A Proofs

A.1 Notation, regularity conditions, and a note on asymptotics

Before presenting the regularity conditions and additional notation to facilitate the proofs, a clarification regarding the theorems in the main article is warranted. Theorems 1-5 are presented conditional upon Neyman’s null $H_N$ being true, such that $n^{-1} \sum_{i=1}^{n} \bar{\tau}_i = 0$; however, for Neyman’s null to be true at each point along an asymptotic sequence it would have to be the case that $\bar{\tau}_i = 0$ for all pairs. For precision, the asymptotics should instead reflect a sequence of observational studies of increasing sample size. For each number of pairs $n$ the treatment effects in each pair $\bar{\tau}_i$ should be adaptively re-centered by the sample average treatment effect in the first $n$ pairs, call it $\bar{\tau}^{(n)}$, such that Neyman’s null would hold after the re-centering. For instance, for each $n$ the random variable $D_{i,\Gamma}$ should be re-defined as

$$D_{i,\Gamma}^{(n)} = \hat{\tau}_i - \bar{\tau}^{(n)} - (2\theta_{\Gamma} - 1)|\hat{\tau}_i - \bar{\tau}^{(n)}|$$
when considering the results in Theorems 1-5. We have omitted this in the text and in the
proofs in this web supplement, trading precision for notational convenience and enhanced
readability.

The following regularity conditions are imposed throughout the proofs that follow. As
a reminder, the quantity $\eta_i$ is defined as in the main article to represent the difference in
the averages of the potential outcomes for the two individuals in a given pair,

$$\eta_i = \frac{Y_{i1}(0) + Y_{i1}(1) - Y_{i2}(0) - Y_{i2}(1)}{2},$$

such that the treated-minus-control paired difference in pair $i$ is

$$\hat{\tau}_i = \bar{\tau}_i + (T_{i1} - T_{i2})\eta_i.$$

**Condition 1.** There exist constants $C > 0$, $\mu_m$ and $\mu_a$ such that as $n \to \infty$

$$n^{-1} \sum_{i=1}^{n} |\eta_i| > C, \quad n^{-1} \sum_{i=1}^{n} \eta_i^2 > C, \quad (1)$$

$$n^{-2} \sum_{i=1}^{n} \eta_i^2 \to 0, \quad n^{-2} \sum_{i=1}^{n} \eta_i^4 \to 0, \quad n^{-2} \sum_{i=1}^{n} \bar{\tau}_i^4 \to 0, \quad (2)$$

$$n^{-1} \sum_{i=1}^{n} (2\pi_i - 1)\eta_i \to \mu_m, \quad n^{-1} \sum_{i=1}^{n} \pi_i |\bar{\tau}_i + \eta_i| + (1 - \pi_i) |\bar{\tau}_i - \eta_i| \to \mu_a. \quad (3)$$

**Condition 2.** There exists a constant $\nu^2 > 0$ such that

$$n^{-1} \sum_{i=1}^{n} \pi_i (\bar{\tau}_i + \eta_i)^2 + (1 - \pi_i) (\bar{\tau}_i - \eta_i)^2 \to \nu^2. \quad (4)$$
The sensitivity model of Rosenbaum (1987) states
\[
\text{pr}(T = t \mid F, T) = \prod_{i=1}^{n} \pi_i^{t_i}(1 - \pi_i)^{1-t_i}, \quad \frac{1}{1+\Gamma} \leq \pi_i \leq \frac{\Gamma}{1+\Gamma} \quad (i = 1, \ldots, n). \tag{5}
\]

In the proofs that follow, let \(\theta = \Gamma/(1 + \Gamma)\), such that if the sensitivity model holds at \(\Gamma\) we have that \(1 - \theta \leq \pi_i \leq \theta\) and that \(2\theta - 1 = (\Gamma - 1)/(1 + \Gamma)\). Further, all results should be viewed as conditional upon \(F\) and \(T\), with \(F\) and \(T\) growing with \(n\); this has been omitted for enhanced readability.

