Variance and Conservative Estimation of the Difference-in-Means Estimator (SATE)*

Thomas Leavitt

This guide derives the variance of the Difference-in-Means estimator for the Sample Average Treatment Effect under complete random assignment, and then develops a conservative estimator of that variance. The setup below is self-contained; readers wanting the full development of the framework can consult the companion guide "Notation and Setup: The Finite Population Potential Outcomes Framework." The companion guide "Variance and Conservative Estimation of the Difference-in-Means Estimator (PATE)" extends these results to a superpopulation.

1 Variance of the Difference-in-Means estimator for the Sample Average Treatment Effect (SATE)

1.1 Setup

Let the index $i \in \{1, \ldots, n\}$ run over n units in a finite sample, $\mathcal{S}_n$, where $n \geq 4$. Of these n units, $n_T \geq 2$ are assigned to the treatment condition and $n_C \geq 2$ are assigned to the control condition, where $n_T + n_C = n$. Although not necessary for the unbiasedness of the estimator or for the derivation of the Difference-in-Means estimator's variance, these assumptions on the sizes of n, n_T and n_C ensure that the conservative (Neyman) estimator of the Difference-in-Means estimator's variance is well defined. That estimator forms a separate sample variance within each treatment arm, dividing by $n_T - 1$ for the treated units and by $n_C - 1$ for the control units, and so requires at least two units in each arm. We therefore assume that $n_T \geq 2$ and $n_C \geq 2$, so that $n \geq 4$. Let

*This is a live document that is subject to updating at any time.

the binary indicator variable $Z_i \in \{0,1\}$ denote whether unit i is assigned to treatment ($Z_i = 1$) or control ($Z_i = 0$), and collect these indicators in the random assignment vector $\boldsymbol{Z} = \begin{bmatrix} Z_1, \ldots, Z_n \end{bmatrix}^\top$. The set of all logically possible assignment vectors is $\{0,1\}^n$. The assignment mechanism, however, places positive probability only on those vectors with exactly n_T treated units; we collect these in the set $\Omega = \left\{ \boldsymbol{z} \in \{0,1\}^n : \sum_{i=1}^n z_i = n_T \right\}$, so that Ω is the support of $\boldsymbol{Z}$. Under complete random assignment, $\boldsymbol{Z}$ is distributed uniformly on Ω, and the number of elements in Ω, denoted $|\Omega|$,[1] is $\binom{n}{n_T}$. By contrast, under n independent Bernoulli assignments, there would be 2^n possible assignment vectors. However, even if n_T is not fixed by design (as in complete random assignment), we can fix n_T by conditioning on its observed value. The randomization distribution conditional on the realized n_T yields the same randomization distribution one would obtain if n_T had been fixed ex ante by design. Hence, this general setup and the proof to follow pertains to both simple and complete random assignment even though the argument by which one can regard n_T as fixed is slightly different under simple and complete assignment mechanisms.

Adopting the terminology of Freedman (2009) and later Gerber and Green (2012), define a potential outcomes schedule as a vector-valued function $\boldsymbol{y} : \{0,1\}^n \to \mathbb{R}^n$ that maps each possible assignment vector $\boldsymbol{z} \in \{0,1\}^n$ to an n-dimensional vector of real-valued outcomes. More intuitively, a potential outcomes schedule is a listing of how each of the n study participants would respond to every assignment $\boldsymbol{z}$ that the experiment could in principle produce. The ith entry of $\boldsymbol{y}(\boldsymbol{z})$ is the outcome that unit i would exhibit under assignment $\boldsymbol{z}$. The potential outcomes are *fixed* features of the units; they do not vary with the random assignment. Randomness enters only later, through which assignment $\boldsymbol{z}$ the mechanism happens to select.

Since the assignment space $\{0,1\}^n$ contains 2^n assignment vectors, the schedule specifies, in principle, 2^n outcome vectors. However, under the Stable Unit Treatment Value Assumption (SUTVA)[2] (Cox, 1958; Rubin, 1980, 1986), unit i's outcome depends only on its own assignment z_i and not on the assignments of the other units. Accordingly, let y_{Ti} denote the common outcome value of unit i across all assignments $\boldsymbol{z}$ with $z_i = 1$, and let y_{Ci} denote the common outcome value of unit i across all assignments $\boldsymbol{z}$ with $z_i = 0$. SUTVA thus collapses the entire schedule down to just two fixed numbers per unit, y_{Ti} and y_{Ci}. The individual causal effect for unit i on the additive scale is $\tau_i = y_{Ti} - y_{Ci}$. The vectors $\boldsymbol{y_T}$ and $\boldsymbol{y_C}$ collect the treatment and control potential outcomes, respectively, for all n units, and $\boldsymbol{\tau}$ collects the n individual, additive effects. These are all *fixed*, finite population quantities. The observed outcome for unit $i \in \{1, \ldots, n\}$ is $Y_i = Z_i y_{Ti} + (1 - Z_i)\, y_{Ci}$, which equals y_{Ti} when $Z_i = 1$ and y_{Ci} when $Z_i = 0$. Because y_{Ti} and y_{Ci} are fixed, the *only* source of randomness in the observed outcome Y_i is the assignment indicator Z_i.

[1] For an arbitrary set W, let $|W|$ denote the cardinality of (i.e., the number of elements in) the set W.

[2] SUTVA implies that (1) units in the experiment respond to only the treatment condition to which each unit is individually assigned and (2) the treatment condition is actually the same treatment for all units assigned to treatment and the control condition is the same for all units assigned to control.

