) \gamma_1, ..., c_J(\Sigma) \gamma_J\}$, as needed. It follows that $\sigma_{\gamma_j}^2(Y, \Sigma) = 0$ for all $(Y, \Sigma)$ in this neighborhood, since $\gamma_j A = 0$ for all $j \leq J_1$, which implies $\gamma_j A \Sigma^* A^\prime \gamma_j = 0$. Additionally, note that by definition, $\nu^0(Y, \Sigma) = -\infty$ and $\nu^{up}(Y, \Sigma) = \infty$ whenever $\sigma_{\gamma_j}^2(Y, \Sigma) = 0$. Thus, $\sigma_{\gamma_j}^2(Y, \Sigma)$, $\nu^0(Y, \Sigma)$, and $\nu^{up}(Y, \Sigma)$ are continuous at $(\xi + \mu^*, \Sigma^*)$.

To show that $\eta(Y, \Sigma) = \nu(Y, \Sigma) = 0$ in a neighborhood of $(\xi + \mu^*, \Sigma^*)$, observe that it is immediate from the definition of $\Omega$ that any $Y \in \Omega$ can be written as $A v - a_1 \cdot d + a_2 \mu^*$, for $v \in \mathbb{R}^K$ and $a_1, a_2 \geq 0$. For any $j \in \{1, ..., J\}$, $\gamma_j A = 0$, and thus $\gamma_j Y = -a_1 \gamma_j d + a_2 \gamma_j \mu^*$. However, since $\gamma_j \geq 0$ and $\mu^* \leq 0$, we have that $a_2 \gamma_j \mu^* \leq 0$. Likewise, since $\Delta$ is assumed to be non-empty, there exists some $\delta$ such that $A \delta - d \leq 0$. Since $\gamma_j A = 0$ and $\gamma_j \geq 0$, it follows that $\gamma_j (d - d) \leq 0$. Hence, $\gamma_j Y \leq 0$ for any $Y \in \Omega$, and thus, in a neighborhood of $\Sigma^*$ sufficiently small such that $c_j(\Sigma) \geq 0$, $c_j(\Sigma) \gamma_j Y \leq 0$. Since we’ve shown that in a neighborhood of $(\xi + \mu^*, \Sigma^*)$, $\bar{\eta}(Y, \Sigma) = c_j(\Sigma) \gamma_j Y$ for some $j$, it follows that $\eta(Y, \Sigma) = \nu(Y, \Sigma) = 0$ for $(Y, \Sigma)$ sufficiently close to $(\xi + \mu^*, \Sigma^*)$.

Next, suppose that $\hat{V}(\xi + \mu^*)$ has a single element $\gamma_* = c_j(\Sigma^*) \gamma_j^\prime (\xi + \mu^*)$ for some $j \in \{1, ..., J\}$ such that $\gamma_j A \neq 0$. Without loss of generality, suppose $j = 1$. We first show that $\hat{V}(Y, \Sigma) = c_1(\Sigma) \gamma_1$ in a neighborhood of $(\xi + \mu^*)$. Indeed, since $\hat{V}(\xi + \mu^*) = c_1(\Sigma^*) \gamma_1^\prime (\xi + \mu^*)$, for all $i > 1$, $c_1(\Sigma^*) \gamma_1^\prime (\xi + \mu^*) > c_i(\Sigma^*) \gamma_i^\prime (\xi + \mu^*)$. However, since we’ve shown the functions on both sides of this inequality to be continuous in $(Y, \Sigma)$ at $(\xi + \mu^*, \Sigma^*)$, there is a neighborhood of $(\xi + \mu^*, \Sigma^*)$ such that for all $i > 1$, $c_1(\Sigma) \gamma_1 Y > c_i(\Sigma) \gamma_i Y$, and hence $\hat{V}(Y, \Sigma) = c_1(\Sigma) \gamma_1$ in this neighborhood. It follows that in a neighborhood of $(\xi + \mu^*)$, $\sigma_{\gamma_j}^2(Y, \Sigma) = c_1(\Sigma) \gamma_1 A \Sigma^* A^\prime c_1(\Sigma) \gamma_1$, which is clearly
continuous in $\Sigma$. Additionally, by Lemma G.3, $c(\Sigma^*) \geq c > 0$, and so $\sigma^2_\eta \geq c^2 \eta A \Sigma^* A' \eta$, which is positive since $\eta A \neq 0$ and $\Sigma^*$ is positive definite. From the continuity of $\sigma^2_\eta$, it follows that there is a neighborhood of $(\xi + \mu^*, \Sigma^*)$ such that $\sigma^2_\eta(Y, \Sigma) > 0$.

Next, consider $v^{lo}(Y, \Sigma)$. Let $\gamma_\eta(\Sigma) = c_1(\Sigma) \bar{\gamma}_1$. For ease of notation, we will make the dependence of $\gamma_\eta$ on $\Sigma$ implicit where it is clear below. The results above imply that in a neighborhood of $(\xi + \mu^*, \Sigma^*)$, $v^{lo}(Y, \Sigma)$ is the minimum of the set

$$C(Y, \Sigma) = \{c : \max_{\gamma \in V(\Sigma)} \gamma'(S_{\gamma_\eta}(Y) + \frac{\Sigma \gamma_\eta}{\gamma_\eta \Sigma \gamma_\eta} c)\},$$

for

$$S_{\gamma_\eta}(Y, \Sigma) = \left(I - \frac{\Sigma \gamma_\eta}{\gamma_\eta \Sigma \gamma_\eta}\right) Y.$$

Rearranging terms, we see that

$$C = \{c : 0 = \max_{\gamma \in V(\Sigma)} a_{\gamma, \gamma_\eta}(Y) + b_{\gamma, \gamma_\eta} c\},$$

where $a_{\gamma, \gamma_\eta}(Y) := \gamma' S_{\gamma_\eta}(Y)$ and $b_{\gamma, \gamma_\eta} := \frac{\gamma' \Sigma \gamma_\eta}{\gamma_\eta \Sigma \gamma_\eta} - 1$. Note that $a_{\gamma_\eta, \gamma_\eta}(Y) = 0 = b_{\gamma_\eta, \gamma_\eta}$, so $0 \leq \max_{\gamma \in V(\Sigma)} a_{\gamma, \gamma_\eta}(Y) + b_{\gamma, \gamma_\eta} c$ for all $c$. Moreover, for $c = \gamma' Y$, the max is attained at $\gamma_\eta$ by construction. Hence, the set $C$ is non-empty.

Intuitively, if we plot $a_{\gamma, \gamma_\eta}(Y) + b_{\gamma, \gamma_\eta}$ as a function of $c$, then each $\gamma \in V(\Sigma)$ defines a line, and the set $C$ represents the values of $c$ for which 0 is the upper envelope of this set. It follows that the lower bound of $C$ is the maximal x-intercept of a line of the form $a_{\gamma, \gamma_\eta}(Y) + b_{\gamma, \gamma_\eta} c$ with $b_{\gamma, \gamma_\eta} < 0$. Hence,

$$v^{lo}(Y, \Sigma) = \max_{\{\gamma \in V(\Sigma) : (\gamma(Y), b_{\gamma, \gamma_\eta} < 0)\}} \frac{-a_{\gamma, \gamma_\eta}(Y)}{b_{\gamma, \gamma_\eta}}.$$

Recall that by Lemma F.1, $V(\Sigma) := \{\gamma_j(\Sigma), ..., \gamma_f(\Sigma)\}$, where $\gamma_j(\Sigma) := c_j(\Sigma) \bar{\gamma}_j$ and $c_j(\Sigma)$ is continuous. Additionally, we showed earlier in the proof that for all $j$, $c_j(\Sigma) \gamma_j^T Y$ is continuous in a neighborhood of $(\xi + \mu^*, \Sigma^*)$. It is then immediate from the definitions of $a_{\gamma, \gamma_\eta}(Y)$ and $b_{\gamma, \gamma_\eta}$ that for all $j$, $a_{\gamma_j(\Sigma), \gamma_\eta}(\Sigma)(Y)$ and $b_{\gamma_j(\Sigma), \gamma_\eta}(\Sigma)$ are continuous in $(\Sigma, \Sigma)$. Without loss of generality, suppose that for $2 \leq k \leq k_1$, $b_{\gamma_k(\Sigma^*), \gamma_\eta(\Sigma^*)} < 0$; for $k_1 < k \leq k_2$, $b_{\gamma_k(\Sigma^*), \gamma_\eta(\Sigma^*)} = 0$; and for $k > k_2$, $b_{\gamma_k(\Sigma^*), \gamma_\eta(\Sigma^*)} > 0$. From the continuity of $b_{\gamma_j(\Sigma), \gamma_\eta(\Sigma)}$, it is clear that in a neighborhood of $(\xi + \mu^*, \Sigma^*)$, $b_{\gamma_k(\Sigma), \gamma_\eta(\Sigma)} > 0$ for all $2 \leq k \leq k_1$ and $b_{\gamma_k(\Sigma), \gamma_\eta(\Sigma)} < 0$ for all $k > k_2$. Hence, in this neighborhood,

$$v^{lo}(Y, \Sigma) = \max_{\gamma_k(\Sigma): 2 \leq k \leq k_1} \max_{\gamma_k(\Sigma): 2 \leq k \leq k_1} \frac{-a_{\gamma_k(\Sigma), \gamma_\eta(\Sigma)}}{b_{\gamma_k(\Sigma), \gamma_\eta(\Sigma)}}, \max_{\gamma \in V(\Sigma)} \frac{-a_{\gamma, \gamma_\eta}(\Sigma)}{b_{\gamma, \gamma_\eta}(\Sigma)}$$

where
\[ V^0(\Sigma) := \{ \gamma_k(\Sigma) : k_1 < k < k_2, b_{\gamma_k(\Sigma)}, \gamma(\Sigma) < 0 \} \]

and we define the max of an empty set to be \(-\infty\). It is clear from the continuity of the functions \(a\) and \(b\) that the inner max on the left side of (70) is continuous. To show that \(v^0_{lo}\) is continuous at \((\xi + \mu^*, \Sigma^*)\), it suffices to show that for any sequence \((Y, \Sigma) \to (\xi + \mu^*, \Sigma^*)\), the max on the right hand side of (70) converges to \(-\infty\). To do this, observe that by construction, \(a_{\gamma, \gamma^*}(Y) + b_{\gamma, \gamma^*} \cdot \gamma^* Y = \gamma^* Y - \gamma Y\). Since for any \(k > 1\), \(\gamma^*(\Sigma^*)'(\xi + \mu^*) > \gamma_k(\Sigma^*)'(\xi + \mu^*)\), it follows that \(a_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*)'(\xi + \mu^*) + b_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*)' \cdot (\xi + \mu^*) < 0\). Additionally, \(b_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*)'(\xi + \mu^*) = 0\) for \(k \in (k_1, k_2]\), and so for such values of \(k\), \(a_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*)'(\xi + \mu^*) < 0\). However, this implies that for any sequence \((Y, \Sigma) \to (\xi + \mu^*)\) and \(k \in (k_1, k_2]\), we have \(-a_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*) Y\) approaching 0, and \(b_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*) Y\) approaching \(0\). For values of \((Y, \Sigma)\) where \(b_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*) > 0\), it follows that \(-a_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*) Y / b_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*)\) becomes arbitrarily negative, whereas for values of \((Y, \Sigma)\) where \(b_{\gamma_k(\Sigma^*)}, \gamma^*(\Sigma^*) \geq 0\), \(\gamma_k^* \) is not included in \(V^0\). It is then immediate that the max on the right hand side of (70) converges to \(-\infty\), which suffices to establish the continuity of \(v^0_{lo}\) at \((\xi + \mu^*, \Sigma^*)\). The continuity of \(v^0_{up}\) can be shown analogously.

To complete the proof, we now demonstrate that in a neighborhood of \((\xi + \mu^*)\), \(v^0_{lo}(Y, \Sigma) < v^0_{up}(Y, \Sigma)\) for almost every \(\xi\). Note that since we have shown \(v^0_{lo}\) and \(v^0_{up}\) to be continuous, it suffices to show that \(v^0_{lo}(\xi + \mu^*, \Sigma^*) < v^0_{up}(\xi + \mu^*, \Sigma^*)\). We showed above that for almost every \(\xi\), either \(\hat{V}(\xi + \mu^*)\) contains only elements such that \(\gamma^* A = 0\), or \(\hat{V}(\xi + \mu^*)\) has a unique element such that \(\gamma^* A \neq 0\). In the former case, we showed that \(v^0_{lo} = -\infty\) and \(v^0_{up} = \infty\). Suppose we are in the latter case. We showed that \(v^0_{lo}(\xi + \mu^*, \Sigma^*)\) is the x-intercept of a line of the form \(a + b \cdot c\), where \(b < 0\) and \(a + b \cdot \hat{\eta} < 0\). Hence, \(v^0_{lo}(\xi + \mu^*, \Sigma) < \hat{\eta}(\xi + \mu^*, \Sigma)\). However, by construction \(v^0_{lo} \leq \hat{\eta} \leq v^0_{up}\), and thus \(v^0_{lo} < \hat{\eta}\) implies \(v^0_{lo} < v^0_{up}\), which completes the proof.

