Online appendices

A Proofs

A.1 Derivation of MSE(SATE,SATT)

$$
\begin{aligned}
\text{MSE}(\text{SATE},\ \text{SATT}) &= \text{MSE}(t_{\text{diff}} - \text{SATE},\ t_{\text{diff}} - \text{SATT}) \\
&= \left[\mathbb{E}\left[\text{SATE} - \text{SATT}\right]\right]^2 + \text{Var}\left(\text{SATT}\right) \\
&= \text{Var}\left(\text{SATT}\right) \\
&= \left(\frac{1}{m} - \frac{1}{N}\right) \cdot \sigma_\tau^2 \\
&= \frac{N-m}{Nm} \cdot \text{Var}\left(\tau_i\right) = \frac{1-p}{m} \cdot \sigma_\tau^2
\end{aligned}
$$

A.2 Proof of Lemma 3.1

The sum of the responses of the treated units can be decomposed into two components, the sum of the potential outcomes under control and the sum of the treatment effects.

$$
\sum_{i=1}^{N} Y_i \cdot T_i = \sum_{i=1}^{N} T_i \cdot Y_i(1) = \sum_{i=1}^{N} T_i \cdot Y_i(0) + \sum_{i=1}^{N} T_i \cdot (Y_i(1) - Y_i(0)) \tag{A.1}
$$

$$
\Rightarrow \frac{1}{m} \cdot \sum_{i=1}^{N} Y_i \cdot T_i = \frac{1}{m} \cdot \sum_{i=1}^{N} T_i \cdot Y_i(0) + \text{SATT} \tag{A.2}
$$

We can re-write the difference in means as,

$$
\begin{aligned}
t_{\text{diff}} &= \frac{1}{m} \cdot \sum_{i=1}^{N} Y_i \cdot T_i - \frac{1}{N-m} \cdot \sum_{i=1}^{N} Y_i \cdot (1 - T_i) \\
&= \frac{1}{m} \cdot \sum_{i=1}^{N} T_i \cdot Y_i(0) + \text{SATT} - \frac{1}{N-m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot (1 - T_i) \\
&= \frac{1}{m} \cdot \sum_{i=1}^{N} T_i \cdot Y_i(0) - \frac{1}{N-m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot (1 - T_i) + \text{SATT} \\
&= \frac{N}{m \cdot (N-m)} \cdot \sum_{i=1}^{N} Y_i(0) T_i - \frac{1}{N-m} \cdot \underbrace{\sum_{i=1}^{N} Y_i(0)}_{\text{constant}} + \text{SATT}
\end{aligned} \tag{A.3}
$$

The decomposition of t_{diff} as a function of $Y(1)$ and the SATC can be proved using a similar proof to the one above.

A.3 Proof of Lemma 3.2

The variance of $t_{\text{diff}} - \text{SATE}$ can be written as,

$$\begin{aligned}
\text{Var}\left(t_{\text{diff}} - \text{SATE}\right) &= \frac{m}{N \cdot (N-m)}\sigma_0^2 + \frac{N-m}{N \cdot m}\sigma_1^2 + \frac{2}{N} \cdot \rho \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{p}{N \cdot (1-p)} \cdot \sigma_0^2 + \frac{N \cdot (1-p)}{N^2 \cdot p} \cdot \sigma_1^2 + \frac{2}{N} \cdot \rho \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{p}{N \cdot (1-p)} \cdot \sigma_0^2 + \frac{(1-p)}{N \cdot p} \cdot \sigma_1^2 + \frac{2}{N} \cdot \rho \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left[p^2 \cdot \sigma_0^2 + (1-p)^2 \cdot \sigma_1^2 + 2p(1-p) \cdot \rho \cdot \sigma_0 \cdot \sigma_1\right]
\end{aligned}$$

The variance derivations of $t_{\text{diff}} - \text{SATT}$ are described in A.5, and the variance calculation of $t_{\text{diff}} - \text{SATC}$ follows exactly the same steps as the variance calculation of $t_{\text{diff}} - \text{SATT}$.

A.4 Proof of Theorem 3.1

Proof of part (1) of Theorem 3.1

The proof is a simple implementation of the intermediate value theorem. We show that the difference in variance is negative when $\rho = -1$, and is positive when $\rho = 1$, and that the difference in variance is a continuous and increasing function with respect to ρ. Then the desired result follows immediately from the intermediate value theorem.

When $\rho = -1$, the variance of $t_{\text{diff}} - \text{SATE}$ can be written as:

$$\begin{aligned}
\text{Var}\left(t_{\text{diff}} - \text{SATE}\right) &= \frac{m}{N \cdot (N-m)}\sigma_0^2 + \frac{N-m}{N \cdot m}\sigma_1^2 + \frac{2}{N} \cdot \rho \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{p}{N \cdot (1-p)} \cdot \sigma_0^2 + \frac{N \cdot (1-p)}{N^2 \cdot p} \cdot \sigma_1^2 - \frac{2}{N} \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{p}{N \cdot (1-p)} \cdot \sigma_0^2 + \frac{(1-p)}{N \cdot p} \cdot \sigma_1^2 - \frac{2}{N} \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left[p^2 \cdot \sigma_0^2 + (1-p)^2 \cdot \sigma_1^2 - 2p(1-p) \cdot \sigma_0 \cdot \sigma_1\right] \\
&= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left(p \cdot \sigma_0 - (1-p) \cdot \sigma_1\right)^2
\end{aligned}$$

and as,

$$p \cdot \sigma_0 < \sigma_0 - (1-p) \cdot \sigma_1 < \sigma_0$$

it follows that when $\rho = -1$, then $\text{Var}\left(t_{\text{diff}} - \text{SATE}\right) < \text{Var}\left(t_{\text{diff}} - \text{SATT}\right)$. When $\rho = 1$, the variance of $t_{\text{diff}} - \text{SATE}$ can be written as:

$$\begin{aligned}
\operatorname{Var}(t_{\text{diff}} - \text{SATE}) &= \frac{p}{N \cdot (1-p)} \cdot \sigma_0^2 + \frac{N \cdot (1-p)}{N^2 \cdot p} \cdot \sigma_1^2 + \frac{2}{N} \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{p}{N \cdot (1-p)} \cdot \sigma_0^2 + \frac{(1-p)}{N \cdot p} \cdot \sigma_1^2 + \frac{2}{N} \cdot \sigma_0 \cdot \sigma_1 \\
&= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left[p^2 \cdot \sigma_0^2 + (1-p)^2 \cdot \sigma_1^2 + 2p(1-p) \cdot \sigma_0 \cdot \sigma_1\right] \\
&= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left(p \cdot \sigma_0 + (1-p) \cdot \sigma_1\right)^2
\end{aligned}$$

and as $\sigma_1 > \sigma_0$, it follows immediately that when $\rho = 1$, $\operatorname{Var}(t_{\text{diff}} - \text{SATE}) > \operatorname{Var}(t_{\text{diff}} - \text{SATT})$.

