Variance and Conservative Estimation of the Difference-in-Means Estimator (PATE)[*]

Thomas Leavitt

This guide extends the finite sample results to a superpopulation. It derives the variance of the Difference-in-Means estimator for the Population Average Treatment Effect (PATE) and then shows that the *same* variance estimator that is merely *conservative* for the Sample Average Treatment Effect (SATE) is *exactly unbiased* for the PATE. The guide is self-contained, but it builds directly on the companion guides "Notation and Setup: The Finite Population Potential Outcomes Framework" and "Variance and Conservative Estimation of the Difference-in-Means Estimator (SATE)," whose key results we recall as needed.

1 Setup

Thus far (in the companion SATE guides) we have considered estimation in only a finite sample. Now consider a superpopulation, $\mathcal{P}_N$, of size $N \geq n \geq 4$, and let the index $i \in \{1, \dots, N\}$ run over the N units in $\mathcal{P}_N$. Let n units from $\mathcal{P}_N$ be randomly selected into an experimental sample, $\mathcal{S}_n$, by simple random sampling, while the remaining $N - n$ units in $\mathcal{P}_N$ are unsampled. Of these n sampled units, $n_T \geq 2$ are randomly assigned to treatment and $n - n_T = n_C \geq 2$ are randomly assigned to control by complete random assignment (at least two units per arm, and hence $n \geq 4$, so that each arm's sample variance in the conservative variance estimator is well defined). As in the finite sample setting, each unit i has two *fixed* potential outcomes, y_{Ti} and y_{Ci}, with individual effect $\tau_i = y_{Ti} - y_{Ci}$, and the observed outcome is $Y_i = Z_i y_{Ti} + (1 - Z_i)\, y_{Ci}$.

---

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

The binary indicator variable $R_i \in \{0,1\}$ denotes whether unit i is included ($R_i = 1$) or excluded ($R_i = 0$) in the random sample from the superpopulation $\mathcal{P}_N$. Let the set $\Pi = \left\{ \boldsymbol{r} \in \{0,1\}^N : \sum_{i=1}^N r_i = n \right\}$ contain all possible values of $\boldsymbol{R} = \left[R_1, \ldots, R_N \right]^\top$.

The target of interest is the Population Average Treatment Effect (PATE), $\tau_{\text{PATE}} \coloneqq N^{-1} \sum\limits_{i=1}^{N} \tau_i$. Define the Difference-in-Means estimator of τ_{PATE} as

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

Here the product $R_i Z_i$ equals 1 only for units that are both sampled and assigned to treatment, and $R_i \left(1 - Z_i\right)$ equals 1 only for units that are both sampled and assigned to control; unsampled units have $R_i = 0$ and so contribute nothing, regardless of Z_i. Otherwise the setup is identical to the finite sample setup, except that now, under simple random sampling of n units from a population of size N and complete random assignment with n_T treated units out of n sampled units, there are $\binom{N}{n}\binom{n}{n_T}$ possible joint configurations: the $\binom{n}{n_T}$ assignments for any single realized sample, times the $\binom{N}{n}$ samples that could be realized.

The PATE estimator in (1) has *two* sources of randomness, $\{R_i\}_{i=1}^N$ and $\{Z_i\}_{i=1}^n$, reflecting the random sampling process and the random assignment process. By contrast, the SATE estimator has only the single source $\{Z_i\}_{i=1}^n$. For the overall expectation and variance of the PATE estimator we write $\mathrm{E}\left[\cdot\right]$ and $\mathrm{Var}\left[\cdot\right]$; we write $\mathrm{E}_\Omega\left[\cdot\right], \mathrm{Var}_\Omega\left[\cdot\right]$ for expectation and variance over the assignment process (conditional on the sample), and $\mathrm{E}_\Pi\left[\cdot\right], \mathrm{Var}_\Pi\left[\cdot\right]$ for those over the sampling process. By the law of total variance, the variance of the PATE estimator is

$$\mathrm{Var}\left[\hat{\tau}\right] = \mathrm{E}_\Pi\left[\mathrm{Var}_\Omega\left[\hat{\tau}\right]\right] + \mathrm{Var}_\Pi\left[\mathrm{E}_\Omega\left[\hat{\tau}\right]\right]. \tag{2}$$