**Lemma G.5.** For any vector \(v \in \mathbb{R}^T\),

\[
\tilde{A}_{(-,1)}(l'v) + \tilde{A}_{(-,1)}\Gamma_{(-,-1)}v = A \begin{pmatrix} 0 \\ I_T^\top \end{pmatrix} v.
\]

**Proof.** By definition,

\[
\tilde{A}_{(-,1)} = A \begin{pmatrix} 0 \\ I \end{pmatrix} \Gamma^{-1} I_{(-,1)}
\]

\[
\tilde{A}_{(-,-1)} = A \begin{pmatrix} 0 \\ I \end{pmatrix} \Gamma^{-1} I_{(-,-1)}
\]

\[
\Gamma_{(-,-)} = I_{(-,-)} \Gamma.
\]

Additionally, the first row of \(\Gamma\) is assumed to be \(l'\), so \(l' = I_{(1,\cdot)} \Gamma\). It follows that
\[ \hat{A}_{(1)} l'v = A \begin{pmatrix} 0 \\ I \end{pmatrix} \Gamma^{-1} I_{(1,1)} \Gamma v \]

\[ \hat{A}_{(-1)} \Gamma_{(-1)} v = A \begin{pmatrix} 0 \\ I \end{pmatrix} \Gamma^{-1} I_{(-1,1)} \Gamma v. \]

Noting that \( I_{(-1)} I_{(-1,1)} + I_{(1,1)} I_{(1)} = 1 \), the two equations in the previous display imply that

\[ \hat{A}_{(1)} (l'v) + \hat{A}_{(-1)} \Gamma_{(-1)} v = A \begin{pmatrix} 0 \\ I \end{pmatrix} \Gamma^{-1} \Gamma v = A \begin{pmatrix} 0 \\ I \end{pmatrix} v, \]

as needed. \( \square \)

### G.2 Proofs and auxiliary lemmas for uniform consistency results

#### Proof of Proposition F.2

*Proof.* By an argument analogous to that in the proof to Proposition F.1 for the conditional test, \( \psi_{*,\kappa,\alpha} (\hat{\beta}_n, A, d, \theta^b_P + x, 1 \tilde{\Sigma}_n) = \psi_{*,\kappa,\alpha} (\sqrt{n} \hat{\beta}_n, A, d, \sqrt{n} \theta^b_P + \sqrt{n} x, \hat{\Sigma}_n) \), so it suffices to show that

\[ \lim_{n \to \infty} \inf_{P \in \mathcal{P}} \mathbb{E}_P \left[ \psi_{*,\kappa,\alpha} (\sqrt{n} \hat{\beta}_n, A, \sqrt{n} d, \sqrt{n} \theta^b_P + \sqrt{n} x, \hat{\Sigma}_n) \right] = 1. \]

Towards contradiction, suppose this is false. Then there exists an increasing sequence of distributions \( P_m \) and sample sizes \( n_m \) such that

\[ \limsup_{m \to \infty} \mathbb{E}_{P_m} \left[ \psi_{*,\kappa,\alpha} (\sqrt{n_m} \hat{\beta}_{n_m}, A, \sqrt{n_m} d, \sqrt{n_m} \theta^b_{P_m} + \sqrt{n_m} x, \hat{\Sigma}_{n_m}) \right] \leq 1 - \omega, \quad (71) \]

for some \( \omega > 0 \). Since \( S \) is compact, we can extract a subsequence \( m_1 \) along which \( \Sigma_{P_{m_1}} \to \Sigma^* \in S \).

For ease of notation, without loss of generality we assume that this holds for the original sequence \( m \). Now, let

\[ \tilde{Y}_m := \sqrt{n_m} \left( A \hat{\beta}_{n_m} - d - \hat{A}_{(-1)} (\theta^b_P + x) \right) \]

\[ = \sqrt{n_m} A \left( \hat{\beta}_{n_m} - \beta_{P_m} \right) + \sqrt{n_m} \left( A \beta_{P_m} - d - \hat{A}_{(-1)} (\theta^b_P + x) \right), \quad (72) \]

and observe that

\[ \psi_{*,\kappa,\alpha} (\sqrt{n_m} \hat{\beta}_{n_m}, A, \sqrt{n_m} d, \sqrt{n_m} \theta^b_{P_m} + \sqrt{n_m} x, \hat{\Sigma}_{n_m}) = \psi_{*,\kappa,\alpha} (\tilde{Y}_m, X, A \hat{\Sigma}_{n_m} A'). \]
Now, from Lemma E.14, there exists a constant \(c > 0\) such that \(\eta(\beta_{P_m}, A, d, \theta_{P_m}^{ub} + x, \Sigma^*) \geq c \cdot x\) for \(\eta(\cdot)\) defined in (45). Reformulating (45) in terms of its dual, and noting that the dual vertices are the same as in the dual problem for \(\hat{\eta}\), we see that there is a dual vertex \(\gamma_j(\Sigma^*) \in V(\Sigma^*)\) such that \(\gamma_j(\Sigma^*)' \left( A\beta_{P_m} - d - \hat{A}_{(-1)}(\theta_{P_m}^{ub} + x) \right) \geq c \cdot x\). From Lemma F.1, \(\gamma_j(\Sigma^*) = c_j(\Sigma^*)\gamma_j\), and there is a vertex of \(V(\hat{\Sigma}_{n_m})\) of the form \(\gamma_j(\hat{\Sigma}_{n_m}) = c_j(\hat{\Sigma}_{n_m})\gamma_j\), where the function \(c_j(\cdot)\) is continuous. Since \(\hat{\Sigma}_{n_m} \to_p \Sigma^*\), it follows that \(\gamma_j(\hat{\Sigma}_{n_m}) \to_p \gamma_j(\Sigma^*)\), and hence \(\gamma_j(\hat{\Sigma}_{n_m})' \left( A\beta_{P_m} - d - \hat{A}_{(-1)}(\theta_{P_m}^{ub} + x) \right) \to_p c \cdot x > 0\). It is then clear from (72) that \(\gamma_j(\hat{\Sigma}_{n_m})\hat{Y}_m \to_p \infty\), since the inner product of \(\gamma_j(\hat{\Sigma}_{n_m})\) with the first term of (72) converges in distribution to a normal distribution with mean 0 and finite variance by Assumption 8 and Slutsky’s lemma, and the second term converges in probability to \(\infty\). Since \(\gamma_j(\hat{\Sigma}_{n_m})\hat{Y}_m\) is feasible in the dual problem for \(\hat{\eta}_{n_m}\), it follows that \(\hat{\eta}_{n_m} \to_p \infty\).

To complete the proof, recall from Section 4 that the unmodified hybrid test rejects whenever \(\hat{\eta}\) exceeds the \(1 - \kappa\) level least-favorable critical value, \(c_{\kappa}^{LF}(\hat{\Sigma}_{n_m})\). Thus, the modified hybrid test rejects whenever \(\hat{\eta}\) exceeds \(\max\{c_{\kappa}^{LF}(\hat{\Sigma}_{n_m}), -C\}\). We argue in the proof to Corollary F.1 that \(c_{\kappa}^{LF}\) is continuous, and hence \(c_{\kappa}^{LF}(\hat{\Sigma}_{n_m}) \to_p c_{\kappa}^{LF}(\Sigma^*) < \infty\). It follows that

\[
\limsup_{m \to \infty} \psi_{\kappa,\alpha}(\hat{Y}_m, X, A\hat{\Sigma}_{n_m}A') = 1,
\]

which contradicts (71).

\[\Box\]

**Proof of Proposition F.3**

*Proof.* The first half of the proof proceeds analogously to the proof of Proposition F.2. Since \(\psi_{*,\alpha}(\hat{\beta}_n, A, d, \theta_{P}^{ub} + x, \frac{1}{n}\hat{\Sigma}_n) = \psi_{*,\alpha}(\sqrt{n}\hat{\beta}_n, A, d, \sqrt{n}\theta_{P}^{ub} + \sqrt{n}x, \hat{\Sigma}_n)\), it suffices to show that

\[
\lim_{n \to \infty} \inf_{P \in \mathcal{P}} \mathbb{E}_P \left[ \psi_{*,\alpha}(\sqrt{n}\hat{\beta}_n, A, \sqrt{n}d, \sqrt{n}\theta_{P}^{ub} + \sqrt{n}x, \hat{\Sigma}_n) \right] = 1.
\]

Towards contradiction, suppose this is false. Then there exists an increasing sequence of distributions \(P_m\) and sample sizes \(n_m\) such that

\[
\limsup_{m \to \infty} \mathbb{E}_{P_m} \left[ \psi_{*,\alpha}(\sqrt{n_m}\hat{\beta}_{n_m}, A, \sqrt{n_m}d, \sqrt{n_m}\theta_{P_m}^{ub} + \sqrt{n_m}x, \hat{\Sigma}_{n_m}) \right] \leq 1 - \omega, \quad (73)
\]

for some \(\omega > 0\). Since \(V\) is compact, we can extract a subsequence \(m_1\) along which \(V_{P_{m_1}} \to V^* = \begin{pmatrix} \Sigma^* & V_{\Sigma}^* \\ V_{\Sigma}^* & V_{\Sigma}^* \end{pmatrix} \in V\).

Now, as in the proof to Proposition F.2, we can write

\[
\psi_{*,\alpha}(\sqrt{n_m}\hat{\beta}_{n_m}, A, \sqrt{n_m}d, \sqrt{n_m}\theta_{P_m}^{ub} + \sqrt{n_m}x, \hat{\Sigma}_{n_m}) = \psi_{*,\alpha}(\hat{Y}_m, X, A\hat{\Sigma}_{n_m}A'),
\]
where
\[
\tilde{Y}_m = \sqrt{\frac{m}{n}} A \left( \beta_n - \beta_{P_m} \right) + \sqrt{\frac{m}{n}} \left( A \beta_{P_m} - d - \bar{A} \right) (\theta_{P_m} + x) .
\]

(74)

By the same argument as in the proof to Proposition F.2, we have that \( \hat{\eta}_{m_n} \to_{p} \infty \). It follows
\[
\mathbb{P} \left( \hat{\eta}_{m_n} < -C \right) \to 0,
\]
sot the modified test agrees with the unmodified test with probability approaching 1. For simplicity, we therefore consider the unmodified test for remainder of the proof.

Now, suppose \( C > \max\{0, z_{1-\alpha}\} \). We showed in the proof to Lemma E.19 that if \( \hat{\eta}(\tilde{Y}, \tilde{\Sigma}) > C \), then \( \psi_{\alpha}(\tilde{Y}, \tilde{\Sigma}) = 1 \) unless \( \sigma_{\gamma} := \sqrt{\gamma' \tilde{\Sigma} \gamma} > 0 \) and \( \frac{1}{\sigma_{\gamma}}(\hat{\eta} - v^{lo}) < \zeta(C) \), where \( \gamma \) is an optimal solution to the dual problem and \( \zeta(\cdot) \) is a function such that \( \zeta(C) \to 0 \) as \( C \to \infty \). Additionally, by Lemma G.6, there exists some vertex \( \gamma \) such that
\[
\frac{1}{\sigma_{\gamma}}(\hat{\eta} - v^{lo}) = \kappa(\gamma, \gamma) \left( \gamma' \tilde{Y} - \gamma' \tilde{Y} \right),
\]
where \( \kappa(\gamma, \gamma) = \frac{\sqrt{\gamma' \tilde{\Sigma} \gamma}}{\gamma' \tilde{\Sigma} \gamma} \left( \gamma' \tilde{Y} - \gamma' \tilde{Y} \right) > 0 \).