As $\frac{\partial \operatorname{Var}(t_{\text{diff}} - \text{SATE}) - \operatorname{Var}(t_{\text{diff}} - \text{SATT})}{\partial \rho} > 0$, it follows directly from the intermediate value theorem that there exists a value of ρ, $\bar{\rho}$, such that:

$$\begin{aligned}
&\rho \leq \bar{\rho} \Rightarrow \operatorname{Var}(t_{\text{diff}} - \text{SATE}) \leq \operatorname{Var}(t_{\text{diff}} - \text{SATT}) \\
&\rho > \bar{\rho} \Rightarrow \operatorname{Var}(t_{\text{diff}} - \text{SATE}) > \operatorname{Var}(t_{\text{diff}} - \text{SATT})
\end{aligned}$$

Proof of part (2) of Theorem 3.1

Consider the case in which $\rho = 0$:

$$\operatorname{Var}(t_{\text{diff}} - \text{SATE}) = \frac{1}{N \cdot (1-p) \cdot p} \cdot \left[p^2 \cdot \sigma_0^2 + (1-p)^2 \cdot \sigma_1^2\right]$$

$\operatorname{Var}(t_{\text{diff}} - \text{SATT})$ will be lower than $\operatorname{Var}(t_{\text{diff}} - \text{SATE})$, when:

$$p^2 \cdot \sigma_0^2 + (1-p)^2 \cdot \sigma_1^2 \geq \sigma_0^2 \Rightarrow \frac{(1-p)^2}{1-p^2} \cdot \sigma_1^2 \geq \sigma_0^2$$

Note, $\frac{(1-p)^2}{1-p^2} = \frac{1+p^2-2p}{1-p^2} \leq 1 \Rightarrow p^2 - 2p \leq -p^2 \Rightarrow 2p^2 \leq 2p \Rightarrow p^2 \leq p$, which is always satisfied. Therefore, $0 \leq \frac{(1-p)^2}{1-p^2} \leq 1$. To conclude, when $\frac{\sigma_1}{\sigma_0} > \sqrt{\frac{1-p^2}{(1-p)^2}}$ the variance difference will be positiv $\operatorname{Var}(t_{\text{diff}} - \text{SATE}) > \operatorname{Var}(t_{\text{diff}} - \text{SATT})$. As the variance difference is negative when $\rho = -1$ it follows directly from the intermediate value theorem that the desired $\bar{\rho}$ exists and is strictly larger than -1 and lower than 0 — $\bar{\rho}$ is negative. This concludes the proof.

A.5 Proof of Theorem 3.2

According to Lemma 3.1, the adjusted difference in means, $\frac{N-m}{N} \cdot (t_{\text{diff}} - \text{SATT})$, has the same distribution of $\frac{1}{m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i$ with an additive shift of $\frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0)$. Hájek (1960), Hájek (1961), Lehmann (1975), and Li and Ding (2016) all provide proofs showing that in a finite population model with complete randomization the standardized mean of the treated (or sampled) units will follow a standard Normal distribution under two regularity conditions,

$$N - m \to \infty, \quad \text{and} \quad m \to \infty \tag{A.4}$$

and

$$\frac{\max\limits_{1\le i\le N}\left(Y(0)_{Ni} - \bar{Y}(0)_N\right)^2}{\sum_{i=1}^{N}\left(Y(0)_{Ni} - \bar{Y}(0)_N\right)^2} \cdot \max\left(\frac{N-m}{m}, \frac{m}{N-m}\right) \to \infty \tag{A.5}$$

The expectation and variance of $\frac{N-m}{N} \cdot (t_{\text{diff}} - \text{SATT})$ are derived next and complete the proof of Theorem 3.2. The expectation is,

$$\begin{aligned}
\mathbb{E}\left[\frac{N-m}{N} \cdot (t_{\text{diff}} - \text{SATT})\right] &= \frac{N-m}{N} \cdot \left[\frac{N}{m(N-m)} \cdot \sum_{i=1}^{N} \mathbb{E}\left[Y(0)\right] \cdot \mathbb{E}\left[T\right] - \frac{1}{N} \cdot \sum_{i=1}^{N} \mathbb{E}\left[Y(0)\right]\right] \\
&= \frac{N}{m} \cdot \mathbb{E}\left[Y(0)\right] \cdot \frac{m}{N} - -\frac{1}{N} \cdot N \cdot \mathbb{E}\left[Y(0)\right] \\
&= 0
\end{aligned}$$

where the first equality follows from Lemma 3.1. The variance of the adjusted difference in means is,

$$\begin{aligned}
\text{Var}\left(\frac{N-m}{N} \cdot (t_{\text{diff}} - \text{SATT})\right) &= \text{Var}\left(\frac{1}{m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i - \frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0)\right) \\
&= \frac{1}{m^2} \cdot \text{Var}\left(\cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i\right)
\end{aligned}$$

where the first equality follows from Lemma 3.1, and the second equality follows as $\frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0)$ is a constant and not a random variable. The variance of the sum $\sum_{i=1}^{N} Y_i(0) \cdot T_i$, is the variance of m units who are randomly sampled from the population $Y_{N1}(0), \ldots, Y_{NN}(0)$.[9] Denote the m chosen variables by $V_1, \ldots, V_m$, and the variance of the sum $V_1 + \cdots + V_m$ is,

$$\text{Var}\left(V_1 + \cdots + V_m\right) = \sum_{i=1}^{m} \text{Var}\left(V_i\right) + \sum_{i=1}^{m} \sum_{j \neq i} \text{Cov}\left(V_i, V_j\right) = m \cdot \sigma_0^2 + m(m-1) \cdot \rho \tag{A.6}$$

where the second equality follows as the covariance between each two units is the same and is

[9] Our calculation of the variance of $\frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0)$ follows the variance calculation of the Wilcoxon Rank Sum Test statistic in Lehmann (1975).

denoted by ρ. When $m = N$, the variance of the sum is zero and therefore,

$$\rho = -\frac{\sigma_0^2}{N-1} \tag{A.7}$$

Substituting (A.7) in (A.6) and re-arranging yields that:

$$\begin{aligned}
\mathrm{Var}\left(\frac{N-m}{N}\cdot(t_{\mathrm{diff}}-\mathrm{SATT})\right) &= \mathrm{Var}\left(\frac{1}{m}\cdot\sum_{i=1}^{N}Y_i(0)\cdot T_i - \frac{1}{N}\cdot\sum_{i=1}^{N}Y_i(0)\right)\\
&= \frac{1}{m^2}\cdot\mathrm{Var}\left(\cdot\sum_{i=1}^{N}Y_i(0)\cdot T_i\right) = \frac{1}{m^2}\cdot\frac{m(N-m)}{N-1}\cdot\sigma_0^2\\
&= \frac{N-m}{m(N-1)}\cdot\sigma_0^2\\
&\approx \frac{N-m}{mN}\cdot\sigma_0^2 = \left(\frac{1}{m}-\frac{1}{N}\right)\cdot\sigma_0^2
\end{aligned}$$

Hence, it follows that:

$$\begin{aligned}
\mathrm{Var}\left((t_{\mathrm{diff}}-\mathrm{SATT})\right) &= \left(\frac{N}{N-m}\right)^2\cdot\mathrm{Var}\left(\frac{N-m}{N}\cdot(t_{\mathrm{diff}}-\mathrm{SATT})\right)\\
&= \left(\frac{N}{N-m}\right)^2\cdot\left(\frac{N-m}{mN}\right)\cdot\sigma_0^2\\
&= \frac{N}{(N-m)\cdot m}\cdot\sigma_0^2\\
&= \frac{1}{Np(1-p)}\cdot\sigma_0^2
\end{aligned}$$

A.6 Proof of Theorem 3.3

The first part of the theorem follows immediately from the proof of Theorem 1 in Rigdon and Hudgens (2015). Next we show the derivations of the second part of the theorem.

$$\begin{aligned}
p\cdot L_{\mathrm{SATT}} + (1-p)\cdot L_{\mathrm{SATC}} &= t_{\mathrm{diff}} - z_{1-\alpha/2}\cdot\sqrt{k(N,m)}\cdot(p\sigma_0+(1-p)\sigma_1)\\
&= t_{\mathrm{diff}} - z_{1-\alpha/2}\cdot\sqrt{k(N,m)}\cdot\sqrt{\hat{\mathbb{V}}_{\rho=1}}
\end{aligned}$$

and

$$\begin{aligned}
p\cdot U_{\mathrm{SATT}} + (1-p)\cdot U_{\mathrm{SATC}} &= t_{\mathrm{diff}} + z_{1-\alpha/2}\cdot\sqrt{k(N,m)}\cdot(p\sigma_0+(1-p)\sigma_1)\\
&= t_{\mathrm{diff}} + z_{1-\alpha/2}\cdot\sqrt{k(N,m)}\cdot\sqrt{\hat{\mathbb{V}}_{\rho=1}}
\end{aligned}$$

The second equality in both of the above equations follows by substituting $\rho = 1$ in the variance formula $\text{Var}\,(t_{\text{diff}} - \text{SATE})$ in Lemma 3.2.