We will also need two ingredients from the finite sample (SATE) analysis. First, conditional on the realized sample, the finite sample variances of the potential outcomes (computed over the n sampled units, with divisor $n-1$) are

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

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

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

Second, as derived in the companion SATE variance guide, conditional on the realized sample the assignment variance of $\hat{\tau}$ is

$$\mathrm{Var}_\Omega\left[\hat{\tau}\right] = \frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}, \tag{6}$$

and $\hat{\tau}$ is unbiased for the (sample) SATE, $\mathrm{E}_\Omega\left[\hat{\tau}\right] = \tau_{\mathrm{SATE}}$.

2 Variance of the Difference-in-Means estimator for the PATE

Proposition 1. *The variance of $\hat{\tau}$ for τ_{PATE} under simple random sampling from the units in $\mathcal{P}_N$ and complete random assignment among the units in $\mathcal{S}_n$ is*

$$\mathrm{Var}\left[\hat{\tau}\right] = \frac{N}{N-1}\left(\frac{\sigma_N^2(\boldsymbol{y_T})}{n_T} + \frac{\sigma_N^2(\boldsymbol{y_C})}{n_C} - \frac{\sigma_N^2(\boldsymbol{\tau})}{N}\right), \tag{7}$$

where

$$\sigma_N^2(\boldsymbol{y_T}) = \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(y_{Ti} - \frac{1}{N}\sum_{i=1}^{N}y_{Ti}\right)^2$$

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

$$\sigma_N^2(\boldsymbol{\tau}) = \frac{1}{N}\sum_{i=1}^{N}\left(y_{Ti} - y_{Ci} - \frac{1}{N}\sum_{i=1}^{N}\tau_i\right)^2.$$

Proof. By the law of total variance, Equation (2), the variance of the Difference-in-Means estimator of the PATE is

$$\mathrm{Var}\left[\hat{\tau}\right] = \mathrm{E}_\Pi\left[\mathrm{Var}_\Omega\left[\hat{\tau}\right]\right] + \mathrm{Var}_\Pi\left[\mathrm{E}_\Omega\left[\hat{\tau}\right]\right].$$

Since $\hat{\tau}$ is unbiased for τ_{SATE} over Ω, we have $\mathrm{E}_\Omega\left[\hat{\tau}\right] = \tau_{\mathrm{SATE}}$. Here we must be careful about a shift in what is random. In the superpopulation setting, the sample average treatment effect $\tau_{\mathrm{SATE}} = \frac{1}{n}\sum_{i=1}^{N} R_i\tau_i$ is *itself* a random quantity, because it depends on which n units happen to be sampled and so varies over the sampling distribution Π. (Within a fixed sample, as in the SATE guide, it is a fixed constant.) This variability of τ_{SATE} across samples is exactly the second component, $\text{Var}_\Pi\left[\text{E}_\Omega\left[\hat{\tau}\right]\right] = \text{Var}_\Pi\left[\tau_{\text{SATE}}\right]$, in the law of total variance. Using the assignment variance recalled in (6), it follows that

$$\begin{aligned}\text{Var}\left[\hat{\tau}\right] &= \text{E}_\Pi\left[\text{Var}_\Omega\left[\hat{\tau}\right]\right] + \text{Var}_\Pi\left[\text{E}_\Omega\left[\hat{\tau}\right]\right] \\ &= \text{E}_\Pi\left[\frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}\right] + \text{Var}_\Pi\left[\tau_{\text{SATE}}\right].\end{aligned}$$

Therefore, we need to derive $\text{E}_\Pi\left[\frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}\right]$ and $\text{Var}_\Pi\left[\tau_{\text{SATE}}\right]$.