It will be convenient to write the finite population means of the treatment and control potential outcomes as $\bar{y}_T \coloneqq n^{-1} \sum_{i=1}^n y_{Ti}$ and $\bar{y}_C \coloneqq n^{-1} \sum_{i=1}^n y_{Ci}$, both of which are fixed quantities. The target of interest is the Sample Average Treatment Effect (SATE),

$$\tau_{\text{SATE}} \coloneqq n^{-1} \sum_{i=1}^{n} \tau_i = \bar{y}_T - \bar{y}_C,$$

which is likewise a fixed (though unknown) quantity. Define the Difference-in-Means estimator of τ_{SATE} under complete random assignment as

$$\hat{\tau} \coloneqq \frac{1}{n_T} \sum_{i=1}^{n} Z_i Y_i - \frac{1}{n_C} \sum_{i=1}^{n} (1 - Z_i) Y_i. \tag{1}$$

This is the same Difference-in-Means estimator whose unbiasedness the companion guide establishes. Under complete random assignment, where $\sum_{i=1}^n Z_i = n_T$ and $\sum_{i=1}^n (1 - Z_i) = n_C$ are fixed, the estimator coincides with the ratio form $\left(\sum_i Z_i Y_i\right) / \left(\sum_i Z_i\right) - \left(\sum_i (1 - Z_i) Y_i\right) / \left(\sum_i (1 - Z_i)\right)$. Unlike the estimand τ_{SATE}, the estimator $\hat{\tau}$ is a *random variable*, since it is a function of the random assignment vector $\boldsymbol{Z}$ and inherits all of its randomness from $\boldsymbol{Z}$. For the expectation and variance of $\hat{\tau}$ in Equation (1), I write $\text{E}_\Omega [\cdot]$ and $\text{Var}_\Omega [\cdot]$ to emphasize that they are taken over only the randomness of the assignment process, i.e., over the distribution of $\boldsymbol{Z}$ on Ω.

1.2 Derivation of variance of Difference-in-Means estimator for the SATE

Proposition 1. *The variance of $\hat{\tau}$ for the τ_{SATE} under complete random assignment is*

$$\text{Var}_\Omega [\hat{\tau}] = \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}, \tag{2}$$

where

$$S_n^2(\boldsymbol{y_T}) = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} \left(y_{Ti} - \frac{1}{n} \sum_{i=1}^{n} y_{Ti} \right)^2 \tag{3}$$

$$S_n^2(\boldsymbol{y_C}) = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} \left(y_{Ci} - \frac{1}{n} \sum_{i=1}^{n} y_{Ci} \right)^2 \tag{4}$$

$$S_n^2(\boldsymbol{\tau}) = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} \left(\tau_i - \frac{1}{n} \sum_{i=1}^{n} \tau_i \right)^2. \tag{5}$$

Proof. Building on Imbens and Rubin (2015, Chapter 6, Appendix A), we will break the proof into several steps:

Step 1: We will show that we can rewrite the Difference-in-Means estimator in (1) in terms of *not* the raw assignment indicators Z_i, but the centered treatment variable $A_i = Z_i - \mathrm{E}_\Omega[Z_i]$. We center because the variance measures the fluctuations of $\hat{\tau}$ around its mean, τ_{SATE}. If we can algebraically separate $\hat{\tau}$ into a fixed part (equal to τ_{SATE}) plus a random part with mean zero, then the variance of $\hat{\tau}$ equals the variance of the random part alone. Centering the assignment variable produces this separation and, because $\mathrm{E}_\Omega[A_i] = 0$, greatly simplifies the algebra that follows.

Step 2: The potential outcomes, $\{y_{Ti}, y_{Ci}\}_{i=1}^n$, are fixed, and the observed outcomes inherit their randomness only from the treatment assignment. Once $\hat{\tau}$ is written as a weighted sum of the A_i, its variance depends only on the first and second moments of the A_i. We will therefore derive $\mathrm{E}_\Omega[A_i]$, $\mathrm{Var}_\Omega[A_i]$ and $\mathrm{Cov}\left[A_i, A_j\right]$ for $i \neq j$. The covariance term is essential and nonzero. Under complete random assignment the total number of treated units is fixed at n_T, so if one unit is pulled into treatment another must be pulled out. This mechanical dependence among the assignment indicators is the source of the covariance term, and we will see that the covariance is negative.

Step 3: We will rely on the expressions for $S_n^2(\boldsymbol{y_T})$ and $S_n^2(\boldsymbol{y_C})$ above, as well as on the following algebraic equivalence that we will need to prove:

$$S_n^2(\boldsymbol{\tau}) = S_n^2(\boldsymbol{y_T}) + S_n^2(\boldsymbol{y_C}) - \frac{2}{n-1}\sum_{i=1}^n \left(y_{Ti} - \frac{1}{n}\sum_{i=1}^n y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^n y_{Ci}\right). \tag{6}$$

Step 4: Once we have completed the three previous steps, we will rely only on the linearity of expectations, rules of variance and (often very messy) algebra.

Step 1

First, note that we can re-write the estimator as

$$\frac{1}{n_T}\sum_{i=1}^n Z_i Y_i - \frac{1}{n_C}\sum_{i=1}^n (1 - Z_i) Y_i = \frac{1}{n}\sum_{i=1}^n \left(\frac{n}{n_T} Z_i y_{Ti} - \frac{n}{n_C}(1 - Z_i) y_{Ci}\right)$$

Because $n = n_C + n_T$, note that $1 = \dfrac{n_C + n_T}{n}$, so instead of $(1 - Z_i)$ we can write $\left(\dfrac{n_C + n_T}{n} - Z_i\right)$:

$$\frac{1}{n}\sum_{i=1}^n \frac{n}{n_T} Z_i y_{Ti} - \frac{n}{n_C}\left(\frac{n_C + n_T}{n} - Z_i\right) y_{Ci},$$

which after a little bit of algebra yields

$$\frac{1}{n}\sum_{i=1}^{n}\frac{n}{n_T}\left(\underbrace{Z_i-\frac{n_T}{n}}_{=A_i}+\frac{n_T}{n}\right)y_{Ti}-\frac{n}{n_C}\left[\frac{n_C}{n}-\left(\underbrace{Z_i-\frac{n_T}{n}}_{=A_i}\right)\right]y_{Ci}, \tag{7}$$

where A_i is the centered treatment variable, $A_i = Z_i - \mathrm{E}_\Omega[Z_i]$, because, under complete random assignment, $\mathrm{E}_\Omega[Z_i] = \dfrac{n_T}{n}$.