To complete the proof, we will show that we can extract a subsequence of \( m \), indexed by \( q \), along with a constant \( C \) such that
\[
\limsup_{q \to \infty} \mathbb{P}_{P_q} \left( \{ \hat{\eta}_{n_q} < C \} \lor \left\{ \sigma_{n_q} > 0 \right\} \right) \leq \omega/2.
\]

This implies a contradiction of (73), since the event in the probability in the previous display is a necessary condition for the conditional test to fail to reject. Further, since we’ve shown that \( \hat{\eta}_{m_n} \to_{p} \infty \), it suffices to construct a subsequence such that
\[
\limsup_{q \to \infty} \mathbb{P}_{P_q} \left( \{ \sigma_{n_q} > 0 \} \right) \leq \omega/2.
\]

(75)

Now, recall from Lemma F.1 that we can write \( V(\tilde{\Sigma}) = \{ c_1(\tilde{\Sigma}) \bar{\gamma}_1, ..., c_J(\tilde{\Sigma}) \bar{\gamma}_J \} \) for positive continuous functions \( c_j \) and distinct non-zero vectors \( \bar{\gamma}_j \). For notational convenience, let \( c_{i,m} = c_i(\tilde{\Sigma}_{m_n}), c_{i,*} = c_i(\Sigma^*), \gamma_{i,m} = c_{i,m} \bar{\gamma}_i \), and \( \gamma_{i,*} = c_{i,*} \bar{\gamma}_i \). Likewise, for a pair \( (i, j) \) let \( \kappa_{ij,m} = \kappa(\gamma_{i,m}, \gamma_{j,m}) \) and \( \kappa_{ij,*} = \kappa(\gamma_{i,*}, \gamma_{j,*}) \). Assumption 8 implies that \( \tilde{\Sigma}_{n_m} \to_p \Sigma^* \). By the continuous mapping theorem, we therefore have \( c_{i,m} \to_p c_{i,*}, \gamma_{i,m} \to_p \gamma_{i,*}, \) and \( \kappa_{ij,m} \to_p \kappa_{ij,*} \).

Note that if \( \gamma_{i,m} \) is optimal and \( \bar{\gamma}_i A = 0 \), then \( \sigma_{n,m} = (c_{i,m} \bar{\gamma}_i)' A \tilde{\Sigma}_{n_m} A' (c_{i,m} \bar{\gamma}_i) = 0 \). Thus, we can only have \( \sigma_{n,m} > 0 \) if the optimal vertex corresponds with an index \( i \) such that \( \bar{\gamma}_i A \neq 0 \). To establish (75), it therefore suffices to extract a subsequence \( q \) such that for any pair \( (i, j) \) with \( i \neq j \) and \( \bar{\gamma}_i A \neq 0 \), either
\[
\lim_{q \to +\infty} \mathbb{P}_q \left( \hat{\eta}_n = \gamma_{i,q} Y_m \right) = 0, \quad \text{OR} \\
\limsup_{q \to +\infty} \mathbb{P}_q \left( \{ \hat{\eta}_n = \gamma_{i,q} Y_m \} \cap \{ \kappa_{ij,q}(\gamma_{i,q} - \gamma_{j,q})'Y_q < \zeta(C) \} \right) \leq \omega/(2m),
\]
where \( m \) is the number of such pairs \((i, j)\).

Consider any such pair \((i, j)\). First, we claim that \( \bar{\gamma}'_i \lambda_m \leq \bar{\gamma}'_i \bar{A}_{(1)} x \). To show this, note that since \( \theta_{P_m}^{ab} \in S(\Delta, \beta_{P_m}) \), \( \exists \tilde{\tau} \in \mathbb{R}^{T-1} \) such that \( \lambda_m + \bar{A}_{(1)} x = A \beta_{nm} - d - \bar{A}_{(1)} \theta_{P_m}^{ab} \tilde{\tau} \leq 0 \). By construction (see the proof to Lemma F.1) \( \bar{\gamma}'_i \bar{A}_{(1)} = 0 \) and \( \bar{\gamma}_i > 0 \), and hence \( \bar{\gamma}'_i (\lambda_m + \bar{A}_{(1)} x) \leq 0 \), which implies \( \bar{\gamma}'_i \lambda_m \leq -\bar{\gamma}'_i \bar{A}_{(1)} x \), as desired.

Since \( \bar{\gamma}'_i \lambda_m \) is bounded above, it follows that either i) \( \bar{\gamma}'_i \lambda_m \to -\infty \), or ii) there exists a subsequence \( m_1 \) such that \( \bar{\gamma}'_i \lambda_{m_1} \to \mu_1 \in \mathbb{R} \). If i) holds, then it is clear from (74) that \( \bar{\gamma}'_{i,m} Y_m \to_{p} -\infty \), since the inner product of \( \gamma_{i,m} \) with the first term in (74) converges in distribution to a normal distribution with mean 0 and finite variance by Assumption 10 and Slutsky’s lemma, and the second term converges in probability to \(-\infty \). Since \( \hat{\eta}_{m_1} \to_{p} \infty \), it follows that \( \mathbb{P} \left( \hat{\eta}_{m_1} = \bar{\gamma}'_{i,m_1} Y_m \right) \to 0 \), and \( \gamma_{i,m} \) is optimal with vanishing probability. Now, suppose ii) holds and consider the sequence \( m_1 \). By an analogous argument for \( \gamma_{j,m} \), we can show that either ii.a) \( \gamma'_{j,m_1} Y_m \to_{p} -\infty \), or ii.b) there exists a further subsequence \( m_2 \) such that \( \bar{\gamma}'_{j,m_2} \lambda_{m_2} \to \mu_2 \in \mathbb{R} \). If ii.a) holds, then it is immediate that for any \( \zeta > 0 \), \( \mathbb{P} \left( \{ \hat{\eta}_{m_2} = \gamma_{i,m_2} Y_{m_2} \} \cap \{ \kappa_{ij,m_2}(\gamma_{i,m_2} - \gamma_{j,m_2})'Y_{m_2} \in [-\zeta, \zeta] \} \right) \to 0 \), since \( \hat{\eta}_{m_2} \to_{p} \infty \), \( \gamma'_{j,m_1} Y_m \to_{p} -\infty \), and \( \kappa_{ij,m_2} \to \kappa'_{ij} > 0 \). Now, suppose iib) holds. Since \( \sqrt{m_2}(\gamma_{i}^* - \gamma_{j}^*)'\lambda_{m_2} \) is non-stochastic, we can choose a subsequence \( m_3 \) such that \( \sqrt{m_3}(\gamma_{i}^* - \gamma_{j}^*)'\lambda_{m_3} \to \mu_3 \in \mathbb{R} \cup \{ \pm \infty \} \). Then

\[
(\gamma_{i,m_3} - \gamma_{j,m_3})'Y_{m_3} = (\gamma_{i,m_3} - \gamma_{j,m_3})'\sqrt{m_3} A(\hat{\beta}_{m_3} - \beta_{m_3}) =: Z_1 \\
+ \sqrt{m_3}(\gamma_{i,m_3} - \gamma_{j,m_3})'\lambda_{m_3} =: Z_2 \\
+ \sqrt{m_3}(\gamma_{i}^* - \gamma_{j}^*)'\lambda_{m_3} =: Z_3
\]

By assumption 10 along with Slutsky’s lemma, \( Z_1 \to_d (\gamma_{i}^* - \gamma_{j}^*) A \xi_{\beta} \), for \( \xi_{\beta} \sim \mathcal{N}(0, \Sigma^*) \). Next, note that we write \( Z_2 = \sqrt{n}(c_{i}(\bar{\Sigma}_{m_3}) - c_{i}(\Sigma^*))\tilde{v}'_{i} \lambda_{m_3} \). Since \( c_{i} \) is continuous, Assumption 10 along with the delta method imply that \( \sqrt{n}(c_{i}(\bar{\Sigma}_{m_3}) - c_{i}(\Sigma^*)) \to_d G_{i}' \xi_{\Sigma} \), where \( G_{i} = D_{\text{vec}(\Sigma)}c_{i}(\Sigma^*) \) is the gradient of \( c_{i} \) at \( \Sigma^* \), and \( \xi_{\Sigma} \sim \mathcal{N}(0, V_{\Sigma}) \). Since \( \tilde{v}'_{i} \lambda_{m_3} \to \mu_{1} \), by Slutsky’s lemma, we have \( Z_2 \to_d \mu_{1} G_{i}' \xi_{\Sigma} \). By an analogous argument, we have that \( Z_3 \to_d \mu_{2} G_{i}' \xi_{\Sigma} \). Finally, recall that \( Z_4 \to \mu_{3} \) by construction, and \( \kappa_{ij,m_3} \to \kappa'_{ij} > 0 \). Combining these results, along with the fact that these convergences hold jointly by Assumption 10, we have that

\[ S\text{-70} \]
we show in the proof to Lemma G.11 that

\[ \eta_{ij} \] is positive definately by Assumption 11. Further, Assumption 9 implies that \((\gamma_i^* - \gamma_j^*)'A\neq 0\), and thus \(\kappa_{ij}^*(\gamma_i^* - \gamma_j^*)'A\xi_{\beta}^*\) has positive variance conditional on \(\xi_{\Sigma}^*\). That the unconditional variance of \(\eta_{ij}\) is positive then follows from the law of total variance. Let \(\sigma_{ij}^2\) denote the unconditional variance of \(\eta_{ij}\). We then see that for any \(\zeta > 0\), \(\mathbb{P}(\xi_{ij} \in [-\zeta, \zeta]) \leq \Phi(\zeta/\sigma_{ij}) - \Phi(-\zeta/\sigma_{ij})\), since the normal distribution is single-peaked and symmetric about its mean, so the maximal probability that a normal variable falls in an interval of length \(2\zeta\) occurs when the interval is centered around the mean. Since \(\zeta(C) \to 0\) as \(C \to \infty\), we can choose \(C\) sufficiently large such \(\Phi(\zeta/\sigma_{ij}) - \Phi(-\zeta/\sigma_{ij}) < \omega/(2m)\). Hence,

\[
\limsup_{m_3 \to \infty} \mathbb{P}(|\eta_{ij,n_3}(\gamma_i,m_3 - \gamma_j,m_3)'Y_{m_3}| < \zeta(C)) \leq \omega/(2m).
\]

We have thus established that we can find a subsequence along which (76) or (77) holds for a single pair \((i,j)\). However, since there are finitely many such pairs \((i,j)\), we can use analogous arguments to further refine our subsequence and constant \(C\) such that this holds for all pairs \((i,j)\).

**Lemma G.6.** Let \(\hat{\eta}(Y, \Sigma)\) be as defined in the proof to Proposition F.1, and \(\gamma_*\) an optimal solution to the dual problem for \(\hat{\eta}(Y, \Sigma)\). Then, if \(v^{lo}(Y, \Sigma)\) is finite,

\[
\hat{\eta} - v^{lo} = \frac{\gamma_*'\Sigma \gamma_*}{\gamma_*' \Sigma \gamma_* - \gamma' \Sigma \gamma_*} \left( \gamma_* \tilde{Y} - \gamma' \tilde{Y} \right),
\]

for some vertex \(\gamma \in V(\Sigma)\) such that \(\frac{\gamma_*' \Sigma \gamma_*}{\gamma_*' \Sigma \gamma_* - \gamma' \Sigma \gamma_*} > 0\).

**Proof.** We show in the proof to Lemma G.11 that

\[
v^{lo} = \min_{\gamma \in V(\Sigma): b_{\gamma, \gamma_*} < 0} \left( -\frac{a_{\gamma, \gamma_*}(Y)}{b_{\gamma, \gamma_*}} \right),
\]

where

\[
b_{\gamma, \gamma_*} = 1 - \frac{\gamma_*' \Sigma \gamma_*}{\gamma_*' \Sigma \gamma_* - \gamma' \Sigma \gamma_*},
\]

\[
a_{\gamma, \gamma_*}(Y) = \gamma' (I - \frac{\Sigma \gamma_*}{\gamma_*' \Sigma \gamma_*}) Y.
\]
Noting that \( \hat{\eta} = \gamma'_b Y \), the result then follows from applying the expressions above and cancelling like terms.

\[ \text{G.3 Proofs and auxiliary lemmas for uniform local asymptotic power results} \]

\[ \text{Proof of Proposition F.4} \]

\[ \text{Proof.}\] Let \( \hat{\gamma}_1, \ldots, \hat{\gamma}_j \) be as defined in Lemma F.1. By Lemma G.16, there exists a value \( C^* \in \mathbb{R} \) such that for any \( \Sigma \in \mathcal{S} \) and any \( j \) such that \( \hat{\gamma}'_j A \neq 0 \),

\[
\Phi \left( \frac{\hat{\eta}}{\sqrt{\gamma'_j \Sigma' A \gamma'_j \Sigma}} \right) > 1 - \alpha
\]

only if \( \hat{\eta} > -C^* \). We suppose throughout the proof that \( -C \leq -C^* \).

