Technical Appendix: Improving precision through design and analysis in experiments with noncompliance

TA-1 Formal Efficiency Gain Comparisons

In this section we formalize the theoretical precision gains under blocking, as well as the relative gains of the principal ignorability approaches over IV. In line with much of the blocking literature (see Nickerson (2005), Miratrix, Sekhon, and Yu (2013), etc.), we demonstrate that the proposed blocking design can help greatly improve statistical precision over complete randomization, and can help off-set some of the precision loss associated with low compliance rates. More specifically, when block principal ignorability holds, blocking can provide precision gains over principal score weighted estimators. Furthermore, in situations where we cannot safely assume block principal ignorability, block-randomized design using variables related to both outcome and compliance at the design stage can help increase the precision of the IV estimator for the CACE.

In the following two subsections, we define $N = 2n$, such that there are $2n$ total units in the experiment. Assume the number of units randomly assigned to encouragement and control is fixed and equal (i.e., there are n units in encouragement, and n units in control). We will consider a matched-pairs setting. As such, the number of blocks is also equal to n (i.e., $B = n$). Consistent with Feller, Mealli, and Miratrix (2017), we assume the principal scores are known (and fixed). Finally, define p as the average probability of compliance. The variances are derived under a finite-sample framework, consistent with Imbens and Rubin (2015). The overall results can be extended to a super-population framework with slight modifications to the theoretical expressions, depending on what sampling design is assumed. See Pashley and Miratrix (2021b) and Pashley and Miratrix (2021a) for more discussion.

TA-1.1 Principal Ignorability Approaches versus Instrumental Variables

In the following subsection, we show that the relative efficiency gain from using principal ignorability approaches over instrumental variables. To directly compare the blocked difference-in-means estimator with the alternative approaches (i.e., principal score weighting and IV), we assume that researchers are using a blocking design.

To begin, we demonstrate that the block difference-in-means estimator can provide efficiency gains over the block IV estimator:

Theorem TA-1.1 (Relative Reduction in Variance from Block DiM over Block IV) *The difference in variance between a blocked IV estimator and the block difference-in-means estimator is approximately:*

$$var(\hat{\tau}_B^{IV}) - var(\hat{\tau}_B) \approx \frac{1}{(np)^2}\left(\sum_{b=1}^{B}\big(1 - P(C_b = 1)\big)var(\hat{\tau}_b) - var(C_b)\cdot\mathbb{E}(\hat{\tau}_b)^2\right) \tag{1}$$

A formal derivation of Theorem TA-1.1 can be found in Appendix TA-2.2.1. Consider the case in which $P(C_b = 1)$ is equal to 1 or 0 for all blocks b (i.e., units within a given block have the same compliance status). In such a case, Equation 1 will simplify to the following:

$$\text{var}(\hat{\tau}_B^{IV}) - \text{var}(\hat{\tau}_B) \approx \frac{1}{(np)^2} \sum_{b=1}^{B} \big(1 - C_b\big) \text{var}(\hat{\tau}_b) \geq 0$$

Under this setting, because the blocked difference-in-means estimator directly drops blocks with no compliers in them, the blocked difference-in-means estimator will be a more efficient estimator than the blocked IV estimator. This is especially helpful when there are high rates of non-compliance, as the number of blocks we are able to drop increases with higher rates of non-compliance.

We can similarly show the efficiency gain from using blocked PSW over blocked IV.

Theorem TA-1.2 (Relative Reduction in Variance from Block PSW over Block IV) *The difference in variance between a blocked IV estimator and the block difference-in-means estimator is approximately:*

$$var(\hat{\tau}_B^{IV}) - var(\hat{\tau}_B^{PSW}) \approx \frac{1}{(np)^2} \sum_{b=1}^{B} \Bigg((1 - P(C_b = 1)) var_b(Y_i(1)) + (1 - \mathbb{E}_b(e_c(X_i)^2)) var_b(Y_i(0)) - \big(\mathbb{E}_b(Y(1))^2 var(C_b) - \mathbb{E}_b(Y_i(0))^2 var(e_c(X_i))\big) \Bigg) + \xi_1, \qquad (2)$$

where ξ_1 represents the difference between the unobservable, individual-level treatment effect.