Therefore, we can re-write (7) as

$$\frac{1}{n}\sum_{i=1}^{n}\frac{n}{n_T}\left(A_i+\frac{n_T}{n}\right)y_{Ti}-\frac{n}{n_C}\left(\frac{n_C}{n}-A_i\right)y_{Ci},$$

which, with a bit more algebra, is equivalent to

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n}(y_{Ti}-y_{Ci})}_{=\tau_{\mathrm{SATE}}}+\frac{1}{n}\sum_{i=1}^{n}A_i\left(\frac{n}{n_T}y_{Ti}+\frac{n}{n_C}y_{Ci}\right)=\tau_{\mathrm{SATE}}+\frac{1}{n}\sum_{i=1}^{n}A_i\left(\frac{n}{n_T}y_{Ti}+\frac{n}{n_C}y_{Ci}\right)$$

Step 2

Thus far, we have shown that the Difference-in-Means estimator in (1) is equivalent to

$$\tau_{\mathrm{SATE}}+\frac{1}{n}\sum_{i=1}^{n}A_i\left(\frac{n}{n_T}y_{Ti}+\frac{n}{n_C}y_{Ci}\right), \tag{8}$$

in which τ_{SATE}, n_C, n_T, n, and $\{y_{Ci}, y_{Ti}\}_{i=1}^{n}$ are all fixed quantities. The only random quantities in (8) are $\{A_i\}_{i=1}^{n}$.

Therefore, to derive the variance of (8), we now need to derive $\mathrm{E}_\Omega[A_i]$, $\mathrm{Var}_\Omega[A_i]$ and $\mathrm{E}_\Omega\left[A_iA_j\right]$ for $i \neq j$.

To do so, let's write the sample space of A_i as follows:

$$A_i = Z_i - \frac{n_T}{n} = \begin{cases} \dfrac{n_C}{n} & \text{if } Z_i = 1 \\ -\dfrac{n_T}{n} & \text{if } Z_i = 0 \end{cases} \tag{9}$$

and recall that, under complete random assignment, $\Pr(Z_i = 1) = \dfrac{n_T}{n}$ and $\Pr(Z_i = 0) = 1 - \dfrac{n_T}{n} = \dfrac{n}{n} - \dfrac{n_T}{n} = \dfrac{n - n_T}{n} = \dfrac{n_C}{n}$.

Derivation of $\mathrm{E}_\Omega[A_i]$**:** It is straightforward to see that $\mathrm{E}_\Omega[A_i] = \mathrm{E}_\Omega\left[Z_i - \frac{n_T}{n}\right] = \mathrm{E}_\Omega[Z_i] - \frac{n_T}{n} = \frac{n_T}{n} - \frac{n_T}{n} = 0$ or, equivalently,

$$\begin{aligned}
\mathrm{E}_\Omega[A_i] &= \frac{n_C}{n}\Pr(Z_i = 1) - \frac{n_T}{n}\Pr(Z_i = 0) \\
&= \frac{n_C}{n}\frac{n_T}{n} - \frac{n_T}{n}\left(1 - \frac{n_T}{n}\right) \\
&= \frac{n_C}{n}\frac{n_T}{n} - \frac{n_T}{n}\frac{n_C}{n} \\
&= 0.
\end{aligned}$$

Derivation of $\mathrm{Var}_\Omega[A_i]$**:** Note that $\mathrm{Var}_\Omega[A_i] = \mathrm{E}_\Omega\left[A_i^2\right] - \mathrm{E}_\Omega[A_i]^2 = \mathrm{E}_\Omega\left[A_i^2\right] - 0$. Thus,

$$\begin{aligned}
\mathrm{Var}_\Omega[A_i] &= \mathrm{E}_\Omega\left[A_i^2\right] \\
&= \left(\frac{n_C}{n}\right)^2\Pr(Z_i = 1) + \left(-\frac{n_T}{n}\right)^2\Pr(Z_i = 0) \\
&= \frac{n_C^2}{n^2}\Pr(Z_i = 1) + \frac{n_T^2}{n^2}\Pr(Z_i = 0) \\
&= \frac{n_C^2}{n^2}\frac{n_T}{n} + \frac{n_T^2}{n^2}\frac{n_C}{n} \\
&= \frac{n_C^2 n_T}{n^3} + \frac{n_T^2 n_C}{n^3} \\
&= \frac{n_C^2 n_T + n_T^2 n_C}{n^3} \\
&= \frac{n_C n_T(n_C + n_T)}{n^3} \\
&= \frac{n_C n_T(n)}{n^3}
\end{aligned}$$

$$= \frac{n_C n_T}{n^2}.$$

Derivation of $\text{Cov}\left[A_i, A_j\right]$ **for** $i \neq j$**:** Note that the $\{A_i\}_{i=1}^{n}$ all have the same expected value, namely $\text{E}_\Omega\left[A_i\right] = 0$ for all i (as we just showed), but they are *not* independent, because fixing the total number of treated units at n_T links the assignment indicators to one another. We therefore will need to derive $\text{Cov}\left[A_i, A_j\right]$ for $i \neq j$. To do so, first note that

$$\begin{aligned}
\text{Cov}\left[A_i, A_j\right] &= \text{E}_\Omega\left[\left(A_i - \text{E}_\Omega\left[A_i\right]\right)\left(A_j - \text{E}_\Omega\left[A_j\right]\right)\right] \\
&= \text{E}_\Omega\left[\left(A_i - 0\right)\left(A_j - 0\right)\right] \\
&= \text{E}_\Omega\left[A_i A_j\right],
\end{aligned}$$

which is easier to work with.