Towards contraction, suppose that the proposition is false. Then there exists a sequence of distributions \( P_m \in \mathcal{P}_e \) and an increasing sequence of sample sizes \( n_m \) such that

\[
\liminf_{n \to \infty} \left| \mathbb{E}_{P_m} \left[ \psi^C_{*,\alpha}(\hat{\gamma}'_{n_m}, A, d, \theta_{P_m}^{ub} + \frac{1}{\sqrt{n_m}} x, \frac{1}{n_m} \hat{\Sigma}_{n_m}) - \rho^*(P_m) \right] \right| \geq \omega
\]

for some \( \omega > 0 \). We showed in the proof to Proposition F.1 that \( \psi^C_{*,\alpha} \) is invariant to scale, so this is equivalent to

\[
\liminf_{n \to \infty} \left| \mathbb{E}_{P_m} \left[ \psi^C_{*,\alpha}(\sqrt{n_m} \hat{\gamma}'_{n_m}, A, \sqrt{n_m} d, \sqrt{n_m} \theta_{P_m}^{ub} + x, \hat{\Sigma}_{n_m}) - \rho^*(P_m) \right] \right| \geq \omega
\]

Define

\[
Y_m = \sqrt{n_m} (A \hat{\gamma}'_{n_m} - d - \hat{A}(\cdot, -1)(\theta_{P_m}^{ub} + x))
\]

and \( X := \hat{A}(\cdot, -1) \). Then

\[
\psi^C_{*,\alpha}(\sqrt{n_m} \hat{\gamma}'_{n_m}, A, \sqrt{n_m} d, \sqrt{n_m} \theta_{P_m}^{ub} + x, \hat{\Sigma}_{n_m}) = \psi^C_{*,\alpha}(Y_m, X, A \hat{\Sigma}_{n_m} A')
\]

For notational convenience, define \( \tau_m := \tau_{P_m} \); define \( \delta_m, \delta_m^{**} \) and \( \Sigma_m \) analogously. Let \( \hat{Y}_m := Y_m - \hat{A}(\cdot, -1) \Gamma(\cdot, -1) \sqrt{n_m} (\tau_{P_m} - \delta_{P_m,\text{post}} + \delta_{P_m,\text{post}}^{**}) \). By Lemma 16 in ARP, \( \psi^C_{*,\alpha}(Y_m, X, A \hat{\Sigma}_{n_m} A') = \psi^C_{*,\alpha}(\hat{Y}_m, X, A \hat{\Sigma}_{n_m} A') \). Additionally, recall from the proof of Lemma E.8 that \( \theta_{P_m}^{ub} = l' (\tau_{P} + \delta_{P,\text{post}} - \delta_{P,\text{post}}^{**}) \). From this, we see that
\[ \tilde{Y}_m = \sqrt{n_m} \left( A\hat{\beta}_{n_m} - d - \tilde{A}_{(\cdot,1)} \hat{\delta}_{P_m} - \tilde{A}_{(\cdot,-1)} \Gamma(\cdot,+) (\tau_{P_m} - \delta_{P_m,post}) \right) - \tilde{A}_{(\cdot,1)} x \]
\[ = \sqrt{n_m} \left( A\hat{\beta}_{n_m} - d - \tilde{A}_{(\cdot,1)} \Gamma(\cdot,+) (\tau_{P_m} + \delta_{P_m,post} - \delta_{P_m,post}^*) - \tilde{A}_{(\cdot,1)} \Gamma(\cdot,+) (\tau_{P_m} - \delta_{P_m,post} + \delta_{P_m,post}^*) \right) - \tilde{A}_{(\cdot,1)} x \]
\[ = \sqrt{n_m} \left( A\hat{\beta}_{n_m} - d - A \begin{pmatrix} 0 \\ I \end{pmatrix} (\tau_{P_m} + \delta_{P_m,post} - \delta_{P_m,post}^*) \right) - \tilde{A}_{(\cdot,1)} x, \]

where the last line follows from Lemma G.5. Additionally, note that by construction, \( \delta_{P_m,pre} = \delta_{P_m,post}^* \). Thus, \( \delta_{P_m} - \delta_{P_m,post}^* = \begin{pmatrix} 0 \\ I \end{pmatrix} (\delta_{P_m,post} - \delta_{P_m,post}^*) \). It follows that

\[ \tilde{Y}_m = \sqrt{n_m} A \left( \hat{\beta}_{n_m} - \delta_{P_m} - \begin{pmatrix} 0 \\ \tau_{P_m} \end{pmatrix} \right) + \sqrt{n_m} (A\delta_{P_m}^* - d) - \tilde{A}_{(\cdot,1)} x. \quad (80) \]

Now, since \( P_m \in \mathcal{P}_\epsilon \), by definition there exists an index \( B_m \) such that

\[ A_{(B_m,\cdot)} \delta_{P_m}^* - d_{B_m} = 0 \]
\[ A_{(-B_m,\cdot)} \delta_{P_m}^* - d_{-B_m} < \epsilon, \]

and \( A_{B_m,post} \) has rank \( |B_m| \). Since there are finitely many possible subindices of the rows of \( A \), we can choose a subsequence \( m_1 \) such that \( B_{m_1} = B \) for some index \( B \) such that \( A_{B,post} \) has rank \(|B|\). Additionally, since \( S \) is compact, we can choose a further subsequence \( m_2 \) along which \( \Sigma_{P_{m_2}} \to \Sigma^* \) for some \( \Sigma^* \in \mathcal{S} \). To avoid notational clutter, we will assume that these convergences hold for the original sequence \((m, n_m)\). Additionally, without loss of generality, we will assume that \( B \) corresponds with the first \(|B| \) rows of \( A \). It follows that

\[ \sqrt{n_m} (A\delta_{P_m}^* - d) - \tilde{A}_{(\cdot,1)} x = \left( \begin{pmatrix} -\tilde{A}_{(B,1)} x \\ \sqrt{n_m} (A_{(-B,\cdot)} \delta_{P_m}^* - d_{-B}) - \tilde{A}_{(-B,1)} x \end{pmatrix}, \right) \]
\[ \leq \left( \begin{pmatrix} -\tilde{A}_{(B,1)} x \\ -\sqrt{n_m} \epsilon - \tilde{A}_{(-B,1)} x \end{pmatrix} \right), \]

from which it is apparent that

\[ \sqrt{n_m} (A\delta_{P_m}^* - d) - \tilde{A}_{(\cdot,1)} x \to \left( \begin{pmatrix} -\tilde{A}_{(B,1)} x \\ -\infty \end{pmatrix} \right) =: \bar{\mu} \]

as \( m \to \infty \). Now, equation (80) along with Assumptions 6 and 8 and the continuous mapping theorem
imply that

\[(\tilde{Y}_m, \tilde{\Sigma}_m) \to_d (\xi + \tilde{\mu}, \Sigma^*),\]

for \(\xi \sim \mathcal{N}(0, A\Sigma^* A')\).

Now, as in the proof to Proposition F.1, note that the (unmodified) conditional test rejects iff \(p(Y, \Sigma) > 1 - \alpha\) for

\[p(Y, \Sigma) := \mathbb{P}\left( \zeta < \tilde{\eta}(Y, \Sigma) \mid \zeta \in [v^{lo}(Y, \Sigma), v^{up}(Y, \Sigma)] , \zeta \sim \mathcal{N}\left(0, \sigma^2_\eta(Y, \Sigma)\right) \right) > 1 - \alpha.\]

It follows that the modified conditional test rejects iff \(\tilde{p}(Y, \Sigma) := p(Y, \Sigma) \cdot 1[\tilde{\eta}(Y, \Sigma) \geq -C] > 1 - \alpha\).

Thus, (79) implies that

\[\liminf_{n \to \infty} \left| \mathbb{P}_{P_m}\left( \tilde{p}(\tilde{Y}_m, \tilde{\Sigma}_m) > 1 - \alpha \right) - p^*(P_m) \right| \geq \omega.\]

Additionally, Proposition 5.2 implies that for all \(m\), \(p^*(P_m) = \Phi(c^* x - z_{1-\alpha})\), where \(c^* = -\gamma_B' A_{(B,1)} / \sigma_B\); for \(\sigma_B = \sqrt{\gamma_B' A_{(B,1)} \Sigma A_{(B,1)'}} \gamma_B\) and \(\gamma_B\) the unique vector such that \(\gamma_B' A_{(B,1)} = 0, \gamma_B \geq 0, \|\gamma_B\| = 1\).

Thus,

\[\liminf_{n \to \infty} \left| \mathbb{P}_{P_m}\left( \tilde{p}(\tilde{Y}_m, \tilde{\Sigma}_m) > 1 - \alpha \right) - \Phi(c^* x - z_{1-\alpha}) \right| \geq \omega. \quad (81)\]

However, Lemma G.14 gives that \(\tilde{p}(Y, \Sigma)\) is continuous at \((\xi + \tilde{\mu}, \Sigma^*)\) for almost every \(\xi \sim \mathcal{N}(0, A\Sigma^* A')\), and so from the continuous mapping theorem,

\[\tilde{p}(\tilde{Y}_m, \tilde{\Sigma}_m) \to_d \tilde{p}(\xi + \tilde{\mu}, \Sigma^*).\]

Additionally, Lemma G.15 gives that the distribution of \(\tilde{p}(\xi + \tilde{\mu}, \Sigma^*)\) is continuous at \(1 - \alpha\), and thus

\[\mathbb{P}_{P_m}\left( \tilde{p}(\tilde{Y}_m, \tilde{\Sigma}_m) > 1 - \alpha \right) \to \mathbb{P}\left( \tilde{p}(\xi + \tilde{\mu}, \Sigma^*) > 1 - \alpha \right).\]

Lemma G.12 implies that with probability 1,

\[p(\xi + \tilde{\mu}, \Sigma^*) = \Phi\left( \frac{\gamma_j(\Sigma^*)' (\xi + \tilde{\mu})}{\sqrt{\gamma_j(\Sigma^*)' A \Sigma^* A' \gamma_j(\Sigma^*)}} \right),\]

where \(\gamma_j(\Sigma^*) = c_j(\Sigma) \tilde{\gamma}_j\) for \(\tilde{\gamma}_j\) the unique element of \(\{\tilde{\gamma}_1, ..., \tilde{\gamma}_J\}\) such that \(\tilde{\gamma}_{j,-B} = 0\). Additionally, Lemma G.9 gives that with probability 1, \(\tilde{\eta}(\xi + \tilde{\mu}, \Sigma^*) = \gamma_j(\Sigma^*)' (\xi + \tilde{\mu})\). Since \(-C \leq -C^*, \Phi\left( \frac{\tilde{\eta}}{\sqrt{\gamma_j(\Sigma^*)' A \Sigma^* A' \gamma_j(\Sigma^*)}} \right) > 1 - \alpha\) only if \(\tilde{\eta} > -C\), from which we see that \(\mathbb{P}(\tilde{p}(\xi + \tilde{\mu}, \Sigma^*) > 1 - \alpha) = S-74\)
\[ \mathbb{P} \left( p(\xi + \hat{\mu}, \Sigma^*) > 1 - \alpha \right) \text{.} \] It follows from the expression for \( p(\xi + \hat{\mu}, \Sigma^*) \) in the previous display that with probability 1, \( p(\xi + \hat{\mu}, \Sigma^*) > 1 - \alpha \) iff

\[ \frac{\gamma_j(\Sigma^*)^\xi}{\sqrt{\gamma_j(\Sigma^*)^\xi A \Sigma^* A^\gamma_j(\Sigma^*)}} > z_{1-\alpha} - \frac{\gamma_j(\Sigma^*)^\hat{\mu}}{\sqrt{\gamma_j(\Sigma^*)^\hat{\mu} A \Sigma^* A^\gamma_j(\Sigma^*)}}. \]

The term on the left-hand side has the standard normal distribution, and thus

\[ \mathbb{P} \left( p(\xi + \hat{\mu}, \Sigma^*) > 1 - \alpha \right) = \Phi\left( \frac{\gamma_j(\Sigma^*)^\hat{\mu}}{\sqrt{\gamma_j(\Sigma^*)^\hat{\mu} A \Sigma^* A^\gamma_j(\Sigma^*)}} - z_{1-\alpha} \right). \]

Next, note that by definition \( \gamma_j(\Sigma^*) = c_j(\Sigma^*) \tilde{\gamma}_j \), where by construction \( \tilde{\gamma}_j \) is equal to the vector \( \tilde{\gamma}_j \) defined above (i.e. the unique vector satisfying the unique vector such that \( \gamma_j^\prime \tilde{A}(B,-1) = 0, \tilde{A}_B \geq 0, ||\gamma|| = 1 \)). It is then immediate from the previous display and the fact that \( \hat{\mu}_B = -\tilde{A}(B,-1) \) that

\[ \mathbb{P} \left( p(\xi + \hat{\mu}, \Sigma^*) > 1 - \alpha \right) = \Phi\left( c^* x - z_{1-\alpha} \right). \]

But this implies that

\[ \liminf_{n \to \infty} \left| \mathbb{P}_{\tilde{p}} \left( \hat{p}(\tilde{Y}_m, \tilde{\Sigma}_m) > 1 - \alpha \right) - \Phi(c^* x - z_{1-\alpha}) \right| = 0, \]

which contradicts (81).