### A.2 Theorem 1

Proof of a detail in Lemma 2

We now show that \(\theta \mid \bar{\tau}_i + |\eta_i| + (1-\theta) |\bar{\tau}_i - |\eta_i| \geq |\eta_i| + (2\theta - 1) \bar{\tau}_i\). We do this in three cases depending upon the values for \(\text{sign}(\bar{\tau}_i + |\eta_i|)\) and \(\text{sign}(\bar{\tau}_i - |\eta_i|)\)

**Case 1** \((\bar{\tau}_i + |\eta_i| \geq 0, \bar{\tau}_i - |\eta_i| \geq 0)\). Here \(\bar{\tau}_i \geq |\eta_i|\). Recalling that \(0 \leq 2\theta - 1 \leq 1\),

\[
\theta \mid \bar{\tau}_i + |\eta_i| + (1-\theta) |\bar{\tau}_i - |\eta_i| = (2\theta - 1) |\eta_i| + \bar{\tau}_i \\
\geq |\eta_i| + (2\theta - 1) \bar{\tau}_i.
\]

**Case 2** \((\bar{\tau}_i + |\eta_i| \geq 0, \bar{\tau}_i - |\eta_i| < 0)\). Here we have that the result holds with equality, as

\[
\theta \mid \bar{\tau}_i + |\eta_i| + (1-\theta) |\bar{\tau}_i - |\eta_i| = |\eta_i| + (2\theta - 1) \bar{\tau}_i.
\]

**Case 3** \((\bar{\tau}_i + |\eta_i| < 0, \bar{\tau}_i - |\eta_i| < 0)\). In this case \(-\bar{\tau}_i \geq |\eta_i|\). Noting \(-\bar{\tau}_i = -2\theta \bar{\tau}_i + (2\theta - 1) \bar{\tau}_i\)
and that $2\theta_\Gamma, (2\theta_\Gamma - 1) \geq 0$,

$$\theta_\Gamma |\bar{\tau}_i + |\eta_i|| + (1 - \theta_\Gamma)|\bar{\tau}_i - |\eta_i|| = (1 - 2\theta_\Gamma)|\eta_i| - \bar{\tau}_i$$

$$\geq |\eta_i| + (2\theta_\Gamma - 1)\bar{\tau}_i.$$  

The inequality thus always holds.

**Proof of Lemma 3**

For any constant $\Gamma \geq 1$

$$E(se(\bar{D}_\Gamma)^2) - \text{var}(\bar{D}_\Gamma)$$

$$= \frac{1}{n(n-1)} \sum_{i=1}^{n} \left\{ E(D_{i,\Gamma}) - E(\bar{D}_\Gamma) \right\}^2 \geq 0.$$  

**Proof.**

$$E(se(\bar{D}_\Gamma)^2) = \frac{1}{n(n-1)} \left\{ \sum_{i=1}^{n} E(D_{i,\Gamma}^2) - n^{-1} \sum_{k,\ell=1}^{n} E(D_{k,\Gamma}D_{\ell,\Gamma}) \right\}$$

$$= n^{-2} \left\{ \sum_{i=1}^{n} E(D_{i,\Gamma}^2) - \frac{1}{n-1} \sum_{k\neq\ell} E(D_{k,\Gamma})E(D_{\ell,\Gamma}) \right\}$$

$$= n^{-2} \left[ \sum_{i=1}^{n} \left\{ \text{var}(D_{i,\Gamma}) + E(D_{i,\Gamma})^2 \right\} - \frac{1}{n-1} \sum_{k\neq\ell} E(D_{k,\Gamma})E(D_{\ell,\Gamma}) \right]$$

$$= \text{var}(\bar{D}_\Gamma) + \frac{1}{n(n-1)} \sum_{i=1}^{n} \left\{ E(D_{i,\Gamma}) - E(\bar{D}_\Gamma) \right\}^2,$$

proving the result.

**Remark 1.** The result of Lemma 3 applies beyond the collection of random variables $\{D_{i,\Gamma}\}$. Take any collection of $n$ independent random variables $\{X_i\}$ with $E(X_i) = \mu_i$
and \( \text{var}(X_i) = \sigma_i^2 \), and consider their random average \( \bar{X} \). Then, \( E(se(\bar{X})^2) - \text{var}(\bar{X}) = ((n-1)n^{-1}\sum_{i=1}^{n}(\mu_i - \bar{\mu})^2 \geq 0. \)

**Lemma 5.** For each \( i \),

\[
16\theta(1 - \theta)^3\eta_i^2 \leq \text{var}(U_{i,\Gamma}) \leq 16\theta^3(1 - \theta)\eta_i^2 
\]

\[
E(U_{i,\Gamma}^4) \leq 128\theta^4_i \left( \eta_i^4 + \tilde{\eta}_i^4 \right) 
\]

Further, if treatment assignment satisfies (5) at \( \Gamma \),

\[
16\theta(1 - \theta)^3\eta_i^2 \leq \text{var}(D_{i,\Gamma}) \leq 4\theta^2\eta_i^2 
\]

\[
E(D_{i,\Gamma}^4) \leq 128\theta^4_i \left( \eta_i^4 + \tilde{\eta}_i^4 \right) 
\]