The possible values that $A_i A_j$ could take on are

$$A_i A_j = \begin{cases} \left(\frac{n_C}{n}\right)\left(\frac{n_C}{n}\right) = \frac{n_C^2}{n^2} & \text{if } Z_i = 1 \text{ and } Z_j = 1 \\ \left(-\frac{n_T}{n}\right)\left(\frac{n_C}{n}\right) = \frac{-n_T n_C}{n^2} & \text{if } Z_i = 0 \text{ and } Z_j = 1 \text{ or } Z_i = 1 \text{ and } Z_j = 0 \\ \left(-\frac{n_T}{n}\right)\left(-\frac{n_T}{n}\right) = \frac{n_T^2}{n^2} & \text{if } Z_i = 0 \text{ and } Z_j = 0. \end{cases}$$

Having derived the sample space of $A_i A_j$, we now need to derive the probabilities that correspond to each of the events in the sample space of $A_i A_j$, namely, $\Pr\left(Z_i = 1, Z_j = 1\right)$, $\Pr\left(Z_i = 0, Z_j = 1\right)$, $\Pr\left(Z_i = 1, Z_j = 0\right)$ and $\Pr\left(Z_i = 0, Z_j = 0\right)$.

Remember that Z_i and Z_j are *not* independent, which implies that, e.g., $\Pr\left(Z_i = 1, Z_j = 1\right) \neq \Pr\left(Z_i = 1\right)\Pr\left(Z_j = 1\right)$. Instead, we will appeal to the definition of joint probability in which

$\Pr\left(Z_i = z, Z_j = z'\right) = \Pr\left(Z_i = z\right)\Pr\left(Z_j = z' \mid Z_i = z\right)$:

$$\Pr\left(Z_i = z, Z_j = z'\right) = \begin{cases} \left(\frac{n_T}{n}\right)\left(\frac{n_T - 1}{n-1}\right) & \text{if } Z_i = 1 \text{ and } Z_j = 1 \\ \left(\frac{n_T}{n}\right)\left(\frac{n_C}{n-1}\right) & \text{if } Z_i = 1 \text{ and } Z_j = 0 \\ \left(\frac{n_C}{n}\right)\left(\frac{n_T}{n-1}\right) & \text{if } Z_i = 0 \text{ and } Z_j = 1 \\ \left(\frac{n_C}{n}\right)\left(\frac{n_C}{n-1}\right) & \text{if } Z_i = 0 \text{ and } Z_j = 0. \end{cases}$$

We can therefore write the probability distribution function (PDF) of A_iA_j as

$$\Pr\left(A_iA_j\right) = \begin{cases} \left(\frac{n_T}{n}\right)\left(\frac{n_T - 1}{n-1}\right) = \frac{n_T\left(n_T - 1\right)}{n\left(n-1\right)} & \text{if } A_iA_j = \frac{n_C^2}{n^2} \\ \left(\frac{n_T}{n}\right)\left(\frac{n_C}{n-1}\right) + \left(\frac{n_C}{n}\right)\left(\frac{n_T}{n-1}\right) = \frac{2\left(n_Tn_C\right)}{n\left(n-1\right)} & \text{if } A_iA_j = \frac{-n_Tn_C}{n^2} \\ \left(\frac{n_C}{n}\right)\left(\frac{n_C}{n-1}\right) = \frac{n_C\left(n_C - 1\right)}{n\left(n-1\right)} & \text{if } A_iA_j = \frac{n_T^2}{n^2}, \end{cases}$$

which implies that

$$\mathrm{E}_\Omega\left[A_iA_j\right] = \frac{n_C^2}{n^2}\left(\frac{n_T\left(n_T - 1\right)}{n\left(n-1\right)}\right) + \frac{-n_Tn_C}{n^2}\left(\frac{2\left(n_Tn_C\right)}{n\left(n-1\right)}\right) + \frac{n_T^2}{n^2}\left(\frac{n_C\left(n_C - 1\right)}{n\left(n-1\right)}\right)$$

or equivalently, after some algebra,

$$\frac{-n_Tn_C}{n^2\left(n-1\right)}.$$

Step 3

In this step, we will prove the algebraic equivalence between the expression for $S_n^2(\boldsymbol{\tau})$ in Equation (5) above and

$$S_n^2(\boldsymbol{\tau}) = S_n^2(\boldsymbol{y_T}) + S_n^2(\boldsymbol{y_C}) - \frac{2}{n-1}\sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right), \tag{10}$$

which will be valuable for our derivation of $\mathrm{Var}_\Omega\left[\hat{\tau}\right]$ in **Step 4** to follow.

Recall from Equation (5) that

$$S_n^2(\boldsymbol{\tau}) = \frac{1}{n-1}\sum_{i=1}^{n}\left(\tau_i - \frac{1}{n}\sum_{i=1}^{n}\tau_i\right)^2 .$$

Hence,

$$\begin{aligned}
S_n^2(\boldsymbol{\tau}) &= \frac{1}{n-1}\sum_{i=1}^{n}\left(\tau_i - \frac{1}{n}\sum_{i=1}^{n}\tau_i\right)^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(\underbrace{(y_{Ti} - y_{Ci})}_{=\tau_i} - \frac{1}{n}\sum_{i=1}^{n}\underbrace{(y_{Ti} - y_{Ci})}_{=\tau_i}\right)^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n}y_{Ti}\right) - \left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n}y_{Ci}\right)\right)^2 ,
\end{aligned}$$

where, in the last line, we have used $\frac{1}{n}\sum_{i=1}^{n}(y_{Ti} - y_{Ci}) = \frac{1}{n}\sum_{i=1}^{n}y_{Ti} - \frac{1}{n}\sum_{i=1}^{n}y_{Ci}$ to regroup each summand into a treated deviation, $\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n}y_{Ti}\right)$, minus a control deviation, $\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n}y_{Ci}\right)$. Expanding the square via $(a-b)^2 = a^2 - 2ab + b^2$ and then distributing the sum, this is equivalent to

$$\begin{aligned}
S_n^2(\boldsymbol{\tau}) &= \underbrace{\frac{1}{n-1}\sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n}y_{Ti}\right)^2}_{=S_n^2(\boldsymbol{y_T})} + \underbrace{\frac{1}{n-1}\sum_{i=1}^{n}\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n}y_{Ci}\right)^2}_{=S_n^2(\boldsymbol{y_C})} \\
&\quad - \frac{2}{n-1}\sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n}y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n}y_{Ci}\right) \\
&= S_n^2(\boldsymbol{y_T}) + S_n^2(\boldsymbol{y_C}) - \frac{2}{n-1}\sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n}y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n}y_{Ci}\right) .
\end{aligned}$$