**Lemma G.7.** Suppose Assumption 3 holds. Let \( B = B(\delta^{**}) \) be the index of the binding moments. Let \( \tilde{\gamma}_1, ..., \tilde{\gamma}_J \) be as defined in Lemma F.1. Then \( \tilde{\gamma}_{j,-B} = 0 \) for exactly one \( j \in \{1, ..., J\} \). Additionally, \( \tilde{\gamma}_j \) is the unique vector in the set \( \{ \gamma_B : \gamma_B^\prime \tilde{A}(B,-1) = 0, \gamma_B \geq 0, ||\gamma|| = 1 \} \).

**Proof.** We first show that there can be at most one \( \tilde{\gamma}_j \) such that \( \tilde{\gamma}_{j,-B} = 0 \). Recall from the proof to Lemma F.1 that for all \( j \), \( \gamma_j^\prime \tilde{A}(\cdot,-1) = 0, \gamma_j \geq 0 \) and \( ||\gamma_j|| = 1 \). Thus, if \( \tilde{\gamma}_{j,-B} = 0 \), we have \( \gamma_j^\prime \tilde{A}(B,-1) = 0 \). However, from Lemma E.8, the set \( \{ \gamma_B : \gamma^\prime \tilde{A}(B,-1) = 0 \} = \{ c\gamma^*_B | c \in \mathbb{R} \} \) for some non-zero vector \( \gamma^*_B \geq 0 \). Thus, there is a single vector in the set \( \{ \gamma_B : \gamma_B^\prime \tilde{A}(B,-1) = 0, \gamma_B \geq 0, ||\gamma_B|| = 1 \} \). In particular, its lone element is \( c^* \gamma^*_B \), for \( c^* = 1/||\gamma^*_B|| \). Hence, if there is such a \( \gamma_j \), it has \( c^* \gamma^*_B \) in the positions corresponding with \( B \) and zeros otherwise.

It thus remains to show that the vector with \( c^* \gamma^*_B \) in the positions corresponding with \( B \) and zeros otherwise is in the set \( \{ \tilde{\gamma}_1, ..., \tilde{\gamma}_J \} \). Denote this vector \( \gamma^* \). Note that by construction, \( \gamma^* \tilde{A}(\cdot,-1) = 0 \). Thus, for any \( \Sigma \) positive definite, \( (\gamma^* \tilde{\sigma})^{-1} \gamma \in F(\Sigma) = \{ \gamma : \gamma^\prime \tilde{A}(\cdot,-1) = 0, \gamma^\prime \tilde{\sigma} = 1 \} \). Moreover, \( (\gamma^* \tilde{\sigma})^{-1} \gamma \) must be the unique vector in \( F(\Sigma) \) with \( \gamma_B = 0 \), since as discussed above, \( \{ \gamma_B : \gamma_B^\prime \tilde{A}(B,-1) = 0 \} = \{ c\gamma^*_B | c \in \mathbb{R} \} \) and so there is a unique vector with \( \gamma_B^\prime \tilde{A}(B,-1) = 0, \gamma \geq 0 \), and \( \gamma^\prime \tilde{\sigma} = 1 \). Let \( \nu \) be the vector with -1 in the positions corresponding with \( -B \) and zeros otherwise. Then \( \nu(\gamma^* \tilde{\sigma})^{-1} \gamma = 0 \), whereas \( \nu \gamma < 0 \) for any other \( \gamma \in F(\Sigma) \), since every \( \gamma \in F(\Sigma) \) satisfies

S-75
$\gamma \geq 0$ and $\gamma_{-B} \neq 0$. Thus, $(\gamma^*/\tilde{\sigma})^{-1}\gamma^*$ is a minimal face of $F(\Sigma)$, and hence a vertex (see Schrijver (1986), Section 8.5). By Lemma F.1, $F(\Sigma) = \{c_1(\Sigma)\tilde{\gamma}_1, ..., c_J(\Sigma)\tilde{\gamma}_J\}$ where $c_J > 0$. It follows that $(\gamma^*/\tilde{\sigma})^{-1}\gamma^* = c_j(\Sigma)\tilde{\gamma}_j$ for some $j$, so $\gamma^*$ is a constant multiple of $\tilde{\gamma}_j$. However, since by construction $\gamma^*$ and $\tilde{\gamma}_j$ are both positive and have a norm of 1, they must be equal, which gives the first result.

Next, note that we showed in the proof to Lemma E.8 that $\gamma^*_B\tilde{A}_{(B,\cdot)} = e_1^\prime$. Since $\tilde{A}_{(B,\cdot)} = A^\prime\begin{pmatrix} 0 & I \end{pmatrix}$, it follows that $\gamma^*_BA \neq 0$. Since $\tilde{\gamma}_{j,B} = c^*\gamma^*_B$ and $\gamma_{j,-B} = 0$, we have that $\gamma^*_jA = c^*\gamma^*_BA \neq 0$, which gives the second result.

Lemma G.8. Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4. Let $\hat{V}(Y, \Sigma) = \arg\max_{\gamma \in V(\Sigma)} \gamma'Y$. By Lemma G.7, there is a unique index $j$ such that $\bar{\gamma}_{j,-B} = 0$. Then for almost every $\xi \sim \mathcal{N}(0, A\Sigma^*A')$, there is a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$ such that $\hat{V}(Y, \Sigma) = c_j(\Sigma)\tilde{\gamma}_j$ for almost every $\xi \sim \mathcal{N}(0, A\Sigma^*A')$.

Proof. Without loss of generality, suppose that $\bar{\gamma}_{1,-B} = 0$. Lemma F.1 implies that

$$\hat{\eta}(Y, \Sigma) := \max_{\gamma \in V(\Sigma)} \gamma'Y = \max\{c_1(\Sigma)\tilde{\gamma}_1^jY, ..., c_J(\Sigma)\tilde{\gamma}_J^jY\},$$

where the functions $c_j(\Sigma)$ are continuous. Each of the elements of the max are continuous functions of $(Y, \Sigma)$ in a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$ by an argument analogous to that in the proof to Lemma G.4 (replacing $\mu^*$ with $\bar{\mu}$). Note, however, that $\tilde{\gamma}_1^j(\mu + \xi) = \tilde{\gamma}_{1,B}^j(\xi_0 + \tilde{A}_{(B,1)}x)$, which is finite with probability 1. On the other hand, for $j > 1$, $\tilde{\gamma}_j^j(\xi + \bar{\mu}) = -\infty$, since $\bar{\gamma}_{j} \geq 0$ and has at least one strictly positive element in the index $-B$. Since $c_j(\Sigma^*) > 0$ for all $j$ by Lemma G.3, it follows that $c_1(\Sigma^*)\tilde{\gamma}_1^j(\xi + \bar{\mu}) > c_j(\Sigma^*)\tilde{\gamma}_j^j(\xi + \bar{\mu})$ for all $j > 2$. Since the functions on both sides of the inequality are continuous at $(\xi + \bar{\mu}, \Sigma^*)$, this implies that $c_1(\Sigma)\tilde{\gamma}_1^jY > c_j(\Sigma)\tilde{\gamma}_j^jY$ in a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$, which gives the desired result.

Lemma G.9. Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4. Let $\hat{\eta}(Y, \Sigma) = \max_{\gamma \in V(\Sigma)} \gamma'Y$. Then for almost every $\xi \sim \mathcal{N}(0, A\Sigma^*A')$, $\eta(Y, \Sigma)$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$. Further, there is a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$ such that $\hat{\eta}(Y, \Sigma) = c_j(\Sigma)\tilde{\gamma}_jY$, where $j$ is the unique index such that $\tilde{\gamma}_{j,-B} = 0$ (which exists by Lemma G.7).

Proof. Follows immediately from the proof to Lemma G.8.

Lemma G.10. Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4. Then for almost every $\xi \sim \mathcal{N}(0, A\Sigma^*A')$, $\sigma_\eta(Y, \Sigma)$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$. Further, there is a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$ such that $\sigma_\eta(Y, \Sigma) = c_j(\Sigma)^2\tilde{\gamma}_jA\Sigma A'\tilde{\gamma}_j > 0$.

Proof. By Lemma G.7, there is a unique index $j$ such that $\tilde{\gamma}_{j,-B} = 0$, and this $\tilde{\gamma}_j$ satisfies $\tilde{\gamma}_jA \neq 0$. Lemma G.8 implies that $\hat{V}(Y, \Sigma) = c_j(\Sigma)\tilde{\gamma}_j$ in a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$. Thus, in that
neighborhood, \( \hat{\sigma}^2_\eta(Y, \Sigma) = c_j(\Sigma)^2 \tilde{\gamma}'_j A \Sigma A' \tilde{\gamma}_j \), which is clearly continuous in \( \Sigma \). Additionally, \( c_j(\Sigma^*) > 0 \) by Lemma G.3, and \( \Sigma^* \) is positive definite, so \( \hat{\sigma}^2_\eta(\xi + \bar{\mu}, \Sigma^*) = c_j(\Sigma^*)^2 \tilde{\gamma}'_j \Sigma^* A' \tilde{\gamma}_j > 0 \). Since \( \hat{\sigma}^2_\eta \) is continuous at \((\xi + \bar{\mu}, \Sigma^*)\), it is also positive in a neighborhood of \((\xi + \bar{\mu}, \Sigma^*)\).

**Lemma G.11.** Let \( \bar{\mu} \) and \( \Sigma^* \) be as defined in the proof to Proposition F.4. Then for almost every \( \xi \sim \mathcal{N}(0, A \Sigma^* A') \), \( v^{lo}(\xi + \bar{\mu}, \Sigma^*) = -\infty \), \( v^{up}(\xi + \bar{\mu}, \Sigma^*) = \infty \), and the functions \( v^{lo} \) and \( v^{up} \) are continuous at \((\xi + \bar{\mu}, \Sigma^*)\).

**Proof.** By Lemma G.7, there is a unique index \( j \) such that \( \tilde{\gamma}_j \cdot B = 0 \), and this \( \tilde{\gamma}_j \) satisfies \( \tilde{\gamma}'_j A \neq 0 \). Without loss of generality, assume this holds for \( j = 1 \). Lemmas G.8 and G.10 then imply that \( \hat{\nu}(Y, \Sigma) = c_1(\Sigma) \tilde{\gamma}_1 \) and \( \hat{\sigma}^2_\eta(Y, \Sigma) > 0 \) in a neighborhood of \((\xi + \bar{\mu}, \Sigma^*)\).

The proof of the continuity of \( v^{lo} \) and \( v^{up} \) is then similar to that in Lemma G.4. Let \( \gamma_*(\Sigma) = c_1(\Sigma) \tilde{\gamma}_1 \). For ease of notation, we will make the dependence of \( \gamma_* \) on \( \Sigma \) implicit where it is clear below. Since in a neighborhood of \((\xi + \bar{\mu}, \Sigma^*)\), \( \hat{\sigma}^2_\eta(Y, \Sigma) > 0 \) and \( \hat{\nu}(Y, \Sigma) = \{\gamma_*(\Sigma)\} \), in that neighborhood \( v^{lo}(Y, \Sigma) \) is the minimum of the set

\[
C = \{ c : \max_{\gamma \in V(\Sigma)} \gamma' \left( S_{\gamma_*(\Sigma)} + \frac{\Sigma \gamma_* \gamma_*'}{\gamma_* \Sigma \gamma_*} - c \right) \},
\]

for

\[
S_{\gamma_*(\Sigma)} = \left( I - \frac{\Sigma \gamma_* \gamma_*'}{\gamma_* \Sigma \gamma_*} \right) Y.
\]

Rearranging terms, we see that

\[
C = \{ c : 0 = \max_{\gamma \in V(\Sigma)} a_{\gamma, \gamma_*, \Sigma} + b_{\gamma, \gamma_*, \Sigma} c \},
\]

where \( a_{\gamma, \gamma_*, \Sigma} := \gamma' S_{\gamma_*(\Sigma)} \) and \( b_{\gamma, \gamma_*, \Sigma} := \gamma' \Sigma \gamma_* \gamma_*' / \gamma_* \Sigma \gamma_* - 1 \). Note that \( a_{\gamma_*, \gamma_*, \Sigma} = b_{\gamma_*, \gamma_*, \Sigma} = 0 \), so \( 0 \leq \max_{\gamma \in V(\Sigma)} a_{\gamma, \gamma_*, \Sigma} + b_{\gamma, \gamma_*, \Sigma} c \) for all \( c \). Moreover, for \( c = \gamma'_* Y \), the max is attained at \( \gamma_* \) by construction. Hence, the set \( C \) is non-empty.