We see that the efficiency gain from using PSW over IV in a blocking design is going to be dependent on (1) how much variation in compliance status (and principal score) there is within a specific block, and (2) how much compliance there is within each block. To illustrate the first point, consider the case in which there is no variation in compliance status and principal score within each block. Then we expect $\text{var}(C_b)$ and $\text{var}(e_c(X_i))$ to be equal to zero. As such, Equation 2 will simplify to:

$$\text{var}(\hat{\tau}_B^{IV}) - \text{var}(\hat{\tau}_B^{PSW}) \approx \frac{1}{(np)^2} \left(\sum_{b=1}^{B} (1 - P(C_b = 1)) \text{var}_b(Y_i(1)) + (1 - \mathbb{E}_b(e_c(X_i)^2)) \text{var}_b(Y_i(0)) \right) + \xi_1.$$

So long as ξ_1 is relatively small, then we expect this expression to be greater than zero.

For the second point, if the compliance rate is very low overall, because the principal score weighting estimator is weighting by probability of compliance, this reduces the overall magnitude of the variance for a block b (i.e., $P(C_b = 1)$ and $\mathbb{E}_b(e_c(X_i)^2)$ will both be small). This can help increase the relative efficiency gain over the block IV estimator.

In practice, if the blocks are formed such that we are able to minimize variation in compliance within a block (i.e., minimize $\text{var}(C_b)$ and $\text{var}(e_c(X_i))$), then the relative reduction in variance from using a block principal ignorability (i.e., block DIM or block PSW) over a block IV estimator will increase. With blocks of size 2, we expect, in practice, there to be minimal variation within each pair. In particular, if block principal ignorability holds, using the blocked PI will generally provide efficiency gains over the blocked IV estimator. We illustrate this in our simulations.

TA-1.2 Advantages to Precision from Blocking Design

In the following subsection, we examine the precision gains from using a blocking design. We show that using blocked difference-in-means can provide efficiency gains over a principal score weighted estimator with complete randomization. Furthermore, we extend the results from Miratrix, Sekhon, and Yu, 2013 and show blocking can be leveraged in the instrumental variables setting to help offset some of the precision loss under a weak instrument.

TA-1.2.1 Blocked Difference-in-Means versus Principal Score Weighting

We now compare the variance of the blocked difference-in-means estimator, identified under block principal ignorability, and the principal score weighting estimator with complete randomization, identified under principal ignorability. We show that the difference in precision between the blocked difference-in-means estimator and the principal score weighted estimator is primarily dependent upon the following: (1) the gains from the blocking design and (2) how much variation from complete randomization is reduced by weighting on the compliance scores. We formalize this in the following theorem.

Theorem TA-1.3 (Relative Efficiency of Blocked Difference-in-Means v. PSW) *Assuming that the principal scores are estimated using the same set of covariates, X_i, that are used to construct blocks, the approximate reduction in variance from using blocked difference-in-means over principal score weighting is written as:*

$$\begin{aligned} var(\hat{\tau}_{PSW}) - var(\hat{\tau}_B) &\approx \underbrace{\frac{1}{2n-1}\frac{1}{p^2}\left(var_B(\hat{\mu}_b(1) + \hat{\mu}_b(0)) - \frac{1}{n}\sum_{b=1}^{B}\frac{n-1}{n} var(\hat{\tau}_b^{PSW})\right)}_{(a)\ \text{Reduction in variance from blocking}} \\ &\quad \underbrace{- \frac{1}{(np)^2}\sum_{b=1}^{B}\left((\mathbb{E}_b(e_c(X_i)) - \mathbb{E}_b(e_c(X_i)^2))(var_b(Y_i(0)) + \mathbb{E}_b(Y_i(0))^2\right) + \zeta}_{(b)\ \text{Reduction in variance from weighting control units}}, \end{aligned} \tag{3}$$

where $\hat{\mu}_b(1)$ and $\hat{\mu}_b(0)$ represent the average complier outcome of the encouragement units and average weighted outcome of the control units (respectively), for a block b. $var_B(X_b) = \frac{1}{n}\sum_{b=1}^{n}(X_b - \frac{1}{n}\sum_{k=1}^{B} X_k)^2$, $var(\hat{\tau}^{(b)})$ represents the within-block variation of block b, and ζ is a function of the unobservable individual-level treatment effect terms (see Appendix TA-2.2.3 for more discussion).