Step 4

Now, to complete the derivation of its variance, note that the variance of the Difference-in-Means estimator based on the expression of the Difference-in-Means in (8) is:

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \mathrm{Var}_\Omega\left[\tau_{\text{SATE}} + \frac{1}{n}\sum_{i=1}^{n} A_i \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)\right]$$

$$= \frac{1}{n^2}\,\mathrm{Var}_\Omega\left[\sum_{i=1}^{n} A_i \underbrace{\left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)}_{\text{constant}}\right],$$

which, since $\mathrm{E}_\Omega\left[A_i\right] = 0$, is equivalent to

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{1}{n^2}\left(\mathrm{E}_\Omega\left[\left(\sum_{i=1}^{n} A_i \underbrace{\left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)}_{\text{constant}}\right)^2\right] - \mathrm{E}_\Omega\left[\sum_{i=1}^{n} A_i \underbrace{\left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)}_{\text{constant}}\right]^2\right)$$

$$= \frac{1}{n^2}\,\mathrm{E}_\Omega\left[\left(\sum_{i=1}^{n} A_i \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)\right)^2\right].$$

Expanding the expression immediately above yields

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{1}{n^2}\,\mathrm{E}_\Omega\left[\left(\sum_{i=1}^{n} A_i \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)\right)^2\right]$$

$$= \frac{1}{n^2}\,\mathrm{E}_\Omega\left[\left(\sum_{i=1}^{n} A_i \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)\right)\left(\sum_{j=1}^{n} A_j \left(\frac{n}{n_T} y_{Tj} + \frac{n}{n_C} y_{Cj}\right)\right)\right]$$

$$= \frac{1}{n^2}\,\mathrm{E}_\Omega\left[\sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)\left(\frac{n}{n_T} y_{Tj} + \frac{n}{n_C} y_{Cj}\right)\right].$$

In **Step 2** above, we showed that

$$\begin{cases} \mathrm{E}_\Omega\left[A_i A_j\right] = \dfrac{-n_T n_C}{n^2\,(n-1)} & \text{when } i \neq j \\ \\ \mathrm{E}_\Omega\left[A_i A_j\right] = \mathrm{E}_\Omega\left[A_i^2\right] = \mathrm{Var}_\Omega\left[A_i\right] = \dfrac{n_C n_T}{n^2} & \text{when } i = j, \end{cases}$$

we now split the double sum $\sum_{i=1}^n \sum_{j=1}^n$ into the n diagonal terms (where $i = j$, so that $A_i A_j = A_i^2$) and the $n\,(n-1)$ off-diagonal terms (where $i \neq j$), because $\mathrm{E}_\Omega\left[A_i A_j\right]$ takes a different value in each case. The split yields

$$\begin{aligned} \mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{1}{n^2}\Bigg(& \sum_{i=1}^n \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)^2 \underbrace{\mathrm{E}_\Omega\left[A_i^2\right]}_{=\frac{n_C n_T}{n^2}} \\ & + \sum_{i=1}^n \sum_{j\neq i} \left(\frac{n}{n_T} y_{Ti} + \frac{n}{n_C} y_{Ci}\right)\left(\frac{n}{n_T} y_{Tj} + \frac{n}{n_C} y_{Cj}\right) \underbrace{\mathrm{E}_\Omega\left[A_i A_j\right]}_{=\frac{-n_T n_C}{n^2\,(n-1)}} \Bigg). \end{aligned}$$

We now substitute the expressions we derived for $\mathrm{E}_\Omega\left[A_i^2\right]$ and $\mathrm{E}_\Omega\left[A_i A_j\right]$. To keep the bookkeeping manageable, it helps to abbreviate the (constant) weight multiplying A_i as $w_i := \dfrac{n}{n_T} y_{Ti} + \dfrac{n}{n_C} y_{Ci}$, so that the expression above becomes

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{1}{n^2}\left(\frac{n_C n_T}{n^2} \sum_{i=1}^n w_i^2 - \frac{n_T n_C}{n^2\,(n-1)} \sum_{i=1}^n \sum_{j\neq i} w_i w_j\right).$$

Using the identity $\sum_{i=1}^n \sum_{j\neq i} w_i w_j = \left(\sum_{i=1}^n w_i\right)^2 - \sum_{i=1}^n w_i^2$ and then collecting terms over the common denominator $n-1$,

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{n_C n_T}{n^4\,(n-1)}\left(n \sum_{i=1}^n w_i^2 - \left(\sum_{i=1}^n w_i\right)^2\right) = \frac{n_C n_T}{n^3\,(n-1)} \sum_{i=1}^n \left(w_i - \bar{w}\right)^2,$$

where $\bar{w} = \frac{1}{n}\sum_{i=1}^n w_i$ and the final equality uses the standard identity $n \sum_i w_i^2 - \left(\sum_i w_i\right)^2 = n \sum_i \left(w_i - \bar{w}\right)^2$. Finally, substituting back the centered weight, $w_i - \bar{w} = \dfrac{n}{n_T}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^n y_{Ti}\right) +$

$\dfrac{n}{n_C}\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right)$, and expanding the square, yields

$$\begin{aligned}\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{n_C}{n n_T\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)^2 + \frac{n_T}{n n_C\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right)^2 \\ + \frac{2}{n\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right).\end{aligned} \tag{11}$$