A.7 Proof of Lemma 3.3

The desired result follows immediately from a comparison of the different variance expressions in Lemma 3.2.

B Super-population model

In the super-population model, the sampling procedure has two steps. First, a sample of N units is drawn from a super-population, $F_{Y(1),Y(0)}(\cdot)$, which can be either finite or infinite. Second, m units are randomly allocated to the treatment regime and the remaining $N - m$ units are assigned to the control regime.[10] Both the treatment indicator and the potential outcomes are random variables, and they are independent, due to the random assignment of units to treatment regimes. SATE, SATT, and SATC are all random variables that are unbiased, and consistent, estimators of PATE. Under the super-population model the difference-in-means test statistic can be used to construct a CI for PATE. Neyman's variance estimator is a consistent estimator for the variance of the difference-in-means test statistic (Imbens and Rubin, 2015). PATE, PATT, and PATC are all equal and recentering the difference-in-means w.r.t. PATE is equivalent to recentering it w.r.t. PATT (or PATC). There are no efficiency gains or inaccuracies in making inference on PATT instead of PATE.

Next we describe the behavior of the difference-in-means when it is recentered w.r.t. different sample average treatment effect estimands (e.g., SATE, SATT). The variance of $t_{\text{diff}} - \text{SATE}$ is lower than that of $t_{\text{diff}} - \text{PATE}$:

$$\text{Var}\,(t_{\text{diff}} - \text{SATE}) = \text{Var}\,(t_{\text{diff}}) - \underbrace{\frac{1}{N} \cdot \sigma_\tau^2}_{\text{Var(SATE)}} \leq \text{Var}\,(t_{\text{diff}}) = \text{Var}\,(t_{\text{diff}} - \text{PATE}) \tag{B.1}$$

and

$$\text{Var}\,(t_{\text{diff}}) = \frac{\sigma_1^2(\text{SP})}{m} + \frac{\sigma_0^2(\text{SP})}{N - m} \tag{B.2}$$

where $\sigma_j^2(\text{SP})$ $(j = 0, 1)$ denotes the variance of units under treatment regime j in the super-population. See Appendix B.1 for a proof of Equation (B.1). In the fixed population case, SATE is a parameter and not a random variable, which changes when sampling uncertainty is introduced and the potential outcomes become random variables as well.

[10] The number of treated and control units is fixed: m is not a random variable.

Lemma B.1 establishes that the variances calculations we derived in Section 3 hold also under the super-population framework. Although the variance calculations are over repeated samples, the correlation term between potential outcomes, ρ, does not go away and to conduct inference on the SATE all the issues we discussed in Section 3 arise here as well. More specifically, the uncertainty calculation over repeated draws of data cancels the ρ term in the variance of the difference-in-means test statistic. However, it also introduces a correlation term ρ in the variance of SATE, which is now a random variable. The two elements cancel each other out and yield a variance formula that is exactly the same as the one that was derived in Lemma 3.2.

Lemma B.1. Under the super-population model, the variance of the difference-in-means when recentered w.r.t. the SATE, SATT, or SATC is

$$\begin{aligned}
\text{Var}\left(t_{\text{diff}} - \text{SATE}\right) &= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left[p^2 \cdot \sigma_0^2(\text{SP}) + (1-p)^2 \cdot \sigma_1^2(\text{SP}) + 2p(1-p) \cdot \rho(\text{SP}) \cdot \sigma_0(\text{SP}) \cdot \sigma_1(\text{SP})\right] \\
\text{Var}\left(t_{\text{diff}} - \text{SATT}\right) &= \frac{1}{p \cdot (1-p) \cdot N} \cdot \sigma_0^2(\text{SP}) \\
\text{Var}\left(t_{\text{diff}} - \text{SATC}\right) &= \frac{1}{p \cdot (1-p) \cdot N} \cdot \sigma_1^2(\text{SP})
\end{aligned}$$

See Appendix B.1 for the proof.

B.1 Proof of Lemma B.1

$$\text{Var}\left(t_{\text{diff}} - \text{PATE}\right) = \text{Var}\left(t_{\text{diff}}\right) = \frac{\sigma_0^2(\text{SP})}{N-m} + \frac{\sigma_1^2(\text{SP})}{m}$$

see Imbens and Rubin (2015) for proof.

Next we derive the variance of $t_{\text{diff}} - \text{SATE}$:

$$\begin{aligned}
\text{Var}\left(t_{\text{diff}} - \text{SATE}\right) &= \mathbb{E}\left[\left(t_{\text{diff}} - \text{SATE}\right)^2\right] - \left(t_{\text{diff}} - \text{SATE}\right)^2 \\
&= \ \mathbb{E}\left[\left(t_{\text{diff}} - \text{SATE}\right)^2\right] \\
&= \mathbb{E}\left[t_{\text{diff}}^2\right] + \mathbb{E}\left[\text{SATE}^2\right] - 2 \cdot \mathbb{E}\left[t_{\text{diff}} \cdot \text{SATE}\right] \\
&= \mathbb{E}\left[t_{\text{diff}}^2\right] + \mathbb{E}\left[\text{SATE}^2\right] - 2 \cdot \left(\text{Var}\left(\text{SATE}\right) + \text{PATE}^2\right) \\
&= \text{Var}\left(t_{\text{diff}}\right) + \text{Var}\left(\text{SATE}\right) - 2 \cdot \text{Var}\left(\text{SATE}\right) \\
&= \text{Var}\left(t_{\text{diff}}\right) - \text{Var}\left(\text{SATE}\right)
\end{aligned}$$

Note that,

$$\begin{aligned}
\mathbb{E}\left[t_{\text{diff}} \cdot \text{SATE}\right] &= \mathbb{E}_{\text{SP}}\left[\mathbb{E}_{\text{T}}\left[t_{\text{diff}} \cdot \text{SATE} \middle| \left(\mathbf{Y}(\mathbf{1}), \mathbf{Y}(\mathbf{0})\right)\right]\right] \\
&= \mathbb{E}_{\text{SP}}\left[\text{SATE}^2\right] \\
&= \text{Var}\left(\text{SATE}\right) + \left[\mathbb{E}_{\text{SP}}\left[\text{SATE}\right]\right]^2 \\
&= \text{Var}\left(\text{SATE}\right) + \text{PATE}^2
\end{aligned}$$

The variance of SATE is:

$$\text{Var}\left(\text{SATE}\right) = \frac{1}{N} \cdot \text{Var}\left(Y_i(1) - Y_i(0)\right) = \frac{1}{N} \cdot \left(\sigma_1^2(\text{SP}) + \sigma_0^2(\text{SP}) - 2\rho(\text{SP})\sigma_1(\text{SP})\sigma_0(\text{SP})\right)$$

Therefore the $\text{Var}\left(t_{\text{diff}} - \text{SATE}\right)$ is:

$$\begin{aligned}
\text{Var}\left(t_{\text{diff}} - \text{SATE}\right) &= \frac{1}{m} \cdot \sigma_1^2(\text{SP}) + \frac{1}{N-m} \cdot \sigma_0^2(\text{SP}) - \frac{1}{N} \cdot \left(\sigma_1^2(\text{SP}) + \sigma_0^2(\text{SP}) - 2\rho(\text{SP})\sigma_1(\text{SP})\sigma_0(\text{SP})\right) \\
&= \frac{N-m}{Nm} \cdot \sigma_1^2(\text{SP}) + \frac{m}{(N-m)N} \cdot \sigma_0^2(\text{SP}) + \frac{2\rho(\text{SP})\sigma_1(\text{SP})\sigma_0(\text{SP})}{N} \\
&= \frac{1-p}{Np} \cdot \sigma_1^2(\text{SP}) + \frac{p}{N(1-p)} \cdot \sigma_0^2(\text{SP}) + \frac{2\rho(\text{SP})\sigma_1(\text{SP})\sigma_0(\text{SP})}{N} \\
&= \frac{1}{Np(1-p)} \cdot \left[(1-p)^2 \cdot \sigma_1^2(\text{SP}) + p^2 \cdot \sigma_0^2(\text{SP})\right] + \frac{2\rho(\text{SP})\sigma_1(\text{SP})\sigma_0(\text{SP})}{N} \\
&= \frac{1}{N \cdot (1-p) \cdot p} \cdot \left[p^2 \cdot \sigma_0^2(\text{SP}) + (1-p)^2 \cdot \sigma_1^2(\text{SP}) + 2p(1-p) \cdot \rho(\text{SP}) \cdot \sigma_0(\text{SP}) \cdot \sigma_1(\text{SP})\right]
\end{aligned}$$