A formal derivation of Theorem TA-1.3 is in Appendix TA-2.2.3. The first component in Theorem TA-1.3 (i.e., Equation 3(a)) represents the reduction in variance from for the principal score weighted estimator that comes from the blocking design. As is generally true with blocking designs, the magnitude of this term will depend on the total within-block and between-block variation, with gains when the between block variation in the outcomes dominates the within block variation. Previous studies have shown that even in cases when blocks are created at random, the variance of an estimator under blocking is typically no worse than the variance of an estimator under complete randomization (Miratrix, Sekhon, and Yu, 2013; Pashley and Miratrix, 2021a). We expect that, at worst, if blocks are formed at random, Equation 3(a) will be roughly equal to zero–i.e., there are no gains to efficiency from blocking. In practice, however, if researchers block on prognostic variables, this term will be positive.

The second component in Theorem TA-1.3 (i.e., Equation 3(b)) represents the reduction in variance from the principal score weighted estimator over the blocked difference-in-means estimator that comes from re-weighting the control units. The overall gains to efficiency from re-weighting is represented by the difference between the average principal score and the average principal score squared across a block b. If the principal scores within a given block are very extreme (i.e., close to 0 or 1), then we expect this term to be relatively small. However, if the principal scores within a block are less extreme (i.e., close to 0.5), then the difference will be larger.

In practice, because reduction in variance from weighting the control units only applies to the within-block variation of the control units, if the within-block variation is sufficiently minimized by the blocking design, it is difficult for the principal score weighted estimator to overcome the gains from blocking.

There are several key takeaways to highlight. First, while *identification* of the block difference-in-means estimator requires blocking on covariates that are related to compliance, *precision* is largely dependent on blocking on covariates related to variation in the outcome. Thus, researchers should be blocking on covariates related to compliance *and* the outcome. This formalizes previous findings in the principal stratification literature, which have shown that including covariates related to both compliance and outcome can be helpful when performing post-hoc adjustments to estimate CACE (Stuart and Jo, 2015; Porcher et al., 2016).

Secondly, this section focused on the precision gains that researchers can achieve from using a blocked experimental design with a blocked difference-in-means estimator to address confounding from noncompliance. However, a blocking design can be implemented in conjunction with principal score weighting, thereby leveraging the precision gains from the experimental design as well any precision gains that result from using model-based approaches. This includes adjusting for in-exact matches in the blocking design. In the context of Theorem TA-1.3, the variance reduction of a block-randomized design for the principal score weighted estimator is represented by Equation 3(a).

TA-1.2.2 Incorporating Blocking with Instrumental Variables

Even when researchers do not wish to invoke the block principal ignorability assumption, block-randomization can still improve precision of the IV estimator. Furthermore, when principal ignorability and the exclusion restriction are both viable assumptions for researchers to make, using the block difference-in-means estimator results in efficiency gains over the block instrumental variables estimator.

Extending the findings of Miratrix, Sekhon, and Yu (2013), we show that blocking can result in efficiency gains in the instrumental variables setting:

Theorem TA-1.4 (Relative Reduction in Variance from Blocking for an IV Estimator)

Let $\hat{\tau}_{IV}^{B}$ denote the IV estimator under block-randomization. The approximate relative reduction in variance from blocking the IV estimator is as follows:

$$var(\hat{\tau}_{IV}) - var(\hat{\tau}_{IV}^{B}) \approx \frac{1}{2n-1} \cdot \frac{1}{p^2} \left(var_B(\bar{Y}_b(1) + \bar{Y}_b(0)) - \frac{1}{n} \sum_{b=1}^{B} \frac{n-1}{n} var(\hat{\tau}_b) \right), \tag{4}$$

where, consistent with Theorem TA-1.3, $var_B(X_b) = \frac{1}{n} \sum_{b=1}^{B} \left(X_b - \frac{1}{n} \sum_{k=1}^{n} X_k \right)^2$, and $var(\hat{\tau}_b)$ rep-

resents the within-block variation of block b.