**Proof.** To prove (6), observe that \( \text{var}(U_{i,\Gamma}) = \theta(1-\theta)(2\eta_i - (2\theta - 1)(|\bar{\tau}_i + \eta_i| - |\bar{\tau}_i - \eta_i|))^2 \), which is at least \( 16\theta(1 - \theta)^3\eta_i^2 \) and at most \( 16\theta^3(1 - \theta)\eta_i^2 \). The proof of (8) simply replaces \( \theta(1 - \theta) \) with \( 1/4 \) in the upper bound.

Proving (7) requires multiple applications of \((a+b)^2 \leq 2a^2 + 2b^2\) for scalars \( a \) and \( b \).

Without loss of generality assume that \( \eta_i \geq -\eta_i \).

\[
E(U_{i,\Gamma}^4) = \theta(\bar{\tau}_i + \eta_i - (2\theta - 1)|\bar{\tau}_i + \eta_i|)^4 + (1 - \theta)((\bar{\tau}_i - \eta_i - (2\theta - 1)|\bar{\tau}_i - \eta_i|)^4
\]

\[
\leq \theta(2\theta(\bar{\tau}_i + \eta_i))^4 + (1 - \theta)(2\theta(\bar{\tau}_i - \eta_i))^4
\]

\[
\leq 128\theta^4_i \left( \eta_i^4 + \tilde{\eta}_i^4 \right) 
\]

The proof of (9) is analogous.
Lemma 6. Both $n^{1/2} \bar{D}_\Gamma$ and $n^{1/2} \bar{U}_\Gamma$ are asymptotically normal. Further, let

$$k_\Gamma^*(\alpha) = \Phi^{-1}(1 - \alpha) \left\{ n^{-2} \sum_{i=1}^n \text{var}(U_{i,\Gamma}) \right\}^{1/2}. \quad (10)$$

Then, if (5) holds at $\Gamma$ and $H_N$ is true,

$$\lim_{n \to \infty} \text{pr}\{ \bar{D}_\Gamma \geq k_\Gamma^*(\alpha) \} \leq \alpha. \quad (11)$$

Proof. We prove asymptotic normality of $n^{1/2} \bar{U}_\Gamma$, and with it (11) by reference to Lemma 1; the proof for $n^{1/2} \bar{D}_\Gamma$ is analogous. The $U_{i,\Gamma}$ are conditionally independent given $\mathcal{F}$ and $\mathcal{T}$. Further, by Lemma 2 and sharpness of $U_{i,\Gamma}$ as a stochastic upper bound we have that $E(n^{-1} \sum_{i=1}^n U_{i,\Gamma}) \leq 0$. To prove asymptotic normality of $n^{1/2} \bar{U}_\Gamma$, it suffices to show that Lyapunov’s condition holds for $\delta = 2$, i.e. that

$$\sum_{i=1}^n E|U_{i,\Gamma} - E(U_{i,\Gamma})|^4 / \left( \sum_{i=1}^n \text{var}(U_{i,\Gamma}) \right)^2 \to 0$$

By (6), $n^{-1} \sum_{i=1}^n \text{var}(U_{i,\Gamma}) \geq 16\theta_\Gamma (1 - \theta_\Gamma)^3 n^{-1} \sum_{i=1}^n \eta_i^2$, which is greater than $16\theta_\Gamma (1 - \theta_\Gamma)^3 C$ for some $C > 0$ as $n \to \infty$ by (1). Applying Jensen’s inequality and utilizing (7) and (2), we have that $n^{-2} \sum_{i=1}^n E|U_{i,\Gamma} - E(U_{i,\Gamma})|^4 \to 0$. Hence,

$$\sum_{i=1}^n E|U_{i,\Gamma} - E(U_{i,\Gamma})|^4 / \left( \sum_{i=1}^n \text{var}(U_{i,\Gamma}) \right)^2 = n^{-2} \sum_{i=1}^n E|U_{i,\Gamma} - E(U_{i,\Gamma})|^4 / \left( n^{-1} \sum_{i=1}^n \text{var}(U_{i,\Gamma}) \right)^2$$

$$\leq n^{-2} \sum_{i=1}^n E|U_{i,\Gamma} - E(U_{i,\Gamma})|^4 / (16\theta_\Gamma (1 - \theta_\Gamma)^3 C)^2 \to 0.$$

This, along with Lemma 1, proves the result. 