Next we derive the variance of $t_{\text{diff}} - \text{SATT}$ under the super-population model.

$$\begin{aligned}
\text{Var}\left(\frac{N-m}{N} \cdot \left(t_{\text{diff}} - \text{SATT}\right)\right) &= \mathbb{E}_{\text{ Y(1),Y(0)}}\left[\text{Var}_{\text{T}}\left(\frac{N-m}{N} \cdot \left(t_{\text{diff}} - \text{SATT}\right) | Y(1), Y(0)\right)\right] \\
&\quad + \text{Var}_{\text{Y(1),Y(0)}}\left(\mathbb{E}_{\text{ T}}\left[\frac{N-m}{N} \cdot \left(t_{\text{diff}} - \text{SATT}\right) | Y(1), Y(0)\right]\right) \\
&= \mathbb{E}_{\text{ Y(1),Y(0)}}\left[\text{Var}_{\text{T}}\left(\frac{N-m}{N} \cdot \left(t_{\text{diff}} - \text{SATT}\right) | Y(1), Y(0)\right)\right] \\
&= \mathbb{E}_{\text{ Y(1),Y(0)}}\left[\frac{N-m}{m(N-1)} \cdot \sigma_0^2(\text{FS})\right] \\
&= \frac{N-m}{m(N-1)} \cdot \mathbb{E}_{\text{ Y(1),Y(0)}}\left[\sigma_0^2(\text{FS})\right] \\
&= \frac{N-m}{m(N-1)} \cdot \sigma_0^2(\text{SP})
\end{aligned}$$

where the first equality follows from the law of total variance (i.e., law of conditional variance). There are a few more technical steps to complete the derivation. We neglect them as they already appeared in detail in Appendix A.3. The calculation of the variance of $t_{\text{diff}} - \text{SATC}$ is similar to the one above.

B.2 Proof: Variance of $(t_{\text{diff}} - \textbf{SATO})$

Consider the following class of average treatment effect estimands:

$$\text{SATO} = \omega \cdot \text{SATT} + (1-\omega) \cdot \text{SATC} \tag{B.3}$$

We discuss inferences on this class of estimands using the difference-in-means test statistic.

$$\text{Var}\left(t_{\text{diff}} - \text{SATO}\right) = \text{Var}\left(t_{\text{diff}}\right) + \text{Var}\left(\text{SATO}\right) - 2 \cdot \text{Cov}\left(\text{SATO},\ t_{\text{diff}}\right)$$

According to the derivations that are detailed later on in this section we can re-write the variance of t_{diff} re-centered w.r.t SATO as:

$$\begin{aligned}
\text{Var}\left(t_{\text{diff}} - \text{SATO}\right) = &\frac{1}{Np(1-p)}\left[p^2\sigma_0^2 + (1-p)^2\sigma_1^2 + 2p(1-p)\cdot\rho\sigma_0\sigma_1\right] \\
&+ \frac{\sigma_\tau^2}{N-1}\cdot\left[\omega^2\cdot\frac{1-p}{p} + (1-\omega)^2\cdot\frac{p}{1-p} - 2\omega(1-\omega)\right] \\
&- 2\cdot\left[\frac{\omega}{(N-1)p}\left[\sigma_1^2 - \rho\sigma_1\sigma_0\right] - \frac{1}{N-1}\cdot\sigma_\tau^2 + \frac{(1-\omega)}{(N-1)(1-p)}\left[\sigma_0^2 - \rho\sigma_1\sigma_0\right]\right]
\end{aligned}$$

Calculation of $\textbf{Var}\left(\textbf{SATO}\right)$

$$\begin{aligned}
\text{Var}\left(\text{SATO}\right) &= \omega^2\cdot\text{Var}\left(\text{SATT}\right) + (1-\omega)^2\cdot\text{Var}\left(\text{SATC}\right) - 2\cdot\omega(1-\omega)\cdot\text{Cov}\left(\text{SATT},\ \text{SATC}\right) \\
&= \omega^2\cdot\frac{N-m}{m(N-1)}\cdot\sigma_\tau^2 + (1-\omega)^2\cdot\frac{m}{(N-m)(N-1)}\cdot\sigma_\tau^2 - \omega(1-\omega)\cdot\frac{2}{N-1}\cdot\sigma_\tau^2 \\
&= \frac{\sigma_\tau^2}{N-1}\cdot\left[\omega^2\cdot\frac{1-p}{p} + (1-\omega)^2\cdot\frac{p}{1-p} - 2\omega(1-\omega)\cdot\right]
\end{aligned}$$

Calculation of $\textbf{Cov}\left(\textbf{SATO},\ t_{\textbf{diff}}\right)$

The covariance term is:

$$\begin{aligned}
&\text{Cov}\left(\omega\cdot\text{SATT} + (1-\omega)\cdot\text{SATC},\ \frac{N}{m\cdot(N-m)}\cdot\sum_{i=1}^{N}Y_i(0)\cdot T_i - \frac{1}{N-m}\cdot\sum_{i=1}^{N}Y_i(0) + \text{SATT}\right) \\
&= \omega\cdot\frac{N}{m\cdot(N-m)}\cdot\text{Cov}\left(\text{SATT},\ \sum_{i=1}^{N}Y_i(0)\cdot T_i\right) + \omega\cdot\text{Var}\left(\text{SATT}\right) \\
&\quad + (1-\omega)\cdot\frac{N}{m\cdot(N-m)}\cdot\text{Cov}\left(\text{SATC},\ \sum_{i=1}^{N}Y_i(0)\cdot T_i\right) + (1-\omega)\cdot\text{Cov}\left(\text{SATC},\text{SATT}\right)
\end{aligned}$$

According to the derivations that are described in detail later on in this section we can re-write the above covariance expression as:

$$
\begin{aligned}
\mathrm{Cov}\left(\mathrm{SATO},\ t_{\mathrm{diff}}\right) &= \omega \cdot \frac{N}{m \cdot (N-m)} \cdot \frac{(1-p)N}{N-1} \cdot \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] + \omega \cdot \frac{N-m}{m(N-1)} \cdot \sigma_\tau^2 \\
&\quad - (1-\omega) \cdot \frac{N}{m(N-m)} \cdot \frac{m}{N-1} \cdot \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] - (1-\omega) \cdot \frac{1}{N-1} \cdot \sigma_\tau^2 \\
&= \omega \cdot \frac{1}{p(N-1)} \cdot \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] + \omega \cdot \frac{(1-p)}{p(N-1)} \cdot \sigma_\tau^2 - \frac{1}{N-1}(1-\omega)\sigma_\tau^2 \\
&\quad - \frac{(1-\omega)}{(N-1)(1-p)} \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] \\
&= \frac{\omega}{(N-1)p} \left[\rho\sigma_1\sigma_0 - \sigma_0^2 + (1-p)\sigma_\tau^2 + p\sigma_\tau^2\right] - \frac{1}{N-1} \cdot \sigma_\tau^2 \\
&\quad - \frac{(1-\omega)}{(N-1)(1-p)} \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] \\
&= \frac{\omega}{(N-1)p} \left[\sigma_1^2 - \rho\sigma_1\sigma_0\right] - \frac{1}{N-1} \cdot \sigma_\tau^2 - \frac{(1-\omega)}{(N-1)(1-p)} \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right]
\end{aligned}
$$