Intuitively, if we plot \( a_{\gamma, \gamma_*, \Sigma} + b_{\gamma, \gamma_*, \Sigma} \) as a function of \( c \), then each \( \gamma \in V(\Sigma) \) defines a line, and the set \( C \) represents the values of \( c \) for which 0 is the upper envelope of this set. It follows that the lower bound of \( C \) is the maximal x-intercept of a line of the form \( a_{\gamma, \gamma_*, \Sigma} + b_{\gamma, \gamma_*, \Sigma} c \) with \( b_{\gamma, \gamma_*, \Sigma} < 0 \). Hence,

\[
v^{lo}(Y, \Sigma) = \max_{\gamma \in V(\Sigma) \setminus \{\gamma_*\}} \frac{-\hat{a}_{\gamma, \gamma_*, \Sigma}}{\hat{b}_{\gamma, \gamma_*, \Sigma}}.
\]

Now, let \( \gamma_{**} = \gamma_*(\Sigma^*) \). Observe that for any \( \gamma \in V(\Sigma^*) \setminus \gamma_{**} \),

\[
\gamma' \left( I - \frac{\Sigma^* \gamma_{**} \gamma_{**}'}{\gamma_{**}' \Sigma^* \gamma_{**}} \right) (\xi + \bar{\mu}) = \gamma'(\xi + \bar{\mu}) - \frac{\gamma' \Sigma^* \gamma_{**} \gamma_{**}'}{\gamma_{**}' \Sigma^* \gamma_{**}}.
\]

Since \( \gamma_{**} \neq 0 \) and has at least one strictly positive element, \( \gamma'(\xi + \bar{\mu}) = -\infty \) with probability 1. On the other hand, \( \gamma_{**, \Sigma} = 0 \), and so \( \gamma_{**}(\xi + \bar{\mu}) \) is finite with probability one. It follows that
$a_{\gamma, \gamma^*, \xi + \bar{\mu}, \Sigma^*} = -\infty$ with probability 1. Hence, $v^{lo}(\xi + \bar{\mu}, \Sigma^*) = -\infty$.

Next, recall that by Lemma F.1, $V(\Sigma) := \{\gamma_1(\Sigma), ..., \gamma_j(\Sigma)\}$, where $\gamma_j(\Sigma) := c_j(\Sigma)\tilde{\gamma}_j$ and $c_j(\Sigma)$ is continuous. Additionally, we showed in the proof to Lemma G.8 that for all $j$, $c_j(\Sigma)\tilde{\gamma}_j$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$. It is then immediate from the definitions of the functions $a_{\gamma, \gamma^*, \Sigma^*}$ and $b_{\gamma, \gamma^*, \Sigma^*}$ that for all $j$, $a_{\gamma_j(\Sigma), \gamma^*_j(\Sigma), \Sigma}$ and $b_{\gamma_j(\Sigma), \gamma^*_j(\Sigma), \Sigma}$ are continuous in $(Y, \Sigma)$ as well. Without loss of generality, suppose that for $2 \leq k \leq k_1$, $b_{\gamma_k(\Sigma^*), \gamma^*_k(\Sigma^*), \Sigma^*} < 0$; for $k_1 < k \leq k_2$, $b_{\gamma_k(\Sigma^*), \gamma^*_k(\Sigma^*), \Sigma^*} = 0$; and for $k > k_2$, $b_{\gamma_k(\Sigma^*), \gamma^*_k(\Sigma^*), \Sigma^*} > 0$. From the continuity of $b_{\gamma_j(\Sigma), \gamma^*_j(\Sigma), \Sigma}$, it is clear that in a neighborhood of $(\xi + \mu^*, \Sigma^*)$, $b_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), \Sigma} > 0$ for all $2 \leq k \leq k_1$ and $b_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), \Sigma} < 0$ for all $k > k_2$. Hence, in this neighborhood,

$$v^{lo}(Y, \Sigma) = \max \left\{ \max_{\gamma_k(\Sigma) : 2 \leq k \leq k_1} \frac{-a_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), Y, \Sigma}}{b_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), \Sigma}}, \max_{\gamma \in V^0(\Sigma)} \frac{-a_{\gamma, \gamma^*(\Sigma), Y, \Sigma}}{b_{\gamma, \gamma^*(\Sigma), \Sigma}} \right\},$$

where

$$V^0(\Sigma) := \{\gamma_k(\Sigma) : k_1 < k \leq k_2, b_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), \Sigma} < 0\}$$

and we define the max of an empty set to be $-\infty$. It is clear from the continuity of the functions $a$ and $b$ that the inner max on the left side of (82) is continuous and converges to $-\infty$. To show that $v^{lo}$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$, it thus suffices to show that for any sequence $(Y, \Sigma) \to (\xi + \bar{\mu}, \Sigma^*)$, the max on the right hand side of (82) converges to $-\infty$. To do this, note that by construction $b_{\gamma_k(\Sigma^*), \gamma^*_k(\Sigma^*), \Sigma^*} = 0$ for $k \in (k_1, k_2)$, and so along any sequence $(Y, \Sigma) \to (\xi + \bar{\mu}, \Sigma^*)$, $b_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), \Sigma} \to 0$ since $b$ is continuous in $(Y, \Sigma)$. Additionally, since $a$ is continuous, along such a sequence, $a_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), Y, \Sigma} \to a_{\gamma_k(\Sigma^*), \gamma^*_k(\Sigma^*), Y, \Sigma}$ becomes arbitrarily negative, whereas for values of $(Y, \Sigma)$ where $b_{\gamma_k(\Sigma), \gamma^*_k(\Sigma), \Sigma} > 0$, $\gamma_k$ is not included in $V^0$. It is then immediate that the max on the right hand side of (82) converges to $-\infty$, which suffices to establish the continuity of $v^{lo}$ at $(\xi + \bar{\mu}, \Sigma^*)$. The continuity of $v^{up}$ can be shown analogously.

\[\square\]

**Lemma G.12.** Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4. Define $p(Y, \Sigma)$ as in Lemma G.1. Then for almost every $\xi \sim \mathcal{N}(0, A\Sigma^*A')$, $p(Y, \Sigma)$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$, and

$$p(\xi + \bar{\mu}, \Sigma^*) = \Phi \left( \frac{\gamma_j(\Sigma^*)'(\xi + \bar{\mu})}{\sqrt{\gamma_j(\Sigma^*)'(A\Sigma^*A')\gamma_j(\Sigma^*)}} \right),$$

where $j$ is the unique index such that $\tilde{\gamma}_{j,B} = 0$ (which exists by Lemma G.7).

**Proof.** Lemmas G.9 to G.11 imply that for almost every $\xi$, $\dot{\eta}(Y, \Sigma)$, $\sigma^2_\eta(Y, \Sigma)$, $v^{lo}(Y, \Sigma)$ and $v^{up}(Y, \Sigma)$ are continuous at $(\xi + \bar{\mu}, \Sigma^*)$, and when evaluated at $(\xi + \bar{\mu}, \Sigma^*)$, $\dot{\eta} = c_j(\Sigma^*)\gamma_j'(\xi + \bar{\mu}) \tilde{\sigma}_\eta^2 = c_j(\Sigma^*)\tilde{\gamma}_j A\Sigma^*A'\gamma_j > 0$, $v^{lo} = -\infty$, and $v^{up} = \infty$. Thus, $\tilde{\sigma}_\eta > 0$ and $v^{lo} < v^{up}$ in a neighborhood of $(\xi + \bar{\mu}, \Sigma^*)$. When $\tilde{\sigma}_\eta > 0$ and $v^{lo} < v^{up}$,
Proof. Follows immediately from Lemmas G.12 and G.13 and the fact that the product of continuous functions is continuous. Additionally, when evaluated at $(Y, \Sigma) = (\xi + \bar{\mu}, \Sigma^*)$, we have

\[
p(Y, \Sigma) = \Phi\left(\frac{\gamma_j(\Sigma^*)(\xi + \bar{\mu})}{\sqrt{\gamma_j(\Sigma^*)^4 A \Sigma^* A' \gamma_j(\Sigma^*)}}\right) - \Phi(-\infty) = \Phi\left(\frac{\gamma_j(\Sigma^*)(\xi + \bar{\mu})}{\sqrt{\gamma_j(\Sigma^*)^4 A \Sigma^* A' \gamma_j(\Sigma^*)}}\right).
\]

Lemma G.13. Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4. For any $\bar{C} \in \mathbb{R}$, the function $1[\bar{\eta}(Y, \Sigma) \geq -\bar{C}]$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$ for almost every $\xi \sim \mathcal{N}(0, A \Sigma^* A')$.

Proof. By Lemma G.9, for almost every $\xi$, the function $\bar{\eta}(Y, \Sigma)$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$. It thus suffices to show that for almost every $\xi$, $\bar{\eta}(\xi + \mu, \Sigma^*) \neq -\bar{C}$. Lemma G.9 gives that $\bar{\eta}(\xi + \mu, \Sigma^*) = c_j(\Sigma^*) \gamma_j'(\xi + \bar{\mu})$ where $\gamma_j$ is the unique element of $\{\gamma_1, \ldots, \gamma_J\}$ such that $\gamma_j, -B = 0$. Thus, $\bar{\eta}(\xi + \mu, \Sigma^*) = -\bar{C}$ only if $c_j(\Sigma^*) \gamma_j'\xi = -\bar{C} - c_j(\Sigma^*) \gamma_j'\bar{\mu}$, where the right-hand side of the previous equation is finite since $\bar{\mu}B$ is finite and $\bar{\gamma}_j, -B = 0$. Observe further that $c_j(\Sigma^*) \gamma_j'\xi$ is normally distributed with variance $c_j(\Sigma^*)^2 \gamma_j' A \Sigma^* A' \gamma_j$, which is positive for almost every $\xi$ by Lemma G.10. Since $c_j(\Sigma^*) \gamma_j'\xi$ is continuously distributed, it follows that $c_j(\Sigma^*) \gamma_j'\xi = -\bar{C} - c_j(\Sigma^*) \gamma_j'\bar{\mu}$ with probability zero, which suffices for the result.

Lemma G.14. Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4. Let the function $p(Y, \Sigma)$ be as defined in Lemma G.12. For any $\bar{C} \in \mathbb{R}$, the function $\bar{p}(Y, \Sigma) := p(Y, \Sigma) \cdot 1[\bar{\eta}(Y, \Sigma) \geq -\bar{C}]$ is continuous at $(\xi + \bar{\mu}, \Sigma^*)$ for almost every $\xi \sim \mathcal{N}(0, A \Sigma^* A')$.

Proof. Follows immediately from Lemmas G.12 and G.13 and the fact that the product of continuous functions is continuous.

Lemma G.15. Let $\bar{\mu}$ and $\Sigma^*$ be as defined in the proof to Proposition F.4 and $\bar{p}(Y, \Sigma)$ as defined in Lemma G.14. For $\xi \sim \mathcal{N}(0, A \Sigma^* A')$, $\bar{p}(\xi + \bar{\mu}, \Sigma^*) = 1 - \alpha$ with probability 0.