Lemma 7. Suppose that treatment assignment satisfies (5) at $\Gamma$ and $H_N$ holds. If (1)
and (2) hold, then for all $\epsilon > 0$, as $n \to \infty$

\[
pr\left(-\epsilon + \bar{D}_T \geq 0\right) \to 0. \tag{12}
\]

\[
pr\left\{\epsilon + n\text{se}(\bar{D}_T)^2 \leq n^{-1}\sum_{i=1}^{n} \text{var}(U_{i,T})\right\} \to 0 \tag{13}
\]

**Proof.** We begin by proving (12). By Lemma 2,\[ pr\left(-\epsilon + \bar{D}_T \geq 0\right) \leq pr\left(-\epsilon + \bar{D}_T - E(\bar{D}_T) \geq 0\right). \]
The variance of $\text{var}(D_{i,T})$ is, by (8), less than $4\theta_T^2\eta_i^2$. Therefore, using (2),

\[
\text{var}(\bar{D}_T) \leq 4\theta_T^2 n^{-2} \sum_{i=1}^{n} \eta_i^2 \to 0
\]
as $n \to \infty$. Chebyshev’s inequality then yields (12).

We now prove (13). Recall that $n\text{se}(\bar{D}_T)^2 = (n-1)^{-1}\sum_{i=1}^{n} D_{i,T}^2 - n/(n-1)(\bar{D}_T)^2$. By Lemma 4,

\[
pr\left\{\epsilon + n\text{se}(\bar{D}_T)^2 \leq n^{-1}\sum_{i=1}^{n} \text{var}(U_{i})\right\}
\leq pr\left\{\epsilon + n\text{se}(\bar{D}_T)^2 \leq (n-1)^{-1}\sum_{i=1}^{n} \text{var}(D_{i,T}) + (n-1)^{-1}\sum_{i=1}^{n} (E(D_{i,T}) - E(\bar{D}_T))^2\right\}
\]

\[
= pr\left\{\epsilon + (n-1)^{-1} \left(\sum_{i=1}^{n} D_{i,T}^2 - \sum_{i=1}^{n} E(D_{i,T}^2)\right) - n(n-1)^{-1}(\bar{D}_T^2 - E(\bar{D}_T)^2) \leq 0\right\}
\]

The proof of (12) along with (3) yields that $\bar{D}_T^2 - E(\bar{D}_T)^2$ converges in probability to 0. We now show that $(n-1)^{-1} \left\{\sum_{i=1}^{n} D_{i,T}^2 - \sum_{i=1}^{n} E(D_{i,T}^2)\right\}$ also converges in probability to 0. Using (9),
\[
\text{var}\left\{ (n - 1)^{-1} \sum_{i=1}^{n} D_{i,\Gamma}^2 \right\} \leq (n - 1)^{-2} \sum_{i=1}^{n} E(D_{i,\Gamma}^4) \\
\leq 128 \theta_{\Gamma}^4 (n - 1)^{-2} \sum_{i=1}^{n} (\hat{\tau}_i^4 + \eta_i^4),
\]
which converges to 0 as \( n \to \infty \) through (2). Applying Chebyshev’s inequality yields the desired convergence in probability, which in turn yields (13). \( \square \)

**Proof of Theorem 1**

Define \( k_{\Gamma}(\alpha) = se(\bar{D}_{\Gamma})\Phi^{-1}(1 - \alpha) \) with \( 0 < \alpha \leq 0.5 \). By (13), taking \( \epsilon \downarrow 0 \),

\[
\lim_{n \to \infty} \text{pr}\{ k_{\Gamma}(\alpha) \geq k^*_\Gamma(\alpha) \} = 1.
\]

This, in combination with (11), yields the conclusion of the theorem.

**A.3 Theorem 2**

**Lemma 8.** Take a vector \( V_{\Gamma} \) distributed as in §2.3 with \( V_{i,\Gamma} = \pm 1 \) and \( \text{pr}(V_{i,\Gamma} = 1) = \theta_{\Gamma} \). Let \( V'_{\Gamma} \) be an iid copy of \( V_{\Gamma} \). Then, under (2) and (4), \( n^{1/2} \bar{B}_{\Gamma}(V_{\Gamma}, \hat{\tau}) \) and \( n^{1/2} \bar{B}_{\Gamma}(V'_{\Gamma}, \hat{\tau}) \) are iid and converge jointly to a bivariate normal, each with mean zero and variance \( 4\theta_{\Gamma}(1 - \theta_{\Gamma})\nu^2 \).