The results of Theorem TA-1.4 follow directly from findings in Pashley and Miratrix (2021a), Miratrix, Sekhon, and Yu (2013), and Imai (2008). In particular, the relative reduction in variance from blocking for the IV estimator depends minimizing within-block variation.

It is worth highlighting that the efficiency gain is a function of p^2; as such, holding all else equal, as noncompliance increases, so does the relative efficiency gain from using blocking in IV estimation. Previous literature has highlighted the deterioration of the IV estimator's performance under a weak instrument (i.e., high rates of non-compliance), specifically with respect to inflated standard errors and finite sample bias (Bound, Jaeger, and Baker, 1995). Blocking can help offset the instability of a weak instrument. Therefore, even if researchers are relying on the exclusion restriction, accounting for compliance (and/or outcome variation) during the design stage can help offset the precision loss associated with high rates of non-compliance.

TA-2 Proofs and Derivations

TA-2.1 Variance Derivations

Lemma TA-2.1 (Approximate Variance of Block Difference-in-Means) *Let there be n total blocks (i.e., each block comprises of 2 units). C_b is an indicator for whether or not block b is considered a "complier block". More specifically, $C_b = 1$ if the unit assigned to treatment in the b-th block is a complier. Then, the approximate variance of the blocked difference-in-means estimator, $\hat{\tau}_B$, is:*

$$var(\hat{\tau}_B) \approx \frac{1}{(np)^2} \sum_{b=1}^{B} \Big(\underbrace{var(\hat{\tau}_b) P(C_b = 1)}_{(a)} + \underbrace{\mathbb{E}(\hat{\tau}_b)^2 var(C_b)}_{(b)} \Big). \tag{5}$$

We can interpret term (a) as simply variance across each block b, weighted by the probability that the block is a complier block. Term (b) is the penalty for any variation in compliance within a given block. More specifically, $C_b = C_i \cdot T_i \mid i \in \mathcal{B}(b)$, where $\mathcal{B}(b)$ refers to the set of indices of the units in the b-th block. As such, if C_b varies based on who is assigned to treatment, this would imply that within a given block b, the complier status of the two units is different. As such, there is additional variance incurred from whether or not we are dropping the b-th block.

Proof: This result follows from an application of the Delta Method. First, we note that we may write $\hat{\tau}_B$ as a ratio estimator:

$$\hat{\tau}_B = \frac{1}{\sum_{b=1}^{B} C_b} \sum_{b=1}^{B} C_b \cdot \hat{\tau}_b$$

Because we are assuming blocks of size 2, $B = n$:

$$\begin{aligned} &= \frac{1}{\sum_{b=1}^{n} C_b} \sum_{b=1}^{n} C_b \cdot \hat{\tau}_b \\ &= \frac{\frac{1}{n}\sum_{b=1}^{n} C_b \cdot \hat{\tau}_b}{\frac{1}{n}\sum_{b=1}^{n} C_b} \end{aligned}$$

We now calculate a couple key quantities (i.e., the expectation, variance, and covariance of both the denominator and numerator terms):

1. $\mathbb{E}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\right) = p$

$$\begin{aligned}\mathbb{E}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\right) &= \frac{1}{n}\sum_{b=1}^{n}\mathbb{E}(C_b) \\ &= \frac{1}{n}\sum_{b=1}^{n}\mathbb{E}(C_i \cdot T_i \mid i \in \mathcal{B}(b))\end{aligned}$$

   By randomization of treatment assignment:

$$= \frac{1}{n}\sum_{b=1}^{n}\mathbb{E}(C_i \mid i \in \mathcal{B}(b))$$

   Because $\mathbb{E}(C_i) = \sum_{b=1}^{n}\mathbb{E}(C_i \mid i \in \mathcal{B}(b))P(i \in \mathcal{B}(b))$, where $P(i \in \mathcal{B}(b)) = 1/n$:

$$= p$$

2. $\mathbb{E}\left(\frac{1}{n}\sum_{b=1}^{n} C_b \cdot \hat{\tau}_b\right) = \mathbb{E}(C_b\hat{\tau}_b)$

$$\mathbb{E}\left(\frac{1}{n}\sum_{b=1}^{n} C_b \cdot \hat{\tau}_b\right) = \frac{1}{n}\sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b) = \mathbb{E}(C_b\hat{\tau}_b)$$