Proof. Note that $\bar{p}(Y, \Sigma) := p(Y, \Sigma) 1[\bar{\eta}(Y, \Sigma) \geq -\bar{C}]$ can equal $1 - \alpha$ only if $1[\bar{\eta}(Y, \Sigma) \geq -\bar{C}] = 1$ and $p(Y, \Sigma) = 1 - \alpha$. It thus suffices to show that $p(\xi + \bar{\mu}, \Sigma^*) = 1 - \alpha$ with probability zero. From Lemma G.12, for almost every $\xi$, $p(\xi + \bar{\mu}, \Sigma^*) = \Phi\left(\frac{\gamma_j(\Sigma^*)'(\xi + \bar{\mu})}{\sqrt{\gamma_j(\Sigma^*)^4 A \Sigma^* A' \gamma_j(\Sigma^*)}}\right)$, where $\gamma_j(\Sigma^*) := c_j(\Sigma^*) \gamma_j$ and $\gamma_j$ is the unique element of $\{\gamma_1, \ldots, \gamma_J\}$ such that $\gamma_j, -B = 0$. Thus, $p(\xi + \bar{\mu}, \Sigma^*) = 1 - \alpha$ if $\gamma_j(\Sigma^*)'(\xi) = \tilde{z}_{1-\alpha} \sqrt{\gamma_j(\Sigma^*)^4 A \Sigma^* A' \gamma_j(\Sigma^*) - \gamma_j(\Sigma^*)'} \bar{\mu}$. However, we showed in the proof to Lemma G.13 that $\gamma_j(\Sigma^*)'$ is continuously distributed, and thus this occurs with probability 0.
Lemma G.16. Let $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_j$ be as defined in Lemma F.1, and $\gamma_j(\Sigma) := c_j(\Sigma)\tilde{\gamma}_j$. There exists a value $C^* \in \mathbb{R}$ such that for any $\Sigma \in \mathcal{S}$ and any $j$ such that $\tilde{\gamma}_j' A \neq 0$,\[
abla \left( \frac{\hat{\eta}}{\sqrt{\gamma_j(\Sigma)' A \Sigma A' \gamma_j(\Sigma)}} \right) > 1 - \alpha
\]only if $\hat{\eta} > C^*$.

Proof. Observe that\[
\Phi \left( \frac{\hat{\eta}}{\sqrt{\gamma_j(\Sigma)' A \Sigma A' \gamma_j(\Sigma)}} \right) > 1 - \alpha
\]iff\[
\hat{\eta} > z_{1-\alpha} \sqrt{\gamma_j(\Sigma)' A \Sigma A' \gamma_j(\Sigma)}.
\]If $z_{1-\alpha} \geq 0$, then the lower bound in the previous display is weakly greater than zero. On the other hand if $z_{1-\alpha} < 0$, then the lower bound is weakly greater than $z_{1-\alpha}$ times the maximum possible value of $\sqrt{\gamma_j(\Sigma)' A \Sigma A' \gamma_j(\Sigma)}$. Note, however, that $\sqrt{\gamma_j(\Sigma)' A \Sigma A' \gamma_j(\Sigma)} = \sqrt{c_j(\Sigma)^2 \tilde{\gamma}_j' A \Sigma A' \tilde{\gamma}_j}$. By Lemma G.3, $c_j(\Sigma) \leq \hat{c}$. Additionally, since the set $\{\tilde{\gamma}_1, \ldots, \tilde{\gamma}_j\}$ is finite, $\max_j ||\tilde{\gamma}_j' A||^2$ is finite. It then follows from Lemma G.3 that $\tilde{\gamma}_j' A \Sigma A' \tilde{\gamma}_j \leq \hat{\lambda} \max_j ||\tilde{\gamma}_j' A||^2 < \infty$, and so we obtain a finite upper bound on $\sqrt{\gamma_j(\Sigma)' A \Sigma A' \gamma_j(\Sigma)}$, which suffices for the result.

Proof of Proposition F.4

Proof. We first claim that the function $m(\beta) = A\beta$ is a maximal invariant of the group $G$. Since by definition $Av = 0$ for any $v \in A^\perp$, it is immediate that $m(\beta) = m(g_0 \beta)$ for any $g_0 \in G$. To show that $m$ is a maximal invariant, consider $\beta_1$ and $\beta_2$ such that $m(\beta_1) = m(\beta_2)$. Then $A(\beta_1 - \beta_2) = 0$ and hence $(\beta_1 - \beta_2) \in A^\perp$. From this we see that $\beta_1 = \beta_2 + (\beta_1 - \beta_2) = g(\beta_1 - \beta_2) (\beta_2)$, and thus $m(\beta)$ is a maximal invariant. Note further that $A\beta_1 = A\beta_2$ iff $A\beta_1 + h = A\beta_2 + h$ for any constant vector $h$, and so the same argument applies to show that $m(\beta) = A\beta + h$ is maximal for any $h$. It follows from Theorem 1 in Lehmann (1986, p. 285) that $C$ can be written as a function of $(m(\beta), \hat{\Sigma})$ only, so that $C(\sqrt{n}\beta_n, \hat{\Sigma}_n) = \tilde{C}(m(\sqrt{n}\beta_n), \hat{\Sigma}_n)$. From Lemma E.8, there exists a vector $\tilde{\tau}$ such that

\begin{align}
A_{(B,\cdot)} \beta_{\cdot \cdot} - d_B - \tilde{A}_{(B,1)} \theta_{\cdot \cdot} - \tilde{A}_{(B,-1)} \hat{\tau} &= 0 \\
A_{(-B,\cdot)} \beta_{\cdot \cdot} - d_{-B} - \tilde{A}_{(-B,1)} \theta_{\cdot \cdot} - \tilde{A}_{(-B,-1)} \hat{\tau} &= -\epsilon < 0.
\end{align}

(83) (84)

We set the constant $h = -[d - \tilde{A}_{(\cdot,1)} \theta_{\cdot \cdot} - \tilde{A}_{(\cdot,-1)} \hat{\tau}]$, so that $\tilde{C}$ is a function of $Y_n := \sqrt{n}[A\beta_n - d - \tilde{A}_{(\cdot,1)} \theta_{\cdot \cdot} - \tilde{A}_{(\cdot,-1)} \hat{\tau}]$ and $\hat{\Sigma}_n$. 

S-80
Observe that

\[ Y_n = \sqrt{n} A(\beta - \beta_p^*) - \sqrt{n}[A\beta_p^* - d - \bar{A}_{(\cdot,-1)}\tilde{\tau}] \]

It follows immediately from (83) and (84) that \( \sqrt{n}[A\beta_p^* - d - \bar{A}_{(\cdot,-1)}\tilde{\tau}] \to \bar{\mu}, \) where \( \bar{\mu}_B = 0 \) and \( \bar{\mu}_{-B} = -\infty. \) Since by assumption \( \sqrt{n}(\tilde{\beta}_n - \beta_{p^*}) \to_d \mathcal{N}(0, \Sigma^*) \) under \( P^\ast, \) the continuous mapping theorem along with Slutsky’s lemma imply that \( Y_n \xrightarrow{P^\ast_d} \xi + \bar{\mu} \) for \( \xi \sim \mathcal{N}(0, A\Sigma^* A'). \)

Similarly, suppose \( \beta_{P_n} = \beta_{p^*} + \frac{1}{\sqrt{n}}(\tilde{\beta} - \beta_{p^*}) \) for some fixed \( \tilde{\beta}. \) Suppose further that \( \sqrt{n}(\tilde{\beta}_n - \beta_{P_n}) \xrightarrow{P_{n_d}} \mathcal{N}(0, \Sigma^*). \) Observe that

\[ Y_n = \sqrt{n} A(\tilde{\beta} - \beta_{P_n}) + A(\beta - \beta_{p^*}) - \sqrt{n}[A\beta_{p^*} - d - \bar{A}_{(\cdot,-1)}\tilde{\tau}] \]

Thus, \( Y_n \xrightarrow{P_{n_d}} \xi + A(\beta - \beta_{p^*}) + \bar{\mu}. \)

Now, as in Lemma E.13, let \( B_0(\tilde{\beta}) := \{ \beta : \exists \tau \text{ s.t. } l'\tau = \tilde{\theta}, A\beta - d - A\left(\begin{array}{c} 0 \\ \tau \end{array}\right) \leq 0 \} \) be the set of values \( \beta \) consistent with \( \theta = \tilde{\theta}, \) and \( B_0^B(\tilde{\theta}) = \{ \beta : \exists \tau \text{ s.t. } l'\tau = \tilde{\theta}, A_{(B,\cdot)}\beta - d_B - A_{(B,\cdot)}\left(\begin{array}{c} 0 \\ \tau \end{array}\right) \leq 0 \} \) be the analogous set using only the moments \( B. \) Suppose that \( \tilde{\beta} \in B_0^B(\theta^{ub} + x). \) We claim that for \( n \) sufficiently large, \( \beta_n := \beta_{p^*} + \frac{1}{\sqrt{n}}(\tilde{\beta} - \beta_{p^*}) \in B_0(\theta^{ub} + \frac{1}{\sqrt{n}} x). \) It follows from the definition of \( B_0^B(\theta^{ub} + x) \) and the construction of the matrix \( \tilde{A} \) that there exists \( \tilde{\tau} \) such that \( A_{(B,\cdot)}\tilde{\beta} - d_B - \tilde{A}_{(B,1)}(\theta_{p^*}^{ub} + x) - \tilde{A}_{(B,-1)}\tilde{\tau} \leq 0. \) This, combined with (83), implies that

\[ A_{(B,\cdot)}\beta_n - d_B - \tilde{A}_{(B,1)}(\theta_{p^*}^{ub} + \frac{1}{\sqrt{n}} x) - \tilde{A}_{(B,-1)}((1 - \frac{1}{\sqrt{n}})\tilde{\tau} + \frac{1}{\sqrt{n}} \tilde{\tau}) \leq 0. \]

However, from (84), it follows that

\[ A_{(B,\cdot)}\beta_n - d_B - \tilde{A}_{(B,1)}(\theta_{p^*}^{ub} + \frac{1}{\sqrt{n}} x) - \tilde{A}_{(B,-1)}((1 - \frac{1}{\sqrt{n}})\tilde{\tau} + \frac{1}{\sqrt{n}} \tilde{\tau}) = (1 - \frac{1}{\sqrt{n}})(-\epsilon + \frac{1}{\sqrt{n}} (A_{(B,\cdot)}\tilde{\beta} - d_B - \tilde{A}_{(B,1)}(\theta_{p^*}^{ub} + x) - \tilde{A}_{(B,-1)}\tilde{\tau}), \]

which is negative for \( n \) sufficiently large since \( -\epsilon < 0. \) The previous two displays imply that for \( n \) sufficiently large, \( \beta_n \in B_0(\theta_{p^*}^{ub} + \frac{1}{\sqrt{n}} x), \) as we desired to show. Hence, for \( n \) sufficiently large, there exists \( \delta_n \in \Delta \) and \( \tau_n \) such that \( \beta_n = \delta_n + \left(\begin{array}{c} 0 \\ \tau_n \end{array}\right) \) and \( l'\tau_n = \theta^{ub} + \frac{1}{\sqrt{n}} x. \)

Now, let \( \varphi_n(Y_n, \hat{\Sigma}_n) = 1[\theta_{p^*}^{ub} + \frac{1}{\sqrt{n}} x \in \hat{C}(Y_n, \hat{\Sigma}_n)]. \) It follows from the previous paragraph along with the assumptions of the proposition that for any sequence \( P_n \) such that \( \sqrt{n}(\tilde{\beta}_n - \beta_{P_n}) \xrightarrow{P_{n_d}} \mathcal{N}(0, \Sigma^*), \) \( \hat{\Sigma}_n \xrightarrow{P_n} \Sigma^*, \) and \( \beta_{P_n} = \beta_{p^*} + \frac{1}{\sqrt{n}}(\tilde{\beta} - \beta_{p^*}) \) for \( \tilde{\beta} \in B_0^B(\theta_{p^*}^{ub}), \) we have that

\[ \limsup_{n \to \infty} \mathbb{E}_{P_n} \left[ \varphi_n(Y_n, \hat{\Sigma}_n) \right] \leq \alpha. \]
It then follows from Theorem 1 in Mueller Müller (2011) that

\[
\limsup_{n \to \infty} \mathbb{E}_{P^*} \left[ \varphi_n(Y_n, \hat{\Sigma}_n) \right] \leq \tilde{\rho},
\]
for \( \tilde{\rho} \) the power of the most powerful test between

\[
H_0 : \tilde{\beta} \in B^B_0 (\theta^*_P + x) \text{ vs. } H_1 : \tilde{\beta} = \beta_{P^*}
\]
given a single observation \( Y \sim \mathcal{N} \left( \bar{\mu} + A(\tilde{\beta} - \beta_{P^*}), A\Sigma^*A' \right) \). Since \( \mu = -\infty, Y = -\infty \) with probability 1 under both the null and alternative, so it suffices to consider tests of \( H_0 \) vs \( H_1 \) given an observation \( Y_B \sim \mathcal{N} \left( \bar{\mu}_B + A_{B,\cdot}(\tilde{\beta} - \beta_{P^*}), A_{B,\cdot}\Sigma^*A'_{(B,\cdot)} \right) \). Recalling that \( \bar{\mu}_B = 0 \) by construction, we see that \( \tilde{\rho} \) is the power of the most powerful test between \( H_0 : \mu \in M_0 := \{ A_{B,\cdot}(\tilde{\beta} - \beta_{P^*}) : \tilde{\beta} \in B^B_0 (\theta^*_P + x) \} \) and \( H_1 : \mu = 0 \) given \( Y \sim \mathcal{N} \left( \mu, A_{B,\cdot}\Sigma^*A'_{(B,\cdot)} \right) \).