**Proof.** Recall that \( B_{i,\Gamma} = V_{i,\Gamma}|\hat{\tau}_i| - (2\theta_{\Gamma} - 1)|\hat{\tau}_i| \) and that \( E(B_{i,\Gamma}) = 0 \). Since uncorrelatedness implies independence for the normal, to show independence of the limiting
distributions for $\bar{B}_\Gamma$ and $\bar{B}'_\Gamma$ it suffices to show that $\text{cov}(\bar{B}_\Gamma, \bar{B}'_\Gamma) = 0$.

$$\text{cov}(\bar{B}_\Gamma, \bar{B}'_\Gamma) = E\left\{\text{cov}(\bar{B}_\Gamma, \bar{B}'_\Gamma \mid \hat{\tau})\right\} + \text{cov}\left\{E(\bar{B}_\Gamma \mid \hat{\tau}), E(\bar{B}'_\Gamma \mid \hat{\tau})\right\}$$

$$= n^{-2}E\left\{\sum_{i=1}^{n} \hat{\tau}_i^2 \text{cov}(V_{i,\Gamma}, V'_{i,\Gamma})\right\} + 0$$

$$= 0$$

By the Cramér-Wold device, to show bivariate asymptotic normality it suffices to show that $n^{1/2}(w_1\bar{B}_\Gamma + w_2\bar{B}'_\Gamma)$ converges to a normal with mean zero and variance $(w_1^2 + w_2^2)4\theta_\Gamma(1 - \theta_\Gamma)\nu^2$ for any vector of constants $(w_1, w_2)$. Fixing $(w_1, w_2)$, we now show this to be the case through Lyapunov’s condition. We have $E(w_1B_{i,\Gamma} + w_2B'_{i,\Gamma}) = 0$, that $E(B_{i,\Gamma}^4) = \pi_i(\bar{\tau}_i + \eta_i)^4 + (1 - \pi_i)(\bar{\tau}_i - \eta_i)^4 \leq 8\bar{\tau}_i^4 + 8\eta_i^4$, and that $E((w_1B_{i,\Gamma} + w_2B'_{i,\Gamma})^4) \leq 8(w_1^4 + w_2^4)E(B_{i,\Gamma}^4)$. Combining this with (2), we have that $n^{-2}\sum_{i=1}^{n} E((w_1B_{i,\Gamma} + w_2B'_{i,\Gamma})^4) \to 0$. By (4), we have that $n^{-1}\sum_{i=1}^{n} \text{var}(w_1B_{i,\Gamma} + w_2B'_{i,\Gamma}) \to (w_1^2 + w_2^2)4\theta(1 - \theta)\nu^2 > 0$.

Hence,

$$\sum_{i=1}^{n} E((w_1B_{i,\Gamma} + w_2B'_{i,\Gamma})^4)/\left(\sum_{i=1}^{n} \text{var}(w_1B_{i,\Gamma} + w_2B'_{i,\Gamma})\right)^2$$

$$= n^{-2} \sum_{i=1}^{n} E((w_1B_{i,\Gamma} + w_2B'_{i,\Gamma})^4)/\left(n^{-1}(w_1^2 + w_2^2)\sum_{i=1}^{n} \text{var}(B_{i,\Gamma})\right)^2 \to 0.$$ 

Lyapunov’s condition is satisfied at $\delta = 2$, proving the result.

\[\square\]

**Lemma 9.** Under the assumptions of Theorem 2, for any point $a$

$$\hat{F}_\Gamma(a/n^{1/2}) \xrightarrow{p} \Phi(a/\nu_\Gamma),$$

9
where $\nu_\Gamma^2 = 4\theta_\Gamma(1 - \theta_\Gamma)\nu^2$

Proof. Observe that

$$E(\hat{F}_\Gamma(a/n^{1/2})) = E(E(1\{n^{1/2}\bar{B}_\Gamma(V_\Gamma, \hat{\tau}) \leq a\} |))$$

$$= E(1\{n^{1/2}\bar{B}_\Gamma(V_\Gamma, \hat{\tau}) \leq a\})$$

$$= \text{pr}(n^{1/2}\bar{B}_\Gamma(V_\Gamma, \hat{\tau}) \leq a)$$

By Lemma 8, $n^{1/2}\bar{B}_\Gamma(V_\Gamma, \hat{\tau})$ converges in distribution to a normal with mean 0 and variance $4\theta_\Gamma(1 - \theta_\Gamma)\nu^2$. Hence, $E(\hat{F}_\Gamma(a/n^{1/2})) \rightarrow \Phi(a/\nu_\Gamma)$. Through Chebyshev’s inequality, to illustrate the desired convergence in probability it suffices to show that $E(\hat{F}^2_\Gamma(a/n^{1/2})) \rightarrow \Phi^2(a/\nu_\Gamma)$, which is equivalent to $\text{var}(\hat{F}_\Gamma(a/n^{1/2})) \rightarrow 0$.