Now recall that

$$S_n^2(\boldsymbol{y_T}) = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)^2 \quad \text{and}$$

$$S_n^2(\boldsymbol{y_C}) = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right)^2,$$

which yields

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{n_C}{n n_T} S_n^2(\boldsymbol{y_T}) + \frac{n_T}{n n_C} S_n^2(\boldsymbol{y_C}) + \frac{2}{n\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right).$$

Finally, at long last, it follows that

$$\begin{aligned}\mathrm{Var}_\Omega\left[\hat{\tau}\right] &= \frac{n_C}{n n_T} S_n^2(\boldsymbol{y_T}) + \frac{n_T}{n n_C} S_n^2(\boldsymbol{y_C}) + \frac{2}{n\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right) \\ &= \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{y_T})}{n} - \frac{S_n^2(\boldsymbol{y_C})}{n} + \frac{2}{n\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right) \\ &= \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{1}{n}\underbrace{\left(S_n^2(\boldsymbol{y_T}) + S_n^2(\boldsymbol{y_C}) - \frac{2}{\left(n-1\right)} \sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)\left(y_{Ci} - \frac{1}{n}\sum_{i=1}^{n} y_{Ci}\right)\right)}_{=S_n^2(\boldsymbol{\tau})}.\end{aligned}$$

As we showed in **Step 3**, the bracketed term in the expression immediately above is equal to $S_n^2(\boldsymbol{\tau})$, which leaves us with

$$\frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}. \tag{12}$$

$\square$

Each of the three terms in this variance has a clear meaning. The estimator $\hat{\tau}$ is the difference between the treated group mean and the control group mean, so its variability comes from two pieces. The first term, $\frac{S_n^2(\boldsymbol{y_T})}{n_T}$, is the variability contributed by the treated mean. It grows with the spread of the treatment potential outcomes, $S_n^2(\boldsymbol{y_T})$, and shrinks as we average over more treated units, n_T. The second term, $\frac{S_n^2(\boldsymbol{y_C})}{n_C}$, is the analogous variability contributed by the control mean. If the treated and control means were computed on two *independently* drawn groups, these first two terms would account for the entire variance.

The two means are *not* independent, however, and the third term, $-\frac{S_n^2(\boldsymbol{\tau})}{n}$, corrects for the dependence that complete random assignment induces. The treated and control groups are complementary, since every unit assigned to treatment is a unit *withheld* from control, so the treated mean and the control mean are negatively coupled. The coupling runs through the individual treatment effects, which is why the correction involves $S_n^2(\boldsymbol{\tau})$, the finite population variance of the individual effects $\tau_i = y_{Ti} - y_{Ci}$. Two consequences follow. First, the correction enters with a minus sign and therefore *lowers* the variance; ignoring the correction would overstate the variance, which is what makes the conservative variance estimator (developed in the next section) conservative rather than anti-conservative. Second, when the treatment effect is constant across units, τ_i is the same for all i, so $S_n^2(\boldsymbol{\tau}) = 0$, the correction vanishes, and the spreads of the two sets of potential outcomes fully determine the variance.

The expression we just derived for $\mathrm{Var}_\Omega\left[\hat{\tau}\right]$ is the *true* variance of the Difference-in-Means estimator. It is mathematically equivalent to Equation 3.4 in Gerber and Green (2012, 57), which differs from the expression in Equation (12) above because the variances and covariance of treated and control potential outcomes in Gerber and Green (2012, Equation 3.4, 57) use a denominator of n as opposed to $n-1$ as in Equations (3), (4) and (5) above. The corollary below establishes this equivalence.

Corollary 1. *An equivalent expression for the finite sample variance of the Difference-in-Means estimator under complete random assignment is*

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{1}{n-1}\left(\frac{n_C\sigma_n^2(\boldsymbol{y_T})}{n_T} + \frac{n_T\sigma_n^2(\boldsymbol{y_C})}{n_C} + 2\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T})\right), \tag{13}$$

where

$$\sigma_n^2(\boldsymbol{y_T}) = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(y_{Ti} - \frac{1}{n}\sum_{i=1}^{n} y_{Ti}\right)^2 \tag{14}$$

$$\sigma_n^2(\boldsymbol{y_C}) = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left(y_{Ci} - \frac{1}{n} \sum_{i=1}^{n} y_{Ci} \right)^2 \tag{15}$$

$$\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T}) = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left(y_{Ci} - \frac{1}{n} \sum_{i=1}^{n} y_{Ci} \right) \left(y_{Ti} - \frac{1}{n} \sum_{i=1}^{n} y_{Ti} \right). \tag{16}$$

Proof. Recall the expression for $\mathrm{Var}_\Omega\left[\hat{\tau}\right]$ in Equation (11):

$$\begin{aligned}\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{n_C}{n n_T\left(n-1\right)} \sum_{i=1}^{n} \left(y_{Ti} - \frac{1}{n} \sum_{i=1}^{n} y_{Ti} \right)^2 + \frac{n_T}{n n_C\left(n-1\right)} \sum_{i=1}^{n} \left(y_{Ci} - \frac{1}{n} \sum_{i=1}^{n} y_{Ci} \right)^2 \\ + \frac{2}{n\left(n-1\right)} \sum_{i=1}^{n} \left(y_{Ti} - \frac{1}{n} \sum_{i=1}^{n} y_{Ti} \right) \left(y_{Ci} - \frac{1}{n} \sum_{i=1}^{n} y_{Ci} \right).\end{aligned}$$