Calculation of $\mathbf{Var}\left(\mathbf{SATT}\right)$ and $\mathbf{Var}\left(\mathbf{SATC}\right)$

$$
\mathrm{Var}\left(\mathrm{SATT}\right) = \frac{N-m}{m(N-1)} \cdot \sigma_\tau^2 = \frac{N-m}{m(N-1)} \cdot \left[\sigma_1^2 + \sigma_0^2 - 2\rho \cdot \sigma_1 \cdot \sigma_0\right]
$$

Similarly,

$$
\mathrm{Var}\left(\mathrm{SATC}\right) = \frac{m}{(N-m)(N-1)} \cdot \sigma_\tau^2
$$

Calculation of $\mathbf{Cov}\left(\mathbf{SATC},\ \sum_{i=1}^N Y_i(0) \cdot T_i\right)$

$$
\begin{aligned}
\mathrm{Cov}\left(\mathrm{SATC},\ \sum_{i=1}^{N} Y_i(0) \cdot T_i\right) &= -\frac{m}{N-m} \cdot \mathrm{Cov}\left(\mathrm{SATT}, \sum_{i=1}^{N} Y_i(0) \cdot T_i\right) \\
&= -\frac{m}{N-m} \cdot \frac{(1-p)N}{N-1} \cdot \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] \\
&= -\frac{m}{N-1} \cdot \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right] \\
&\approx -p \cdot \left[\rho\sigma_1\sigma_0 - \sigma_0^2\right]
\end{aligned}
$$

where the second equality follows from the derivations in the next subsection below.

Calculation of Cov $\left(\textbf{SATT},\ \sum_{i=1}^N Y_i(0)\cdot T_i\right)$

$$\begin{aligned}
\text{Cov}\left(\text{SATT},\ \sum_{i=1}^N Y_i(0)\cdot T_i\right) &= \frac{1}{m}\cdot\text{Cov}\left(\sum_{i=1}^N Y_i(1)\cdot T_i - \sum_{i=1}^N Y_i(0)\cdot T_i,\ \sum_{i=1}^N Y_i(0)\cdot T_i\right)\\
&= \frac{1}{m}\cdot\left[\text{Cov}\left(\sum_{i=1}^N Y_i(1)\cdot T_i, \sum_{i=1}^N Y_i(0)\cdot T_i\right) - \text{Var}\left(\sum_{i=1}^N Y_i(0)\cdot T_i\right)\right]\\
&= \frac{1}{m}\cdot\left[\frac{p(1-p)N^2}{N-1}\cdot\rho\cdot\sigma_1\sigma_0 - \frac{m(N-m)}{N-1}\cdot\sigma_0^2\right]\\
&= \frac{(1-p)N}{N-1}\cdot\left[\rho\sigma_1\sigma_0 - \sigma_0^2\right]
\end{aligned}$$

Derivations of Cov $\left(\sum_{i=1}^N Y_i(1)\cdot T_i, \sum_{i=1}^N Y_i(0)\cdot T_i\right)$

$$\begin{aligned}
\mathbb{E}\left[\sum_{l=1}^N Y_l(1)\cdot T_l\cdot\sum_{i=1}^N Y_i(0)\cdot T_i\right] &= \mathbb{E}\left[\sum_{i=1}^N\sum_{l=1}^N Y_l(1)\cdot Y_i(0)\cdot T_i\cdot T_l\right]\\
&= \mathbb{E}\left[T_i\cdot\sum_{i=1}^N Y_i(0)\cdot Y_i(1) + \sum_{i=1}^N\sum_{l\neq i} Y_i(0)\cdot Y_l(1)\cdot T_i\cdot T_l\right]\\
&= \Pr(T_i=1)\cdot\sum_{i=1}^N Y_i(0)\cdot Y_i(1) + \mathbb{E}\left[T_i\cdot T_l|i\neq l\right]\cdot\sum_{i=1}^N\sum_{l\neq i} Y_i(0)\cdot Y_l(1)\\
&= \frac{m}{N}\cdot\sum_{i=1}^N Y_i(0)\cdot Y_i(1) + \frac{m}{N}\cdot\frac{m-1}{N-1}\cdot\sum_{i=1}^N\sum_{l\neq i} Y_i(0)\cdot Y_l(1)\\
&= \frac{m}{N}\cdot\frac{N-m}{N-1}\cdot\sum_{i=1}^N Y_i(0)\cdot Y_i(1) + \frac{m}{N}\cdot\frac{m-1}{N-1}\cdot\sum_{i=1}^N\sum_{l=1}^N Y_i(0)\cdot Y_l(1)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[\sum_{l=1}^N Y_l(1)\cdot T_l\right] &= \frac{m}{N}\sum_{l=1}^N Y_l(1)\\
\mathbb{E}\left[\sum_{l=1}^N Y_l(0)\cdot T_l\right] &= \frac{m}{N}\sum_{l=1}^N Y_l(0)\\
\Rightarrow \mathbb{E}\left[\sum_{l=1}^N Y_l(1)\cdot T_l\right]\cdot\mathbb{E}\left[\sum_{i=1}^N Y_l(0)\cdot T_i\right] &= \left(\frac{m}{N}\right)^2\cdot\sum_{i=1}^N\sum_{l=1}^N Y_i(0)\cdot Y_l(1)
\end{aligned}$$

Hence,

$$
\begin{aligned}
\mathrm{Cov}\left(\sum_{i=1}^{N} Y_i(1)\cdot T_i, \sum_{i=1}^{N} Y_i(0)\cdot T_i\right) &= \frac{m}{N}\cdot\left[\frac{N-m}{N-1}\cdot\sum_{i=1}^{N} Y_i(0)\cdot Y_i(1) - \frac{N-m}{N(N-1)}\cdot\sum_{i=1}^{N}\sum_{l=1}^{N} Y_i(0)\cdot Y_l(1)\right]\\
&= \frac{m}{N}\cdot\frac{N-m}{N-1}\cdot\left[\sum_{i=1}^{N} Y_i(0)\cdot Y_i(1) - \frac{1}{N}\cdot\sum_{i=1}^{N}\sum_{l=1}^{N} Y_i(0)\cdot Y_l(1)\right]\\
&= \frac{p(1-p)N}{N-1}\cdot\left[\sum_{i=1}^{N} Y_i(0)\cdot Y_i(1) - \frac{1}{N}\cdot\sum_{i=1}^{N}\sum_{l=1}^{N} Y_i(0)\cdot Y_l(1)\right]
\end{aligned}
$$

Re-arranging the above expression yields:

$$
\mathrm{Cov}\left(\sum_{i=1}^{N} Y_i(1)\cdot T_i, \sum_{i=1}^{N} Y_i(0)\cdot T_i\right) = \frac{p(1-p)N^2}{N-1}\cdot\rho\cdot\sigma_1\sigma_0
$$

Calculation of Cov (SATT, SATC)

Note that the covariance between SATT and SATC is,

$$
\begin{aligned}
\mathrm{Cov}\left(\mathrm{SATT},\mathrm{SATC}\right) &= \mathrm{Cov}\left(\frac{1}{m}\cdot\sum_{i=1}^{N}(Y_i(1)-Y_i(0))\cdot T_i, \frac{1}{N-m}\cdot\sum_{i=1}^{N}(Y_i(1)-Y_i(0))\cdot(1-T_i)\right)\\
&= -\frac{1}{m}\cdot\frac{1}{N-m}\cdot\mathrm{Cov}\left(\sum_{i=1}^{N}(Y_i(1)-Y_i(0))\cdot T_i, \sum_{i=1}^{N}(Y_i(1)-Y_i(0))T_i\right)\\
&= -\frac{1}{N\cdot p(1-p)}\cdot\frac{\mathrm{Var}\left(\sum_{i=1}^{N}(Y_i(1)-Y_i(0))T_i\right)}{N}\\
&= -\frac{1}{N\cdot p(1-p)}\cdot\frac{m^2}{N}\cdot\mathrm{Var}\left(\mathrm{SATT}\right)\\
&= -\frac{1}{N\cdot p(1-p)}\cdot\frac{m^2}{N}\cdot\frac{N-m}{m(N-1)}\cdot\sigma_\tau^2\\
&= -\frac{1}{N-1}\cdot\sigma_\tau^2
\end{aligned}
$$