$$E(\hat{F}^2_\Gamma(a/n^{1/2}))$$

$$= E \left[ \sum_{t, t' \in \Omega} 1\{n^{1/2}\bar{B}_\Gamma(t_1 - t_2, \hat{\tau}) \leq a\} 1\{n^{1/2}\bar{B}_\Gamma(t'_1 - t'_2, \hat{\tau}) \leq a\} \prod_{i=1}^n \theta_t^{i_1 + i'_1}(1 - \theta_\Gamma)^2 - t_{i_1} - t'_{i_1} \right]$$

$$= \text{pr}(n^{1/2}\bar{B}_\Gamma(V_\Gamma, \hat{\tau}) \leq a, n^{1/2}\bar{B}_\Gamma(V'_\Gamma, \hat{\tau}) \leq a) \rightarrow \Phi^2(a/\nu_\Gamma)$$

as desired, where the last line uses Lemma 8. $\square$

Lemma 10.

$$E\{se(\bar{B}_\Gamma)^2\} = \text{var}(\bar{B}_\Gamma)$$

Proof. The lemma follows by Remark 1 along with the fact that by construction $B_{i, \Gamma}$ is centered, such that $E(\bar{B}_{i, \Gamma}) = 0$. $\square$
Lemma 11. Under the assumptions of Theorem 2,

\[ \text{lse}(\bar{B}_\Gamma)^2 \overset{p}{\to} 4\theta_\Gamma(1 - \theta_\Gamma)\nu^2 \]

Proof. Decompose \( \text{lse}(\bar{B}_\Gamma)^2 = (n - 1)^{-1}\sum_{i=1}^n B_{i,\Gamma}^2 - n/(n - 1)\bar{B}_\Gamma \). \( E(\bar{B}_\Gamma) \) is 0, while by Lemma 8 \( \bar{B}_\Gamma \) has limiting variance \( 4\theta_\Gamma(1 - \theta_\Gamma)\nu^2/n \to 0 \). Hence, \( n/(n - 1)\bar{B}_\Gamma \) converges in probability to 0 by Chebyshev’s inequality. Meanwhile, \( E((n - 1)^{-1}\sum_{i=1}^n B_{i,\Gamma}^2) \to 4\theta_\Gamma(1 - \theta_\Gamma)\nu^2 \) by Lemma 10 and (4). To show that the variance of this term goes to zero, observe that \( B_{i,\Gamma}^4 \leq 8(1 + (2\theta_\Gamma - 1)^4)\tilde{\tau}_i^4 \). Similar arguments to those of Lemma 7, utilizing (2), then yield that the variance goes to zero, thus yielding the result through Chebyshev’s inequality.

Proof of Theorem 2

By Lemma 9, \( \hat{F}_\Gamma(a/n^{1/2}) \), the biased randomization distribution of \( n^{1/2}\bar{B}_\Gamma \), converges in probability to \( \Phi(a/\nu_\Gamma) \) for all points \( a \), where again \( \nu_\Gamma^2 = 4\theta_\Gamma(1 - \theta_\Gamma)\nu^2 \). By Lemma 11 and the continuous mapping theorem \( n^{1/2}\text{se}(\bar{B}_\Gamma) \) converges in probability to \( \nu_\Gamma \). Recall that \( \hat{G}_\Gamma(t) \) is the biased randomization distribution of the studentized statistic \( n^{1/2}\bar{B}_\Gamma/\{n^{1/2}\text{se}(\bar{B}_\Gamma)\} \).

Setting \( a = tn^{1/2}\text{se}(\bar{B}_\Gamma) \) and using Slutsky’s theorem for randomization distributions (Chung and Romano, 2013, Lemma 5.2), we have that \( \hat{G}_\Gamma(t) \) then converges in probability to \( \Phi(t\nu_\Gamma/\nu_\Gamma) = \Phi(t) \) for all points \( t \) as desired.