Elementary theory from survey sampling (Cochran, 1977; Kish, 1965; Lohr, 2010) implies that

$$\text{Var}_\Pi\left[\tau_{\text{SATE}}\right] = \text{Var}_\Pi\left[\frac{1}{n}\sum_{i=1}^{N} R_i\tau_i\right] = \frac{N-n}{(N-1)}\frac{\sigma_N^2(\boldsymbol{\tau})}{n}.$$

Now we only need to derive $\text{E}_\Pi\left[\frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}\right]$:

$$\text{E}_\Pi\left[\frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}\right] = \frac{1}{n_T}\text{E}_\Pi\left[S_n^2(\boldsymbol{y_T})\right] + \frac{1}{n_C}\text{E}_\Pi\left[S_n^2(\boldsymbol{y_C})\right] - \frac{1}{n}\text{E}_\Pi\left[S_n^2(\boldsymbol{\tau})\right],$$

since, under simple random sampling with a fixed n and complete random assignment, n_T, n_C and n are all fixed constants. Recalling the definitions of $S_n^2(\boldsymbol{y_T})$, $S_n^2(\boldsymbol{y_C})$ and $S_n^2(\boldsymbol{\tau})$ in (3), (4) and (5), we can re-write each as a function of the sample inclusion indicators $\{R_i\}_{i=1}^N$:

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

which make explicit that the randomness of these three sample variances (all of which would be fixed in a finite sample setting) stems from the N sample inclusion variables $\{R_i\}_{i=1}^N$. They remain *sample* variances, computed over the n sampled units; we keep the subscript n to distinguish them from the population variances $\sigma_N^2(\cdot)$, which are computed over all N units in $\mathcal{P}_N$.

Adapting Cochran (1977, Theorem 2.4), the standard result that the sample variance is unbiased for the population variance gives

$$\mathrm{E}_{\Pi}\left[S_n^2(\boldsymbol{y_T})\right] = \frac{N}{N-1}\sigma_N^2(\boldsymbol{y_T})$$

$$\mathrm{E}_{\Pi}\left[S_n^2(\boldsymbol{y_C})\right] = \frac{N}{N-1}\sigma_N^2(\boldsymbol{y_C})$$

$$\mathrm{E}_{\Pi}\left[S_n^2(\boldsymbol{\tau})\right] = \frac{N}{N-1}\sigma_N^2(\boldsymbol{\tau}),$$

which implies that

$$\mathrm{E}_{\Pi}\left[\frac{S_n^2(\boldsymbol{y_T})}{n_T} + \frac{S_n^2(\boldsymbol{y_C})}{n_C} - \frac{S_n^2(\boldsymbol{\tau})}{n}\right] = \frac{N}{N-1}\left(\frac{\sigma_N^2(\boldsymbol{y_T})}{n_T} + \frac{\sigma_N^2(\boldsymbol{y_C})}{n_C} - \frac{\sigma_N^2(\boldsymbol{\tau})}{n}\right).$$

With these two expressions in hand, it follows that

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

□

As we can see from Equation (7), as $N \to \infty$ (with the population variances bounded), the factor $\frac{N}{N-1} \to 1$ and the term $\frac{\sigma_N^2(\boldsymbol{\tau})}{N} \to 0$. Write $\sigma^2(\boldsymbol{y_T})$, $\sigma^2(\boldsymbol{y_C})$ and $\sigma^2(\boldsymbol{\tau})$ for the superpopulation variances of the treatment potential outcomes, the control potential outcomes, and the individual effects (the limiting values of $\sigma_N^2(\cdot)$ as the superpopulation grows, equivalently the variances of a single unit's quantities under independent sampling from the superpopulation). In that limit the variance of the PATE estimator becomes

$$\operatorname{Var}[\hat{\tau}] = \frac{\sigma^2(\boldsymbol{y_T})}{n_T} + \frac{\sigma^2(\boldsymbol{y_C})}{n_C}. \tag{8}$$

Equation (8) is the variance of the Difference-in-Means estimator about the PATE under the standard superpopulation (or "infinite population") sampling model, in which we regard the n study units as independent draws from the superpopulation. Unlike the finite sample (SATE) variance, Equation (8) has *no* effect heterogeneity correction term. The additional sampling variability of $\hat{\tau}$ about the PATE absorbs the $-\sigma^2(\boldsymbol{\tau})/\cdot$ term that the finite population expression carries. The disappearance of that correction term is exactly why, as we now show, the conservative estimator is not merely conservative for the PATE but *exactly unbiased.*