Now, it follows from the proof to Lemma E.13 that

\[
B^B_0 (\theta^*_P + x) = \{ \beta : \tilde{\gamma}_B' \left( A_{B,(\cdot)} \beta - d_B - \tilde{A}_{(B,1)} (\theta^*_P + x) \right) \leq 0 \},
\]
for \( \tilde{\gamma}_B \) the unique vector such that \( \tilde{\gamma}'_B \tilde{A}_{(B,1)} = 0, \tilde{\gamma}_B \geq 0, ||\tilde{\gamma}_B|| = 1 \). This, combined with (83) and the fact that \( \gamma_{(B,1)} = 0 \), implies that \( B^B_0 (\theta^*_P + x) = \{ \beta : \tilde{\gamma}_B' \left( A_{B,(\cdot)} (\beta - \beta_{P^*}) \right) \leq \tilde{\gamma}_B' \tilde{A}_{(B,1)} x \} \) is then immediate that \( M_0 \subseteq \{ v : \tilde{\gamma}_B v \leq \tilde{\gamma}_B' \tilde{A}_{(B,1)} x \} \). Additionally, since \( \delta_{P^*} \) satisfies Assumption 3, \( A_{(B,\cdot)} \) has rank \( B \), and thus its image is \( \mathbb{R}^{|B|} \). This implies inclusion in the opposite direction, and hence \( M_0 = \{ v : \tilde{\gamma}_B v \leq \tilde{\gamma}_B' \tilde{A}_{(B,1)} x \} \). It then follows from Lemma E.12 that \( \tilde{\rho} = \Phi \left( -\tilde{\gamma}_{(B)}' \tilde{A}_{(B,1)} x / \sigma^*_B - z_{1-\alpha} \right) \), for \( \sigma_B^* = \sqrt{\tilde{\gamma}'_B \tilde{A}_{(B,\cdot)} \Sigma^*A'_{(B,\cdot)} \tilde{\gamma}_B} \). This accords with the formula for \( \rho^*(P^*, x) \) given in Proposition 5.2, which completes the proof.

Proof of Corollary F.2

Proof. The hybrid test rejects whenever the second stage conditional test rejects. The second stage of the hybrid is identical to the size \( \tilde{\alpha} \) conditional test, except that it replaces \( v^{lo} \) with the minimum of \( v^{lo} \) and the least favorable critical value \( c_{LF,\kappa} \). However, \( p(\zeta < \tilde{\eta} | \zeta \in [v^{lo}, v^{up}], \xi \sim \mathcal{N}(0, \sigma^2_\tilde{\eta}) \) is decreasing in \( v^{lo} \), and thus the second stage of the hybrid rejects whenever the size \( \tilde{\alpha} \) conditional test rejects. The result then follows immediately from Proposition F.4.

\[\text{\footnotesize\(^{46}\)See also Section 3.2 of Müller (2011) on applying Theorem 1 to invariant tests.}\]
H Additional Simulation Results

This section contains results from additional simulations that complement the analysis in Section 8. Section H.1 contains additional results from the normal data-generating process considered in Section 8. Section H.2 presents results from a non-normal data-generating process in which the covariance matrix is estimated from the data.

H.1 Additional Results for Normal Simulations

In the main text, we report efficiency in terms of excess length for the median paper considered in our simulations. Figure I1 supplements these results by reporting the minimum and maximum excess length efficiency ratios over the 12 papers used in our simulation design in the main text.

Figures I2 and I3 show results using the average of the post-period causal effects as the target parameter, rather than the first period after treatment.

Throughout the figures in this section, we report results for the conditional-LF hybrid introduced in Section C in addition to the conditional, FLCI, and conditional-FLCI hybrid approaches considered in the main text.

In addition to the specification for \( \Delta \) considered in the main text, we also report a small set of simulations using \( \Delta = \Delta^{RMI} \). Recall that the FLCIs are always infinite-length under \( \Delta^{RMI} \), so in this simulation design we merely consider the conditional and conditional-LF hybrid. Our simulations are conducted under the assumption of parallel trends, and thus the conditions for the asymptotic optimality of the conditional test do not hold for any value of the parameter \( \bar{M} \). Nonetheless, we find decent performance for both procedures, with modest improvements for the conditional-LF hybrid over the conditional approach. The results are reported in Figure I5.
Figure I1: Excess Length Ratios — Minimum, Maximum, and Median Over the Surveyed Papers When $\theta = \tau_1$

Note: This figure shows the ratio of average excess length relative to the excess length of the shortest confidence set that controls size for $\theta = \tau_1$ and each choice of $\Delta = \Delta^{SD}(M), \Delta^{SPB}(M), \Delta^{SI}(M)$. Simulations are calibrated to each of the papers reviewed in Roth (2019), and the median, minimum, and maximum over these papers is shown. The procedures considered are the FLCIs, conditional confidence sets, conditional-FLCI, and conditional-least favorable hybrid confidence sets. Results are averaged over 1000 simulations. See Section 8 for details.
Figure I2: Median efficiency ratios for proposed procedures when $\theta = \bar{\tau}_{post}$.

Note: This figure shows the median efficiency ratios for our proposed confidence sets for $\theta = \bar{\tau}_{post}$. The efficiency ratio for a procedure is defined as the optimal bound divided by the procedure’s expected excess length. Results are averaged over 1000 simulations for each of the 12 papers surveyed, and the median across papers is reported here. See Section 8 for details.
Figure I3: Excess Length Ratios — Minimum, Maximum, and Median Over the Surveyed Papers When $\theta = \bar{\tau}_{\text{post}}$

\[ \Delta^{SD}(M), \theta = \bar{\tau}_{\text{post}} \]

\[ \Delta^{SDP}(M), \theta = \bar{\tau}_{\text{post}} \]

\[ \Delta^{SDI}(M), \theta = \bar{\tau}_{\text{post}} \]

Note: This figure shows the ratio of average excess length relative to the excess length of the shortest confidence set that controls size for $\theta = \bar{\tau}_{\text{post}}$ and each choice of $\Delta = \Delta^{SD}(M), \Delta^{SDP}(M), \Delta^{SDI}(M)$. Simulations are calibrated to each of the papers reviewed in Roth (2019), and the median, minimum, and maximum over these papers is shown. The procedures considered are the FLCIs, conditional confidence sets, conditional-FLCI, and conditional-least favorable hybrid confidence sets. Results are averaged over 1000 simulations. See Section 8 for details.

S-86
Figure I4: Comparison of median efficiency ratios for conditional and conditional-LF hybrid.

Note: This figure shows the median efficiency ratios for the conditional confidence set and conditional least-favorable hybrid confidence set. The efficiency ratio for a procedure is defined as the optimal lower bound divided by the procedure’s expected excess length. Results are averaged over 1000 simulations for each of the papers surveyed in Roth (2019), and the median across papers is reported here. See Section 8 for details.
Figure I5: Simulation results: Median efficiency ratios for proposed procedures, $\Delta = \Delta^{RMI}(\bar{M})$ and $\theta = \tau_1$

Note: This figure shows the median efficiency ratios for the conditional confidence set and conditional least-favorable hybrid confidence set for $\Delta = \Delta^{RMI}(\bar{M})$. The efficiency ratio for a procedure is defined as the optimal bound divided by the procedure’s expected excess length. Results are averaged over 1000 simulations for each of the 12 papers surveyed, and the median across papers is reported here. See Section 8 for details.
H.2 Non-normal simulation results with estimated covariance matrix

In Section 8, we present simulations results where $\hat{\beta}$ is normally distributed and its covariance matrix is treated as known. In this section, we present Monte Carlo results using a data-generating process in which $\hat{\beta}$ is not normally distributed and the covariance matrix is estimated from the data. Specifically, we consider considerations based on the empirical distribution in Bailey and Goodman-Bacon (2015). We find that all of our procedures achieve (approximate) size control, and our results on the relative power of the various procedures are quite similar to those presented in Section 8.

H.2.1 Simulation design

The simulations are calibrated using the empirical distribution of the data in Bailey and Goodman-Bacon (2015).

Let $\hat{\beta}$, $\hat{\Sigma}$ denote the original, estimated event-study coefficients and variance-covariance matrix from the event-study regression in the paper. We simulate data using a clustered bootstrap sampling scheme at the county level (i.e., the level of clustering used by the authors in their event-study regression). For each bootstrap sample $b$, we re-estimate the event-study coefficients $\hat{\beta}_b$ and the variance-covariance matrix $\hat{\Sigma}_b$ also using the clustering scheme specified by the authors. We then re-center the bootstrapped coefficient so that under our simulated data-generating process parallel trends holds, $\hat{\beta}_{b,\text{centered}} = \hat{\beta}_b - \hat{\beta}$. We then construct our proposed confidence sets for bootstrap draw $b$ using the pair $(\hat{\beta}_{b,\text{centered}}, \hat{\Sigma}_b)$.

As in Section 8, we focus on three choices of $\Delta$ to highlight the performance of the proposed confidence sets under a range of conditions: $\Delta^{SD}(M), \Delta^{SDPB}(M)$ and $\Delta^{SDI}(M)$. The parameter of interest in these simulations is the causal effect in the first post-period ($\theta = \tau_1$). We report the performance of the FLCI, conditional confidence set, conditional-FLCI hybrid confidence set and the conditional least-favorable confidence set. All results are averaged over 1000 bootstrap samples.

H.2.2 Size control simulations

Table 2 reports the maximum rejection rate of each procedure over a grid of parameter values $\theta$ within the identified set $S(\Delta, 0)$. We report results for each choice of $\Delta$ and $M = 1, 2, 3, 5$. The table shows that all our procedures approximately control size, with null rejection rates never substantially exceeding the nominal rate of 0.05.

---

47 Since implementing the bootstrap in practice is logistically challenging, we do so for one paper rather than the full 12 papers in the survey. We chose the first paper alphabetically to minimize concerns about cherry-picking.
<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>$M$</th>
<th>Conditional</th>
<th>FLCI</th>
<th>C-F Hybrid</th>
<th>C-LF Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta^{SD}(M)$</td>
<td>0</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.05</td>
<td>0.07</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>$\Delta^{SDP}(M)$</td>
<td>0</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.05</td>
<td>0.08</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>$\Delta^{SDI}(M)$</td>
<td>0</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.06</td>
<td>0.04</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.06</td>
<td>0.04</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.07</td>
<td>0.05</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table 2: Maximum null rejection probability over the identified set using the empirical distribution from Bailey and Goodman-Bacon (2015).
H.2.3 Comparison with normal simulations

We next compare results from the non-normal simulations with estimated covariance discussed above to the simulations in Section 8, in which $\hat{\beta}$ is normal and $\Sigma$ is treated as known. Figure I7 shows the rejection probabilities at different values of the parameter $\theta$ using both simulation methods. Specifically, we plot results for each choice of $\Delta$ using $M = 0$ and $M = 5$. (The results are quite similar for all values of $M$ considered, and we thus omit the intermediate values to preserve space). As can be seen, the estimated average rejection rates of each procedure are quite similar in the non-normal simulations and the normal simulations across each choice of $\Delta$. As a result, the relative rankings the procedures in terms of power are the same in the non-normal simulations as in the normal simulations discussed in Section 8.
Figure I6: Comparison of rejection probabilities using bootstrap and normal simulations. Results are shown for $\theta = \tau_1$, and each choice of $\Delta = \Delta^{SD}(M), \Delta^{SDPB}(M), \Delta^{SDI}(M)$, and $M = 0$. The average rejection rate for the non-normal simulations are in red and the average rejection rate for the normal simulations are in blue; the dashed black lines indicate the identified set bounds. Results are averaged over 1000 simulations.
Figure I7: Comparison of rejection probabilities using bootstrap and normal simulations. Results are shown for $\theta = \tau_1$, and each choice of $\Delta = \Delta^{SD}(M), \Delta^{SDPB}(M), \Delta^{SDI}(M)$, and $M = 5$. The average rejection rate for the non-normal simulations are in red and the average rejection rate for the normal simulations are in blue; the dashed black lines indicate the identified set bounds. Results are averaged over 1000 simulations.
Supplement References