A.4 Proof of Theorem 3

We first prove exactness of \( \varphi_{S+}(\alpha, \Gamma) \) under \( H_F \). Re-arrange the pairs such that the first individual in each pair has the larger response. Define \( q_{i1} = 2(1 - \theta_\Gamma)|\eta_i| \) and \( q_{i2} = -2\theta_\Gamma|\eta_i| \), and recall that \( \theta_\Gamma \geq 1 - \theta_\Gamma \). For any treatment assignment \( t \), the positive
part statistic can be expressed as

\[ f(t, q) = \max \left\{ 0, -\frac{n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{2} q_{ij} t_{ij}}{\sqrt{\frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{2} t_{ij} \left( q_{ij} - n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{2} q_{ij} t_{ij} \right)^2}} \right\} \]  

(14)

Let the vector \( q[i12] \) equal the vector obtained from \( q \) by exchanging the first and second elements in pair \( i \) while leaving the other elements fixed, such that \( (q[i12])_{i2} = q_{i1} \). A function \( h(t, q) \) is called an arrangement increasing function in pairs if for all pairs \( i \) \( h(t, q) \geq h(t, q[i12]) \) whenever \( (t_{i1} - t_{i2})(q_{i1} - q_{i2}) \geq 0 \) (Rosenbaum, 2002, §2.4.4). In words, this says the function \( h \) takes on a larger value when the elements \( t \) and \( q \) are arranged in the same order within a pair than it does when they are out of order.

We now show that \( f(t, q) \) in (14) is arrangement increasing in pairs. We do so for the \( n \)th pair without loss of generality. For each \( t \in \Omega \), let \( d_i = t_{i1}q_{i1} + t_{i2}q_{i2} \). Consider fixed values for \( d_1, \ldots, d_{n-1} \) and consider the two possibilities for \( d_n \), either \( d_n = 2(1 - \theta_T)|\eta_n| \) or \( d_n = -2\theta_T|\eta_n| \). It suffices to show that the function \( f(t, q) \) is at least as large when \( d_n = 2(1 - \theta_T)|\eta_n| \) as it is when \( d_n = -2\theta_T|\eta_n| \) for any fixed values of \( d_1, \ldots, d_{n-1} \).

If \( \sum_{i=1}^{n-1} d_i \leq 0 \), then this is trivially true, as the test statistic will either be positive when \( d_n = 2(1 - \theta_T)|\eta_n| \) and zero otherwise, or will be zero in both cases due to the positive part modification. We thus restrict attention the case \( \sum_{i=1}^{n-1} d_i \geq 0 \). As the numerator of \( f(t, q) \) would be larger when \( d_n = 2(1 - \theta_T)|\eta_n| \); it is enough to show that the denominator will be smaller when \( d_n = 2(1 - \theta_T)|\eta_n| \) than it would be if \( d_n = -2\theta_T|\eta_n| \) when \( \sum_{i=1}^{n-1} d_i \geq 0 \). Algebra yields that this is true if and only if \( 4(n-1)/n(1-\theta_T)^2\eta_n^2 - 4/n(1-\theta_T)|\eta_n| \sum_{i=1}^{n-1} d_i \leq 4(n-1)/n(\theta_T)^2\eta_n^2 + 4/n(1-\theta_T)|\eta_n| \sum_{i=1}^{n-1} d_i \), which holds as \( \sum_{i=1}^{n-1} d_i \geq 0 \) and \( \theta_T \geq (1-\theta_T) \).

The function \( f(t, q) \) is thus arrangement increasing over randomizations in \( \Omega \), and
the proof of exactness of $\varphi_{S+}$ under Fisher's sharp null follows by applying Theorem 2 of Rosenbaum (1987). The asymptotic correctness of $\varphi_{S+}$ under Neyman's weak null follows in a straightforward way from Theorem 2, and the proof is omitted.

A.5 Proof of Theorem 5

Note that $\theta_{\Gamma+\epsilon} > \theta_\Gamma$ for any $\epsilon > 0$. Consider $D_{i,\Gamma+\epsilon} = \hat{r}_i - (2\theta_{\Gamma+\epsilon} - 1)|\hat{r}_i|$. Since (5) holds at $\Gamma$ under the Theorem's conditions, by arguments parallel to those in Lemma 1 $D_{\Gamma+\epsilon}$ is stochastically bounded by the random variable $W_{\Gamma,\epsilon}$, where

$$W_{i,\Gamma,\epsilon} = \hat{r}_i + V_{i,\Gamma,\epsilon} |\eta_i| - (2\theta_{\Gamma+\epsilon} - 1)\{(1 + V_{i,\Gamma,\epsilon})|\hat{r}_i| + |\eta_i| + (1 - V_{i,\Gamma})|\hat{r}_i| - |\eta_i|\}/2.$$ 