Then, appealing to the definitions $\sigma_n^2(\boldsymbol{y_T})$, $\sigma_n^2(\boldsymbol{y_C})$ and $\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T})$ above, yields

$$\begin{aligned}\mathrm{Var}_\Omega\left[\hat{\tau}\right] &= \frac{n_C}{n_T\left(n-1\right)}\sigma_n^2(\boldsymbol{y_T}) + \frac{n_T}{n_C\left(n-1\right)}\sigma_n^2(\boldsymbol{y_C}) + \frac{2}{\left(n-1\right)}\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T}) \\ &= \frac{1}{n-1}\left(\frac{n_C\sigma_n^2(\boldsymbol{y_T})}{n_T} + \frac{n_T\sigma_n^2(\boldsymbol{y_C})}{n_C} + 2\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T}) \right),\end{aligned}$$

which completes the proof. □

2 Conservative estimation of the variance

Having derived the true variance $\mathrm{Var}_\Omega\left[\hat{\tau}\right]$, we now turn to estimating it from the observed data. Absent additional assumptions, we *cannot* estimate $\mathrm{Var}_\Omega\left[\hat{\tau}\right]$ unbiasedly. The obstruction is the same in both equivalent expressions for the variance. Each expression depends on a quantity that couples a unit's two potential outcomes, namely the covariance $\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T})$ in the Gerber–Green form or the effect variance $S_n^2(\boldsymbol{\tau})$ in the Imbens–Rubin form, and we never observe both potential outcomes for any single unit (the fundamental problem of causal inference). What we *can* do is construct an estimable quantity that is at least as large as the true variance and estimate that quantity instead. The resulting estimator is *conservative*, in the sense that on average it weakly overstates the variance, which yields valid confidence intervals (possibly too wide) rather than overconfident ones.

We establish the upper bound in two ways, starting from each of the two equivalent expressions for the variance. Both routes arrive at the *same* conservative estimator. We first record the two expressions for reference. The Imbens–Rubin form (Proposition 1) is

$$\operatorname{Var}_\Omega[\hat{\tau}] = \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}, \tag{17}$$

and the Gerber–Green form (the Corollary above) is

$$\operatorname{Var}_\Omega[\hat{\tau}] = \frac{1}{n-1}\left(\frac{n_C\sigma_n^2(\boldsymbol{y_T})}{n_T} + \frac{n_T\sigma_n^2(\boldsymbol{y_C})}{n_C} + 2\sigma_n(\boldsymbol{y_C},\boldsymbol{y_T})\right). \tag{18}$$

2.1 Approach 1 (Gerber–Green expression): bounding the covariance

In Equation (18), we can estimate the variances $\sigma_n^2(\boldsymbol{y_T})$ and $\sigma_n^2(\boldsymbol{y_C})$ unbiasedly, but we cannot estimate the covariance $\sigma_n(\boldsymbol{y_C},\boldsymbol{y_T})$. The first approach replaces the unestimable covariance with an estimable upper bound. The Cauchy–Schwarz inequality implies that

$$\sigma_n(\boldsymbol{y_C},\boldsymbol{y_T}) \le \sqrt{\sigma_n^2(\boldsymbol{y_C})\sigma_n^2(\boldsymbol{y_T})},$$

and the AM–GM inequality further implies that

$$\sqrt{\sigma_n^2(\boldsymbol{y_C})\sigma_n^2(\boldsymbol{y_T})} \le \frac{\sigma_n^2(\boldsymbol{y_C}) + \sigma_n^2(\boldsymbol{y_T})}{2}.$$

Hence $2\sigma_n(\boldsymbol{y_C},\boldsymbol{y_T}) \le \sigma_n^2(\boldsymbol{y_C}) + \sigma_n^2(\boldsymbol{y_T})$. Substituting this bound for the covariance term in Equation (18), and then collecting the two terms in each potential outcome's variance, yields a simple upper bound. The key simplification is that pairing the cross term with the matching squared term gives, for example,

$$\frac{n_C\,\sigma_n^2(\boldsymbol{y_T})}{n_T} + \sigma_n^2(\boldsymbol{y_T}) = \frac{n_C + n_T}{n_T}\,\sigma_n^2(\boldsymbol{y_T}) = \frac{n\,\sigma_n^2(\boldsymbol{y_T})}{n_T},$$

and likewise for the control variance. Carrying this out,

$$\begin{aligned}
\operatorname{Var}_\Omega[\hat{\tau}] &= \frac{1}{n-1}\left(\frac{n_C\sigma_n^2(\boldsymbol{y_T})}{n_T} + \frac{n_T\sigma_n^2(\boldsymbol{y_C})}{n_C} + 2\sigma_n(\boldsymbol{y_C},\boldsymbol{y_T})\right) \\
&\le \frac{1}{n-1}\left(\frac{n_C\sigma_n^2(\boldsymbol{y_T})}{n_T} + \frac{n_T\sigma_n^2(\boldsymbol{y_C})}{n_C} + \sigma_n^2(\boldsymbol{y_T}) + \sigma_n^2(\boldsymbol{y_C})\right) \\
&= \frac{1}{n-1}\left(\underbrace{\frac{n_C\sigma_n^2(\boldsymbol{y_T})}{n_T} + \sigma_n^2(\boldsymbol{y_T})}_{=n\sigma_n^2(\boldsymbol{y_T})/n_T} + \underbrace{\frac{n_T\sigma_n^2(\boldsymbol{y_C})}{n_C} + \sigma_n^2(\boldsymbol{y_C})}_{=n\sigma_n^2(\boldsymbol{y_C})/n_C}\right)
\end{aligned}$$

$$= \frac{n}{n-1}\left(\frac{\sigma_n^2(\boldsymbol{y_T})}{n_T} + \frac{\sigma_n^2(\boldsymbol{y_C})}{n_C}\right).$$

We can estimate the two remaining unknowns, $\sigma_n^2(\boldsymbol{y_T})$ and $\sigma_n^2(\boldsymbol{y_C})$, unbiasedly. Following Cochran (1977, Theorem 2.4), unbiased estimators are

$$\widehat{\sigma}_n^2(\boldsymbol{y_T}) = \left(\frac{n-1}{n\,(n_T-1)}\right)\sum_{i=1}^{n} Z_i \left(Y_i - \frac{1}{n_T}\sum_{i=1}^{n} Z_i Y_i\right)^2$$

$$\widehat{\sigma}_n^2(\boldsymbol{y_C}) = \left(\frac{n-1}{n\,(n_C-1)}\right)\sum_{i=1}^{n}(1-Z_i) \left(Y_i - \frac{1}{n_C}\sum_{i=1}^{n}(1-Z_i) Y_i\right)^2.$$