3. $\text{var}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\right) = \frac{1}{n}\text{var}(C_b)$

$$\text{var}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\right) = \frac{1}{n^2}\sum_{b=1}^{n}\text{var}(C_b) = \frac{1}{n}\text{var}(C_b)$$

4. $\text{cov}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\hat{\tau}_b, \frac{1}{n}\sum_{b=1}^{n} C_b\right)$

$$\begin{aligned}\text{cov}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\hat{\tau}_b, \frac{1}{n}\sum_{b=1}^{n} C_b\right) &= \frac{1}{n^2}\text{cov}\left(\sum_{b=1}^{n} C_b\hat{\tau}_b, \sum_{b=1}^{n} C_b\right) \\ &= \frac{1}{n^2}\sum_{b=1}^{n}\sum_{k=1}^{n}\text{cov}\left(C_b\hat{\tau}_b, C_k\right)\end{aligned}$$

For $b \neq k$, the covariance is zero:

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{b=1}^{n} \operatorname{cov}\left(C_b\hat{\tau}_b, C_b\right) \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \left(\mathbb{E}(C_b \cdot C_b\hat{\tau}_b) - \mathbb{E}(C_b\hat{\tau}_b)\mathbb{E}(C_b)\right) \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \left(\mathbb{E}(C_b\hat{\tau}_b) - \mathbb{E}(C_b\hat{\tau}_b)\mathbb{E}(C_b)\right) \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \mathbb{E}(C_b\hat{\tau}_b)\left(1 - \mathbb{E}(C_b)\right) \\
&= \frac{1}{n}\mathbb{E}(C_b\hat{\tau}_b)\left(1 - \mathbb{E}(C_b)\right)
\end{aligned}
$$

We denote $\hat{\mu}_{CB} = \frac{1}{n}\sum_{b=1}^{n} C_b\hat{\tau}_b$. Then, applying Delta Method:

$$
\begin{aligned}
\operatorname{var}(\hat{\tau}_B) &\approx \left(\frac{\mathbb{E}(C_b\hat{\tau}_b)}{p}\right)^2 \left(\frac{1}{\mathbb{E}(C_b\hat{\tau}_b)^2} \cdot \operatorname{var}\left(\hat{\mu}_{CB}\right) - \frac{2}{np}\frac{\mathbb{E}(C_b\hat{\tau}_b)(1-\mathbb{E}(C_b)}{\mathbb{E}(C_b\hat{\tau}_b)} + \frac{1}{np^2}\operatorname{var}(C_b)\right) \\
&= \frac{1}{p^2}\operatorname{var}\left(\hat{\mu}_{CB}\right) \underbrace{- \frac{2}{n} \cdot \frac{1}{p^3} \cdot \mathbb{E}(C_b\hat{\tau}_b)^2(1-\mathbb{E}(C_b)) + \frac{1}{n} \cdot \frac{1}{p^4}\operatorname{var}(C_b) \cdot \mathbb{E}(C_b\hat{\tau}_b)^2}_{(*)} \qquad (6) \\
&= \frac{1}{p^2}\operatorname{var}\left(\hat{\mu}_{CB}\right) + o(1) \qquad (7)
\end{aligned}
$$

Therefore, we see that as $n \to \infty$, the asymptotic variance of $\hat{\tau}_B$ is dominated by term 1. We will first derive the expression for $\operatorname{var}(\hat{\mu}_{CB})$. Then, we will show that approximating the variance of $\hat{\tau}_B$ by omitting the $o(1)$ component results in a *conservative* estimate of the variance.

Deriving var($\hat{\mu}_{CB}$):

$$
\begin{aligned}
\operatorname{var}(\hat{\mu}_{CB}) &= \operatorname{var}\left(\frac{1}{n}\sum_{b=1}^{n} C_b\hat{\tau}_b\right) \\
&= \frac{1}{n^2}\operatorname{var}\left(\sum_{b=1}^{n} C_b\hat{\tau}_b\right)
\end{aligned}
$$

Because we are using a blocking estimator, the variance between-block will be equal to zero, so we are left with the sum of within-block variation.