3 Estimating the variance for the PATE: the conservative estimator is exactly unbiased

In the finite sample (SATE) analysis, the *same* estimator we are about to use was *conservative*, overstating the SATE variance by exactly $S_n^2(\boldsymbol{\tau})/n$. We now establish the complementary fact that the same estimator is *exactly unbiased*, not merely conservative, for the variance of the PATE estimator. No new estimator is needed. Recall the estimator from the SATE guide,

$$\widehat{\operatorname{Var}}[\hat{\tau}] = \frac{1}{n_T}\widehat{S}_n^2(\boldsymbol{y_T}) + \frac{1}{n_C}\widehat{S}_n^2(\boldsymbol{y_C}), \tag{9}$$

where the sample variances within each group, computed from the sampled units, are

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

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

Proposition 2. *Suppose the n study units are obtained by independent random sampling from the superpopulation (equivalently, take $N \to \infty$ in the finite population setup above) and treatment is assigned by complete random assignment. Then the estimator $\widehat{\operatorname{Var}}[\hat{\tau}]$ in (9) is exactly unbiased for the variance of the PATE estimator:*

$$\operatorname{E}\left[\widehat{\operatorname{Var}}[\hat{\tau}]\right] = \frac{\sigma^2(\boldsymbol{y_T})}{n_T} + \frac{\sigma^2(\boldsymbol{y_C})}{n_C} = \operatorname{Var}[\hat{\tau}] .$$

That is, the same estimator that is conservative for the SATE is exactly right for the PATE.

Proof. We compute $\mathrm{E}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right]$ by iterating expectations over the assignment process (conditional on the realized sample) and then the sampling process, $\mathrm{E}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right] = \mathrm{E}_{\Pi}\left[\mathrm{E}_{\Omega}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right]\right]$. Three facts do the work.

Fact 1 (the conditional, finite sample result). Condition on the realized sample. Within it, the analysis is exactly the finite sample (SATE) analysis, and the sample variances within each group are unbiased for the sample variances over the n sampled units (Cochran, 1977, Theorem 2.4), so $\mathrm{E}_{\Omega}\left[\widehat{S}^2_n(\boldsymbol{y_T})\right] = S^2_n(\boldsymbol{y_T})$ and $\mathrm{E}_{\Omega}\left[\widehat{S}^2_n(\boldsymbol{y_C})\right] = S^2_n(\boldsymbol{y_C})$. Hence the estimator is *conditionally* conservative:

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

where the second equality is the SATE variance (6). The conditional overstatement is exactly $S^2_n(\boldsymbol{\tau})/n$.

Fact 2 (law of total variance). The PATE variance decomposes as

$$\mathrm{Var}\left[\hat{\tau}\right] = \mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \mathrm{Var}_{\Pi}\left[\mathrm{E}_{\Omega}\left[\hat{\tau}\right]\right] = \mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \mathrm{Var}_{\Pi}\left[\tau_{\mathrm{SATE}}\right],$$

using $\mathrm{E}_{\Omega}\left[\hat{\tau}\right] = \tau_{\mathrm{SATE}}$.

Fact 3 (i.i.d. sampling moments). Because the n study units are independent draws from the superpopulation, the sample average treatment effect $\tau_{\mathrm{SATE}} = \frac{1}{n}\sum_{i\in\mathcal{S}_n}\tau_i$ is the mean of n independent draws of the individual effect, and $S^2_n(\boldsymbol{\tau})$ is the corresponding sample variance. Hence

$$\mathrm{Var}_{\Pi}\left[\tau_{\mathrm{SATE}}\right] = \frac{\sigma^2(\boldsymbol{\tau})}{n} \quad \text{and} \quad \mathrm{E}_{\Pi}\left[S^2_n(\boldsymbol{\tau})\right] = \sigma^2(\boldsymbol{\tau}),$$

both of which are standard facts for an i.i.d. sample, the first for the variance of a sample mean and the second for the unbiasedness of the sample variance. (The same reasoning gives $\mathrm{E}_{\Pi}\left[S^2_n(\boldsymbol{y_T})\right] = \sigma^2(\boldsymbol{y_T})$ and $\mathrm{E}_{\Pi}\left[S^2_n(\boldsymbol{y_C})\right] = \sigma^2(\boldsymbol{y_C})$.)