Another way of deriving the covariance between the SATT and SATC is to use the fact that

$\text{Var}\,(\text{SATE}) = 0$: First note that,

$$\begin{aligned}
\text{Var}\,(\text{SATE}) &= \text{Var}\left(\frac{m}{N}\cdot \text{SATT}\right) + \text{Var}\left(\frac{N-m}{N}\cdot \text{SATC}\right) + 2\cdot p(1-p)\cdot \text{Cov}\,(\text{SATT},\text{SATC}) \\
&= p^2 \cdot \frac{1}{m^2}\cdot \text{Var}\left(\sum_{i=1}^{N}(Y_i(1)-Y_i(0))T_i\right) + (1-p)^2\cdot \frac{1}{(N-m)^2}\cdot \text{Var}\left(\sum_{i=1}^{N}(Y_i(1)-Y_i(0))T_i\right) \\
&+ 2\cdot p(1-p)\cdot \text{Cov}\,(\text{SATT},\text{SATC}) \\
&= 2\cdot \frac{1}{N^2}\cdot \text{Var}\left(\sum_{i=1}^{N}(Y_i(1)-Y_i(0))T_i\right) + 2\cdot p(1-p)\cdot \text{Cov}\,(\text{SATT},\text{SATC})
\end{aligned}$$

Hence,

$$\begin{aligned}
&\text{Var}\,(\text{SATE}) = 0 \\
\Rightarrow \text{Cov}\,(\text{SATT},\text{SATC}) &= -\frac{1}{N\cdot p(1-p)}\cdot \frac{\text{Var}\left(\sum_{i=1}^{N}(Y_i(1)-Y_i(0))T_i\right)}{N}
\end{aligned}$$

C Additional Monte-Carlo simulations

C.1 Binary outcomes

Consider a finite population of size N: $\Pi_N = \{(Y(0)_{1N}, Y(1)_{1N}), \ldots, (Y(0)_{NN}, Y(1)_{NN})\}$, where $(Y(0)_{iN}, Y(1)_N) \in \{0,1\}^2$, and $p_0 = \frac{1}{N}\cdot\sum_{i=1}^{N} Y_i(0)$ and $p_1 = \frac{1}{N}\cdot\sum_{i=1}^{N} Y_i(1)$. The outcome variable is binary and we assume the treatment has a positive effect on average:

$$0 \leq p_0 \leq 1/2 \quad \text{and} \quad p_0 < p_1 \leq 1/2 \tag{C.1}$$

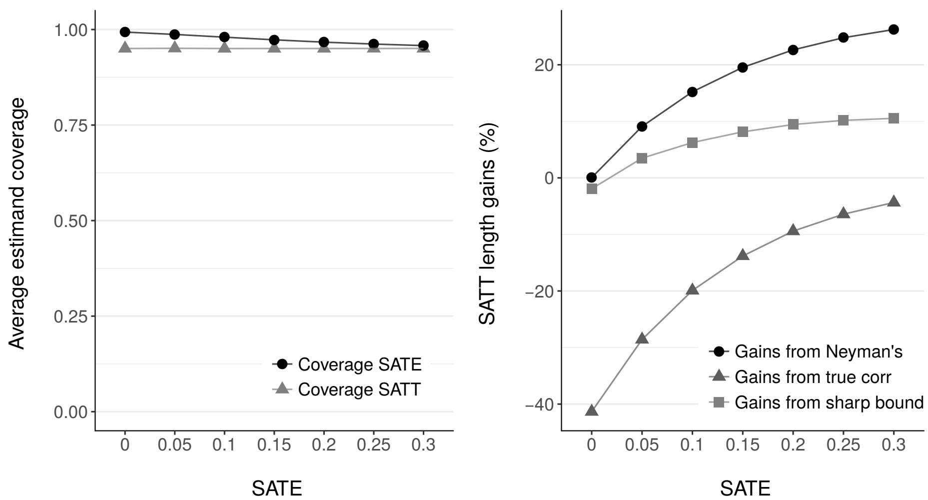
These conditions insure that the positive treatment effect increases the variance of the outcomes among the treated units. SATE is equal to $p_1 - p_0$, which implies that when the conditions in Equation (C.1) are met, the variance of the treated units is strictly higher than that of the controls, $p_0(1-p_0) < p_1(1-p_1)$, and the variance ratio is a function of the treatment effect,

$$\frac{\sigma_1^2}{\sigma_0^2} = \frac{(p_0 + SATE)(1 - p_0 - SATE)}{p_0(1-p_0)} \tag{C.2}$$

Figure C.1 shows the interval length gains (in percentages) from using a PI for SATT in the case of a binary outcome. When $p_0 = 0.1$ and $p_1 = 0.2$, the interval for SATT is lower by 23.3% relative to the standard CI that is based on Neyman's variance estimator.

Next we decompose how much of the efficiency gain is due to a change of estimand, and how much is due to the conservative variance estimation. The lower line (i.e., triangles) in Figure C.1 shows the percentage difference in length relative to a CI for SATE when ρ is known. This CI is substantially shorter than the ones for SATT, which demonstrates that in this simulation the efficiency differences are due to the conservative variance estimation and not because of the change in estimand. Unlike the random coefficient model above, there are small differences in coverage under this data-generating model, although, there are differences in accuracy. The binary outcomes model has less heterogeneity in the treatment effect distribution (i.e., σ_τ^2 is smaller) which mitigates the differences between the different estimands.