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{b=1}^{n} \operatorname{var}\left(C_b\hat{\tau}_b\right) \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \mathbb{E}\left(C_b^2\hat{\tau}_b^2\right) - \mathbb{E}(C_b\hat{\tau}_b)^2 \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \mathbb{E}\left(C_b\hat{\tau}_b^2\right) - \mathbb{E}(C_b\hat{\tau}_b)^2
\end{aligned}
$$

Under block principal ignorability, $C_b \perp\!\!\!\perp \hat{\tau}_b$:

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{b=1}^{n} \mathbb{E}(C_b)\mathbb{E}(\hat{\tau}_b^2) - \mathbb{E}(C_b)^2\mathbb{E}(\hat{\tau}_b)^2 \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \mathbb{E}(C_b)\mathbb{E}(\hat{\tau}_b^2) - \mathbb{E}(C_b)\mathbb{E}(\hat{\tau}_b)^2 + \mathbb{E}(C_b)\mathbb{E}(\hat{\tau}_b)^2 - \mathbb{E}(C_b)^2\mathbb{E}(\hat{\tau}_b)^2 \\
&= \frac{1}{n^2}\sum_{b=1}^{n} \mathbb{E}(C_b)\text{var}(C_b) + \mathbb{E}_b(C_b)(1-\mathbb{E}_b(C_b)) \cdot \mathbb{E}_b(\hat{\tau}_b)^2 \\
&= \frac{1}{n^2}\sum_{b=1}^{n} P(C_b = 1) \cdot \text{var}(\hat{\tau}_b) + P(C_b = 1)(1 - P(C_b = 1)) \cdot \mathbb{E}(\hat{\tau}_b)^2
\end{aligned}
$$

Now we will show that omitting the $o(1)$ component from Equation (6) results in a conservative estimate of the variance in finite-samples. To do so, we re-write the term in $(*)$:

$$
\begin{aligned}
&-\frac{2}{n} \cdot \frac{1}{p^3} \cdot \mathbb{E}(C_b\hat{\tau}_b)^2(1-\mathbb{E}(C_b)) + \frac{1}{n} \cdot \frac{1}{p^4}\text{var}(C_b) \cdot \mathbb{E}(C_b\hat{\tau}_b)^2 \\
&\qquad = -\frac{\mathbb{E}(C_b\hat{\tau}_b)^2}{np^3}\left(2(1-\mathbb{E}(C_b)) - \frac{\text{var}(C_b)}{p}\right)
\end{aligned}
$$

Making note of the fact that $\mathbb{E}(C_b) = p$ and $\text{var}(C_b) = p(1-p)$:

$$
\begin{aligned}
&= -\frac{\mathbb{E}(C_b\hat{\tau}_b)^2}{np^3}\left(2(1-p) - \frac{(1-p)\cdot p}{p}\right) \\
&= -\frac{\mathbb{E}(C_b\hat{\tau}_b)^2}{np^3}(1-p) \qquad (8)
\end{aligned}
$$

Because p is the probability of compliance, it is therefore bound between $[0, 1]$, and as such, Equation (8) is upper bounded by zero. □

Theorem TA-2.1 (Conservative Variance Estimator for Blocked Difference-in-Means)

For blocks of size 2*, the following is a conservative estimate of the variance of the blocked difference-in-means estimator:*

$$
\widehat{var}(\hat{\tau}_B) = \frac{1}{n_c(n_c-1)}\sum_{b=1}^{B} C_b(\hat{\tau}_b - \hat{\tau}_B)^2,
$$

where n_c is equal to the total number of revealed complier blocks, and $\hat{\tau}_b$ is equal to the difference-in-means estimated across block b.