Now take the expectation of Fact 1 over the sampling distribution:

$$\mathrm{E}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right] = \mathrm{E}_{\Pi}\left[\mathrm{E}_{\Omega}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right]\right] = \mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right] + \frac{S^2_n(\boldsymbol{\tau})}{n}\right] = \mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \frac{\mathrm{E}_{\Pi}\left[S^2_n(\boldsymbol{\tau})\right]}{n} = \mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \frac{\sigma^2(\boldsymbol{\tau})}{n},$$

using Fact 3 in the last step. Subtracting the PATE variance from Fact 2,

$$\mathrm{E}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right] - \mathrm{Var}\left[\hat{\tau}\right] = \left(\mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \frac{\sigma^2(\boldsymbol{\tau})}{n}\right) - \left(\mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \mathrm{Var}_{\Pi}\left[\tau_{\mathrm{SATE}}\right]\right)$$

$$= \frac{\sigma^2(\boldsymbol{\tau})}{n} - \mathrm{Var}_{\Pi}\left[\tau_{\mathrm{SATE}}\right]$$

$$= \frac{\sigma^2(\boldsymbol{\tau})}{n} - \frac{\sigma^2(\boldsymbol{\tau})}{n} = 0,$$

again by Fact 3. Hence $\mathrm{E}\left[\widehat{\mathrm{Var}}\left[\hat{\tau}\right]\right] = \mathrm{Var}\left[\hat{\tau}\right]$. Computing the common value directly, take expectations of Fact 1 and use $\mathrm{E}_{\Pi}\left[S_n^2(\boldsymbol{y_T})\right] = \sigma^2(\boldsymbol{y_T})$ and $\mathrm{E}_{\Pi}\left[S_n^2(\boldsymbol{y_C})\right] = \sigma^2(\boldsymbol{y_C})$:

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

the last equality by (8). ☐

The cancellation in the proof has a simple interpretation. Conditional on the sample, the conservative estimator overshoots the SATE variance by exactly $S_n^2(\boldsymbol{\tau})/n$, the effect heterogeneity penalty that makes the estimator conservative for the SATE. The PATE estimator must also contend with a source of variability that the SATE estimator does not, namely the randomness in which units are sampled, which makes the sample's SATE fluctuate about the PATE with variance $\mathrm{Var}_{\Pi}\left[\tau_{\mathrm{SATE}}\right] = \sigma^2(\boldsymbol{\tau})/n$. In expectation the conditional overstatement $S_n^2(\boldsymbol{\tau})/n$ equals the extra sampling variance $\sigma^2(\boldsymbol{\tau})/n$. The slack that makes the estimator conservative for the SATE is precisely the variability the estimator must absorb to be exactly right for the PATE. So the same estimator is conservative for the SATE and exactly unbiased for the PATE.

The finite population case differs slightly from the exact unbiasedness just established. Suppose that, instead of independent draws from the superpopulation, the n units are obtained by simple random sampling *without replacement* from a *finite* population of fixed size N, the setting of Theorem 1. Then the two quantities above no longer cancel exactly. A finite population correction leaves a residual $\mathrm{E}_{\Pi}\left[S_n^2(\boldsymbol{\tau})\right]/n - \mathrm{Var}_{\Pi}\left[\tau_{\mathrm{SATE}}\right] = \sigma_N^2(\boldsymbol{\tau})/(N-1) \geq 0$, so the estimator is mildly conservative, with bias $\sigma_N^2(\boldsymbol{\tau})/(N-1)$. This residual vanishes as $N \to \infty$, which recovers the exact unbiasedness above. The exact result is therefore the one that pertains to the PATE proper, the average effect in the (effectively infinite) superpopulation from which study samples are drawn.

References

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

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

Lohr, S. L. (2010). *Sampling: Design and Analysis* (2nd ed.). Boston: Brooks/Cole.