Hence, $E(W_{\Gamma,\epsilon}) \geq E(D_{\Gamma+\epsilon})$. Further define $U_{i,\Gamma+\epsilon}$ as before, namely

$$U_{i,\Gamma+\epsilon} = \hat{r}_i + V_{i,\Gamma+\epsilon} |\eta_i| - (2\theta_{\Gamma+\epsilon} - 1)\{(1 + V_{i,\Gamma+\epsilon})|\hat{r}_i| + |\eta_i| + (1 - V_{i,\Gamma+\epsilon})|\hat{r}_i| - |\eta_i|\}/2.$$ 

We now show that even the limit, $E(U_{\Gamma+\epsilon}) > E(W_{\Gamma,\epsilon})$.

$$E(U_{\Gamma+\epsilon}) - E(W_{\Gamma,\epsilon})$$

$$= n^{-1} \sum_{i=1}^{n} \left\{ (\theta_{\Gamma+\epsilon} - \theta_\Gamma) \{|\eta_i| - (2\theta_{\Gamma+\epsilon} - 1)|\hat{r}_i| + |\eta_i|\} \right.$$ 

$$+ (\theta_\Gamma - \theta_{\Gamma+\epsilon}) \{ -|\eta_i| + (2\theta_{\Gamma+\epsilon} - 1)|\hat{r}_i| - |\eta_i|\} \right\}$$

$$= (\theta_{\Gamma+\epsilon} - \theta_\Gamma)n^{-1} \sum_{i=1}^{n} \left\{ 2|\eta_i| + (2\theta_{\Gamma+\epsilon} - 1)(|\hat{r}_i| - |\eta_i| - |\hat{r}_i| + |\eta_i|) \right\}$$

$$\geq 4(1 - \theta_{\Gamma+\epsilon})(\theta_{\Gamma+\epsilon} - \theta_\Gamma)n^{-1} \sum_{i=1}^{n} |\eta_i|.$$
where the last line follows arguments similar to those used to prove (6). In the limit, the last line is greater than or equal to $4(1 - \theta \Gamma) (\theta \Gamma - \theta \Gamma) C > 0$ by (1). Hence, if (5) holds at $\Gamma$ but a sensitivity analysis is conducted at $\Gamma + \epsilon$, $E(\bar{D}_{\Gamma+\epsilon})$ is strictly less than $E(\bar{U}_{\Gamma+\epsilon})$ asymptotically, which is itself less than or equal to zero if the average treatment effect equals $\tau$ by Lemma 2.

Let $\mu_D = E(\bar{D}_{\Gamma+\epsilon}) < 0$, $\sigma_D^2/n = \text{var}(\bar{D}_{\Gamma+\epsilon})$, and $\nu^2 = 4\theta (1 - \theta \Gamma) \nu^2$. Asymptotically, the unstudentized procedure rejects if $n^{1/2} \bar{D}_{\Gamma+\epsilon} \geq \nu_{\Gamma+\epsilon} \Phi^{-1}(1 - \alpha)$ by Theorem 2.

$$
\lim_{n \to \infty} E\{\varphi_F(\alpha, \Gamma + \epsilon) \mid H_N\} = \lim_{n \to \infty} \text{pr}\{n^{1/2} \bar{D}_{\Gamma+\epsilon} \geq \nu_{\Gamma+\epsilon} \Phi^{-1}(1 - \alpha) \mid H_N\} = \lim_{n \to \infty} \text{pr}\{n^{1/2}(D_{\Gamma+\epsilon} - \mu_D)/\sigma_D \geq (\Phi^{-1}(1 - \alpha)\nu_{\Gamma+\epsilon} - n^{1/2}\mu_D)/\sigma_D \mid H_N\} = \lim_{n \to \infty} 1 - \Phi\left\{\frac{-n^{1/2}\mu_D + \Phi^{-1}(1 - \alpha)\nu_{\Gamma+\epsilon}}{\sigma_D}\right\} = 0,
$$

where the last line stems from $\mu_D < 0$ and asymptotic normality of $n^{1/2} \bar{D}_\Gamma$ by Lemma 6.

**References**

Chung, E. and Romano, J. P. (2013). Exact and asymptotically robust permutation tests. *The Annals of Statistics*, 41(2):484–507.

Rosenbaum, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika*, 74(1):13–26.

Rosenbaum, P. R. (2002). *Observational Studies*. Springer, New York.