Proof: The proof closely follows the proof from Imbens and Rubin (2015), in which the authors show that their proposed matched pairs variance estimator is conservative. We will apply a first-order Taylor series expansion to show that the expectation is approximately equal to a value that is greater than the true variance of $\hat{\tau}_B$. A similar proof will show that $\widehat{\text{var}}(\hat{\tau}_B)$ consistently estimates a value that is greater than the true variance of $\hat{\tau}_B$.

$$
\begin{aligned}
\sum_{b=1}^{B} C_b(\hat{\tau}_b - \hat{\tau}_B)^2 &= \sum_{b=1}^{n} C_b(\hat{\tau}_b - \hat{\tau}_B)^2 \\
&= \mathbb{E}\left(\sum_{b=1}^{n} C_b(\hat{\tau}_b - \hat{\tau}_B)^2\right) \\
&= \mathbb{E}\left(\sum_{b=1}^{n} C_b\left(\hat{\tau}_b - \frac{1}{n_c}\sum_{k=1}^{n} C_k\hat{\tau}_k\right)\right) \\
&= \mathbb{E}\left(\sum_{b=1}^{n} C_b\hat{\tau}_b^2 - \frac{2}{n_c}\sum_{b=1}^{n} C_b\hat{\tau}_b \cdot \sum_{k=1}^{n} C_k\hat{\tau}_k + C_b\left(\frac{1}{n_c}\sum_{k=1}^{n} C_k\hat{\tau}_k\right)^2\right) \\
&= \underbrace{\sum_{b=1}^{n} \mathbb{E}(C_b\hat{\tau}_b^2)}_{(1)} - \underbrace{\mathbb{E}\left(\frac{2}{n_c}\sum_{b=1}^{n}\sum_{k=1}^{n} C_bC_k\hat{\tau}_b\hat{\tau}_k\right)}_{(2)} + \underbrace{\mathbb{E}\left(\sum_{b=1}^{n} C_b\left(\frac{1}{n_c}\sum_{k=1}^{n} C_k\hat{\tau}_k\right)^2\right)}_{(3)}
\end{aligned}
$$

Using another Taylor series approximation and linearity of expectation allows us to write term (2) as follows:

$$
\begin{aligned}
\mathbb{E}\left(\frac{2}{n_c}\sum_{b=1}^{n}\sum_{k=1}^{n} C_bC_k\hat{\tau}_b\hat{\tau}_k\right) &\approx \frac{2}{np}\sum_{b=1}^{n}\sum_{k=1}^{n} \mathbb{E}(C_bC_k\hat{\tau}_b\hat{\tau}_k) \\
&= \frac{2}{np}\left(\sum_{b=1}^{n} \mathbb{E}(C_b^2\hat{\tau}_b^2) + \sum_{b=1}^{n}\sum_{k\neq b} \mathbb{E}(C_bC_k\hat{\tau}_b\hat{\tau}_k)\right) \\
&= \frac{2}{np}\left(\sum_{b=1}^{n} \mathbb{E}(C_b^2\hat{\tau}_b^2) + \sum_{b=1}^{n}\sum_{k\neq b} \mathbb{E}(C_b)\mathbb{E}(C_k)\mathbb{E}(\hat{\tau}_b)\mathbb{E}(\hat{\tau}_k)\right) \\
&= \frac{2}{np}\left(\sum_{b=1}^{n} \mathbb{E}(C_b^2\hat{\tau}_b^2) + \sum_{b=1}^{n}\sum_{k\neq b} \mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b \cdot \tau_k\right) \\
&= \frac{2}{np}\left(\sum_{b=1}^{n} \mathbb{E}(C_b\hat{\tau}_b^2) + \sum_{b=1}^{n}\sum_{k\neq b} \mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b \cdot \tau_k\right)
\end{aligned}
$$

Similarly for term (3):

$$
\begin{aligned}
\mathbb{E}\left(\sum_{b=1}^{n} C_b\left(\frac{1}{n_c}\sum_{k=1}^{n} C_k\hat{\tau}_k\right)^2\right) &\approx \frac{1}{np} \cdot \mathbb{E}\left(\sum_{b=1}^{n} C_b\right) \cdot \mathbb{E}\left(\left(\sum_{k=1}^{n} C_k\hat{\tau}_k\right)^2\right) \\
&= \frac{1}{(np)^2} \cdot np\mathbb{E}\left(\left(\sum_{k=1}^{n} C_k\hat{\tau}_k\right)^2\right) \\
&= \frac{1}{np}\mathbb{E}\left(\left(\sum_{k=1}^{n} C_k\hat{\tau}_k\right)\left(\sum_{j=1}^{n} C_j\hat{\tau}_j\right)\right) \\
&= \frac{1}{np}\mathbb{E}\left(\sum_{k=1}^{n}\sum_{j=1}^{n} C_kC_j\hat{\tau}_k\hat{\tau}_j\right)
\end{aligned}
$$