Figure C.1: Binary outcome simulation: Length and coverage differences

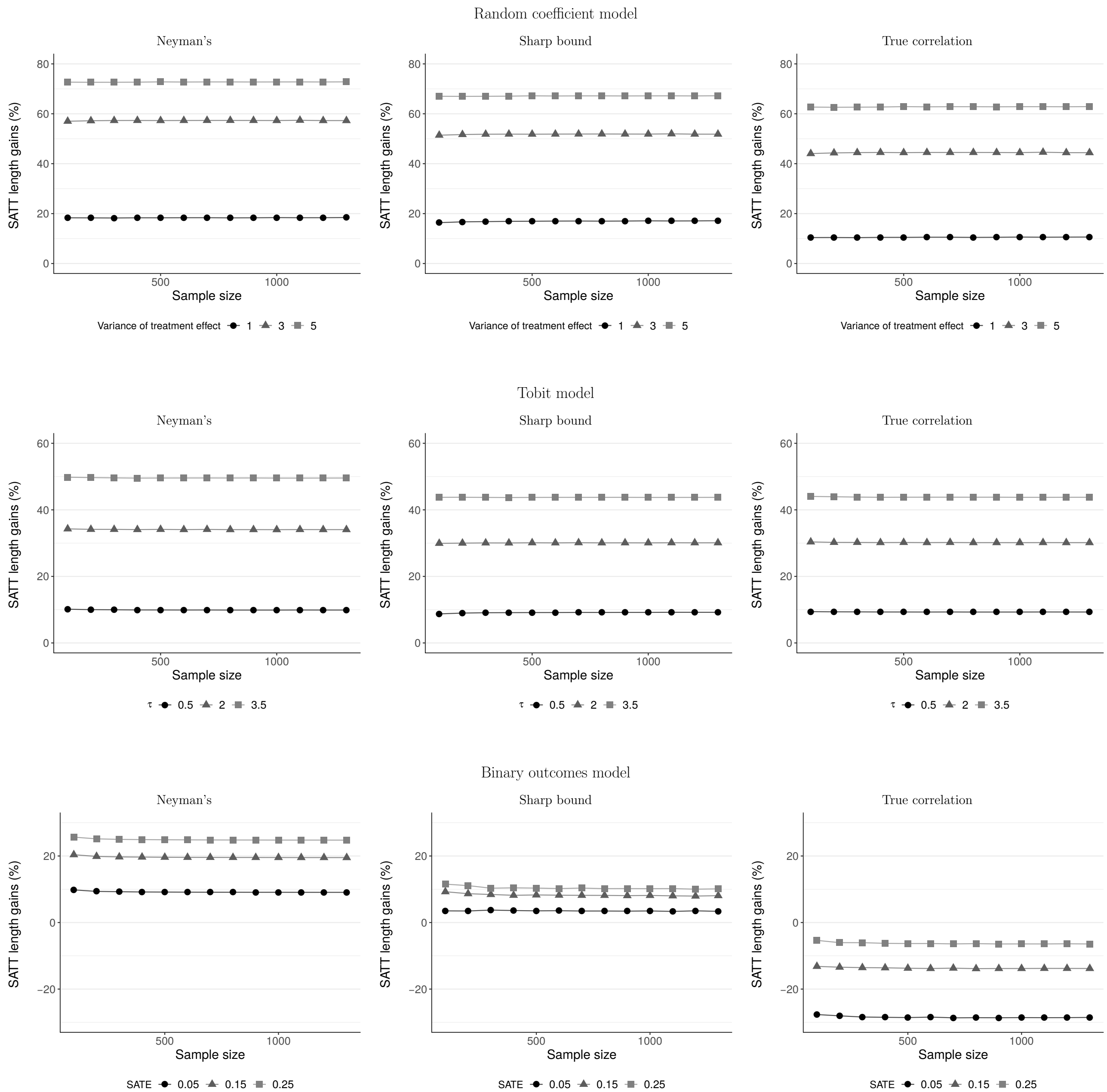


Notes: See the notes in Figure 1.

C.2 Simulation results for multiple sample sizes

This section examines the robustness of the Monte Carlo simulations that have been previously discussed to the chosen sample size. Figure C.2 shows that our choice of using $N = 1,000$ has no impact on the results of the Monte Carlo simulations.

Figure C.2: Monte Carlo simulations by sample size



D Inference on a general class of average treatment effects

In this section we derive inference for the Sample Average Treatment Effect Optimal (SATO), which is the weighting of SATT and SATC that minimizes MSE. The inference is conditional on the sample and uses the difference-in-means test statistics. Formally, SATO is defined as

$$\begin{aligned}
&\text{SATO} \equiv \omega^* \cdot \text{SATT} + (1-\omega^*) \cdot \text{SATC} \\
&s.t. \\
&\omega^* = \underset{\omega}{\operatorname{argmin}} \ \text{Var}\left(\hat{\bar{Y}}_1 - \hat{\bar{Y}}_0 - \text{SATO}\right) = \underset{\omega}{\operatorname{argmin}} \ \text{MSE}\left(\hat{\bar{Y}}_1 - \hat{\bar{Y}}_0 - \text{SATO},\ 0\right)
\end{aligned}$$

Lemma D.1. The variance of the difference-in-means when recentered w.r.t. any choice of weighting between SATT and SATC, $\text{Var}\left(t_{\text{diff}} - (\omega\text{SATT} + (1-\omega)\text{SATC})\right)$, is

$$\begin{aligned}
&\frac{1}{Np(1-p)}\left[p^2\sigma_0^2 + (1-p)^2\sigma_1^2 + 2p(1-p)\cdot\rho\sigma_0\sigma_1\right] \\
&\quad + \frac{\sigma_\tau^2}{N-1}\cdot\left[\omega^2\cdot\frac{1-p}{p} + (1-\omega)^2\cdot\frac{p}{1-p} - 2\omega(1-\omega)\right] \\
&\quad - 2\cdot\left[\frac{\omega}{(N-1)p}\left[\sigma_1^2 - \rho\sigma_1\sigma_0\right] - \frac{1}{N-1}\cdot\sigma_\tau^2 + \frac{(1-\omega)}{(N-1)(1-p)}\left[\sigma_0^2 - \rho\sigma_1\sigma_0\right]\right]
\end{aligned}$$

See Appendix B.2 for the proof.

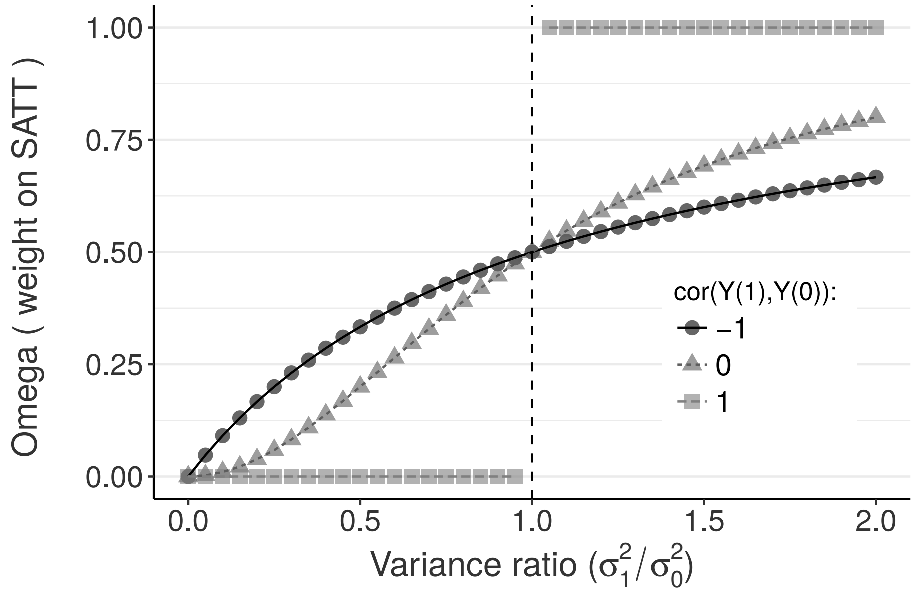
Lemma D.1 presents the variance of the difference-in-means when it is recentered w.r.t. $\omega\text{SATT} + (1-\omega)\text{SATC}$ for any choice of ω. It follows immediately from Lemma D.1, that the optimal accuracy-maximizing value of ω is

$$\omega^* = \frac{\left(\frac{\sigma_1}{\sigma_0}\right)^2 - \rho\cdot\frac{\sigma_1}{\sigma_0}}{\left(\frac{\sigma_1}{\sigma_0}\right)^2 + 1 - 2\rho\left(\frac{\sigma_1}{\sigma_0}\right)}$$

The optimal choice of ω depends on two parameters, ρ and $\frac{\sigma_1}{\sigma_0}$, and is *independent* of p. This is in stark contrast to SATE, which is the weighting of SATT and SATC according *only* to the probability of treatment assignment. Figure D.1 illustrates how the value of ω changes depending on ρ and $\frac{\sigma_1}{\sigma_0}$. Two clear patterns stand out. First, the weight that is assigned to SATT is monotonically increasing w.r.t. the variance ratio, $\frac{\sigma_1}{\sigma_0}$. Second, the relationship between ω^* and ρ is ambiguous and depends on $\frac{\sigma_1}{\sigma_0}$. When $\frac{\sigma_1}{\sigma_0} > 0$, as $\uparrow \rho$, the weight assigned to SATT increases, and the opposite is true when $\frac{\sigma_1}{\sigma_0} < 0$.