These estimators depend only on the *observed* outcomes. The factor Z_i selects the treated units, for which $Y_i = y_{Ti}$, and $(1-Z_i)$ selects the control units, for which $Y_i = y_{Ci}$. The leading factor $\frac{n-1}{n}$ converts the ordinary sample variance, which is unbiased for $S_n^2(\cdot)$, into an unbiased estimator of $\sigma_n^2(\cdot)$, since $\sigma_n^2(\cdot) = \frac{n-1}{n} S_n^2(\cdot)$. Substituting these into the upper bound gives the conservative variance estimator

$$\widehat{\operatorname{Var}}_\Omega\left[\hat{\tau}\right] = \frac{n}{n-1}\left(\frac{\widehat{\sigma}_n^2(\boldsymbol{y_T})}{n_T} + \frac{\widehat{\sigma}_n^2(\boldsymbol{y_C})}{n_C}\right). \tag{19}$$

2.2 Approach 2 (Imbens–Rubin expression): dropping the effect variance term

The second approach starts from Equation (17) and is more direct. The only term we cannot estimate is $S_n^2(\boldsymbol{\tau})$, the finite population variance of the individual effects, since computing $S_n^2(\boldsymbol{\tau})$ would require knowing both potential outcomes for each unit. Notice that $S_n^2(\boldsymbol{\tau})$ enters with a *minus* sign and is always nonnegative. Therefore

$$\operatorname{Var}_\Omega\left[\hat{\tau}\right] = \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \underbrace{\frac{S_n^2(\boldsymbol{\tau})}{n}}_{\geq 0} \leq \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C}.$$

That is, simply *dropping* the nonnegative term we cannot estimate produces an upper bound. The two remaining terms depend on one potential outcome per unit and can be estimated unbiasedly by the corresponding sample variances within each group,

$$\widehat{S}_n^2(\boldsymbol{y_T}) = \left(\frac{1}{n_T-1}\right)\sum_{i=1}^{n} Z_i \left(Y_i - \frac{1}{n_T}\sum_{i=1}^{n} Z_i Y_i\right)^2$$

$$\widehat{S}_n^2(\boldsymbol{y_C}) = \left(\frac{1}{n_C-1}\right)\sum_{i=1}^{n}(1-Z_i) \left(Y_i - \frac{1}{n_C}\sum_{i=1}^{n}(1-Z_i) Y_i\right)^2,$$

which are well defined because $n_T \geq 2$ and $n_C \geq 2$. Substituting these into the upper bound gives the conservative variance estimator

$$\widehat{\mathrm{Var}}_\Omega\left[\hat{\tau}\right] = \frac{1}{n_T}\widehat{S}_n^2(\boldsymbol{y_T}) + \frac{1}{n_C}\widehat{S}_n^2(\boldsymbol{y_C}). \tag{20}$$

2.3 The two approaches give the same estimator, and how conservative it is

The two routes produce the *same* estimator. To see the equivalence, recall that $\widehat{\sigma}_n^2(\boldsymbol{y_T}) = \frac{n-1}{n}\widehat{S}_n^2(\boldsymbol{y_T})$ (compare the two pairs of estimators above), so that

$$\frac{n}{n-1}\frac{\widehat{\sigma}_n^2(\boldsymbol{y_T})}{n_T} = \frac{n}{n-1}\cdot\frac{n-1}{n}\cdot\frac{\widehat{S}_n^2(\boldsymbol{y_T})}{n_T} = \frac{\widehat{S}_n^2(\boldsymbol{y_T})}{n_T},$$

and likewise for the control term. Hence Equation (19) and Equation (20) are identical; the Cauchy–Schwarz/AM–GM bound on the covariance and the dropping of the nonnegative effect variance term are two ways of describing the same approximation.

Finally, the conservative estimator overstates the variance by an amount we can compute exactly. Because $\widehat{S}_n^2(\boldsymbol{y_T})$ and $\widehat{S}_n^2(\boldsymbol{y_C})$ are unbiased for $S_n^2(\boldsymbol{y_T})$ and $S_n^2(\boldsymbol{y_C})$,

$$\mathrm{E}_\Omega\left[\widehat{\mathrm{Var}}_\Omega\left[\hat{\tau}\right]\right] = \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} = \mathrm{Var}_\Omega\left[\hat{\tau}\right] + \frac{S_n^2(\boldsymbol{\tau})}{n},$$

where the last equality uses Equation (17). Therefore

$$\mathrm{E}_\Omega\left[\widehat{\mathrm{Var}}_\Omega\left[\hat{\tau}\right]\right] - \mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{S_n^2(\boldsymbol{\tau})}{n} \geq 0,$$

so the estimator's upward bias is exactly $S_n^2(\boldsymbol{\tau})/n$. The estimator is conservative in general, and it is *exactly* unbiased when the individual treatment effects are constant across units, i.e., when $S_n^2(\boldsymbol{\tau}) = 0$ (equivalently, when $2\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T}) = \sigma_n^2(\boldsymbol{y_C}) + \sigma_n^2(\boldsymbol{y_T})$, so that the Cauchy–Schwarz/AM–GM bound in Approach 1 holds with equality).

References

Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons.

Cox, D. R. (1958). *Planning of Experiments*. New York: John Wiley & Sons.

Freedman, D. A. (2009). *Statistical Models: Theory and Practice* (Revised ed.). Cambridge:

Cambridge University Press.

Gerber, A. S. and D. P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation.* New York: W. W. Norton.

Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* New York: Cambridge University Press.

Rubin, D. B. (1980). Comment on "randomization analysis of experimental data: The Fisher randomization test" by D. Basu. *Journal of the American Statistical Association 75*(371), 591–593.

Rubin, D. B. (1986). Comment: Which ifs have causal answers. *Journal of the American Statistical Association 81*(396), 961–962.