A key question is under what conditions SATE is equal to SATO, which is equivalent to asking whether SATE can ever be the estimand that is estimated most accurately. Two scenarios in which SATO coincides with SATE are (i) a constant treatment effect model and (ii) when $\sigma_1 = \sigma_0$ and $p = 1/2$. When the variance ratio is equal to 1, the optimal weighting between SATT and

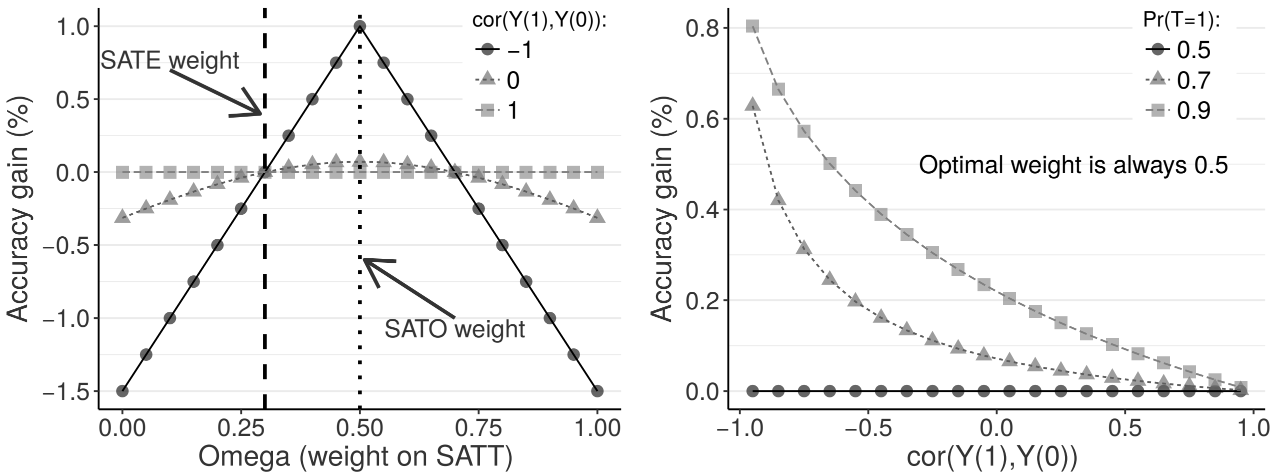
Figure D.1: **Optimal ω weight of SATT in SATO for different $\frac{\sigma_1}{\sigma_0}$ and ρ**



Notes: The above figure describes the relationship between the variance ratio of treated and control units ($\frac{\sigma_1}{\sigma_0}$) and the weight which is assigned to SATT in SATO for three different values of ρ.

SATC is half and half ($\omega = 0.5$). Equality of variance between treatment and control units does *not* imply that SATO collapses to SATE. The right plot of Figure D.2 shows the accuracy gain of SATO relative to SATE for different values of ρ and p. SATO can be estimated at a higher degree of accuracy for most levels of ρ and for any $p \neq 1/2$, even when $\sigma_1 = \sigma_0$. The left plot of Figure D.2 compares the accuracy of inference on SATE to a mixing between SATT and SATC – not SATO as ω^* is not used – for a variety of different ω weights and three values of ρ. SATE can be estimated more accurately than many mixes of SATT and SATC, but it will always – for a known ρ – be more noisily estimated than SATO. To conclude, it is hard to justify a weighting of SATT and SATC that will yield the SATE when the objective is to maximize accuracy. SATE is the only mix of SATT and SATC that yields a parameter and *not* a random variable.

Figure D.2: **The accuracy gain (in percentages) from estimating SATO instead of SATE when $\frac{\sigma_1}{\sigma_0} = 1$ for different values of the treatment assignment probability (p) and the correlation between potential outcomes (ρ)**



Notes: The left plot describes the relationship between the accuracy gain of conducting inference on a mixing between SATT and SATC relative to SATE as a function of the weight which is assigned to SATT (ω) for three different values of ρ. The right plot illustrates the relationship between ρ and the accuracy gains of a PI for SATO relative to a CI for SATE for three different levels of the probability of treatment assignment. In both plots the relationships that are described are for the case in which $\frac{\sigma_1}{\sigma_0} = 0$, which implies that accuracy gains are not the result of a difference in the variance of the outcome among treated units relative to the control units.

E Inference under a Bernoulli trials randomization mechanism)

Theorem E.1. (Limiting distribution of t_{diff} - SATT) When treatment is assigned according to random independent Bernoulli trials, the standardized and recentered, w.r.t. SATT, difference-in-means follows a standard Normal distribution under two regularity conditions. When the following is satisfied:

$$N - m \to \infty, \quad m \to \infty, \quad \text{and} \quad \sigma_1^2,\ \sigma_0^2 < \infty \tag{E.1}$$

then:

$$\frac{\frac{N-m}{N} \cdot (t_{\text{diff}} - \text{SATT})}{\sqrt{\text{Var}\left(\frac{1}{m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i\right)}} \xrightarrow{d} N(0,1) \tag{E.2}$$

where

$$\text{Var}\left(\frac{1}{m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i\right) = \begin{pmatrix} \frac{1}{m} \\ -\frac{\sum_{i=1}^{N} Y_i(0) \cdot T_i}{m^2} \end{pmatrix}^T \cdot \Sigma \cdot \begin{pmatrix} \frac{1}{m} \\ -\frac{\sum_{i=1}^{N} Y_i(0) \cdot T_i}{m^2} \end{pmatrix} \tag{E.3}$$

and

$$\Sigma = \begin{pmatrix} N \cdot p(1-p)\left[\sigma_0^2 + \mathbb{E}\left[Y(0)\right]^2\right] & p(1-p) \cdot N \cdot \mathbb{E}\left[Y(0)\right] \\ p(1-p) \cdot N \cdot \mathbb{E}\left[Y(0)\right] & N \cdot p(1-p) \end{pmatrix} \tag{E.4}$$

Proof:

As both m and $\sum_{i=1}^{N} Y_i(0) \cdot T_i$ are random variables, the distribution of the ratio can be derived using the Delta method. The result that the limiting distribution is Normal follows directly from standard results of the Delta method (Casella and Berger, 2002). The derivations of the variance-covariance matrix of the limiting distribution are detailed below. First notice that:

$$\frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i \xrightarrow{p} \mathbb{E}\left[Y(0)\right]$$
$$\frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i \xrightarrow{d} N\left(\mathbb{E}\left[Y(0)\right],\ \frac{1}{N} \cdot p(1-p)\left[\sigma_0^2 + \mathbb{E}\left[Y(0)\right]^2\right]\right)$$

where

$$\begin{aligned} \mathrm{Var}\left(\frac{1}{N} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i\right) &= \frac{1}{N^2} \cdot p(1-p) \sum_{i=1}^{N} Y_i(0)^2 \\ &= \frac{1}{N} \cdot p(1-p)\left[\sigma_0^2 + \mathbb{E}\left[Y(0)\right]^2\right] \end{aligned}$$

and that m follows a Binomial distribution of $m \sim \mathrm{Binom}(n, p)$ and the variance of m is $\mathrm{Var}(m) = N \cdot p(1-p)$. The covariance between m and $\sum_{i=1}^{N} Y_i(0) \cdot T_i$ is

$$\begin{aligned} \mathrm{Cov}\left(m, \sum_{i=1}^{N} Y_i(0) \cdot T_i\right) = \mathrm{Cov}\left(\sum_{i=1}^{N} T_i, \sum_{i=1}^{N} Y_i(0) \cdot T_i\right) &= \sum_{i=1}^{N} Y_i(0) \cdot \mathrm{Var}\left(T_i\right) \\ &= p(1-p) \cdot \sum_{i=1}^{N} Y_i(0) \\ &= p(1-p) \cdot N \cdot \mathbb{E}\left[Y(0)\right] \end{aligned}$$

and therefore the covariance-variance matrix is,

$$\Sigma = \begin{pmatrix} N \cdot p(1-p)\left[\sigma_0^2 + \mathbb{E}\left[Y(0)\right]^2\right] & p(1-p) \cdot N \cdot \mathbb{E}\left[Y(0)\right] \\ p(1-p) \cdot N \cdot \mathbb{E}\left[Y(0)\right] & N \cdot p(1-p) \end{pmatrix} \tag{E.5}$$

The variance of $\frac{1}{m} \cdot \sum_{i=1}^{N} Y_i(0) \cdot T_i$ is,

$$\begin{pmatrix} \frac{1}{m} \\ -\frac{\sum_{i=1}^{N} Y_i(0) \cdot T_i}{m^2} \end{pmatrix}^T \cdot \Sigma \cdot \begin{pmatrix} \frac{1}{m} \\ -\frac{\sum_{i=1}^{N} Y_i(0) \cdot T_i}{m^2} \end{pmatrix}$$

where a consistent estimator for p is $\frac{1}{N} \cdot \sum_{i=1}^{N} T_i$ and a consistent estimator for $\sum_{i=1}^{N} Y_i(0) \cdot T_i$ is $\sum_{i=1}^{N} Y_i(0) \cdot (1 - T_i)$.