As such, combining the three terms:

$$\begin{aligned}
&\mathbb{E}\left(\sum_{b=1}^{n} C_b(\hat{\tau}_b - \hat{\tau}_B)^2\right) \\
&\approx \frac{np-1}{np}\sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b^2) - \frac{1}{np}\sum_{b=1}^{n}\sum_{k\neq b}\mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b\cdot\tau_k \\
&= \frac{np-1}{np}\sum_{b=1}^{n}\left(\mathbb{E}(C_b\hat{\tau}_b^2) - \mathbb{E}(C_b\hat{\tau}_b)^2\right) + \frac{np-1}{np}\sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 - \frac{1}{np}\sum_{b=1}^{n}\sum_{k\neq b}\mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b\cdot\tau_k \\
&= (np-1)\cdot np\cdot\frac{1}{p^2}\cdot\mathrm{var}(\hat{\mu}_{CB}) + \underbrace{\frac{np-1}{np}\sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 - \frac{1}{np}\sum_{b=1}^{n}\sum_{k\neq b}\mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b\tau_k}_{(**)}
\end{aligned}$$

Finally, we will show that the term denoted as $(**)$ is equal to $\sum_{b=1}^{n}\mathbb{E}(C_b)(\tau_b - \tau_{CACE})^2$.

$$\begin{aligned}
&\frac{np-1}{np}\sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 - \frac{1}{np}\sum_{b=1}^{n}\sum_{k\neq b}\mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b\tau_k \\
&\quad= \sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 - \frac{1}{np}\left(\sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 + \sum_{b=1}^{n}\sum_{k\neq b}\mathbb{E}(C_b)\mathbb{E}(C_k)\tau_b\tau_k\right) \\
&\quad= \sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 - \frac{1}{np}\left(\sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b\right)^2 \\
&\quad= \sum_{b=1}^{n}\mathbb{E}(C_b\hat{\tau}_b)^2 - np\cdot\tau_{CACE}^2 \\
&\quad= \sum_{b=1}^{n}\mathbb{E}(C_b)\mathbb{E}(\tau_b)^2 - \sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b\tau_{CACE} + \sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b\tau_{CACE} - np\cdot\tau_{CACE}^2
\end{aligned}$$

Noting that $\sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b\tau_{CACE} = np\cdot\tau_{CACE}^2$:

$$\begin{aligned}
&= \sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b^2 - 2\sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b\tau_{CACE} + np\cdot\tau_{CACE}^2 \\
&= \sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b^2 - 2\sum_{b=1}^{n}\mathbb{E}(C_b)\tau_b\tau_{CACE} + \sum_{b=1}^{n}\mathbb{E}(C_b)\tau_{CACE}^2 \\
&= \sum_{b=1}^{n}\mathbb{E}(C_b)(\tau_b - \tau_{CACE})^2
\end{aligned}$$

As such, we can write the numerator as:

$$\mathbb{E}\left(\sum_{b=1}^{n} C_b(\hat{\tau}_b - \hat{\tau}_B)^2\right) = (np-1)\cdot np\cdot\frac{1}{p^2}\mathrm{var}(\hat{\mu}_{CB}) + \sum_{b=1}^{n}\mathbb{E}(C_b)(\tau_b - \tau_{CACE})^2$$

Now taking the expectation of the denominator, we see that $\mathbb{E}(n_c(n_c-1)) = np(np-1)$.

Therefore,

$$\mathbb{E}(\widehat{\mathrm{var}}(\hat{\tau}_B)) \approx \frac{1}{p^2}\mathrm{var}(\hat{\mu}_{CB}) + \frac{1}{np(np-1)}\sum_{b=1}^{n}\mathbb{E}(C_b)(\tau_b - \tau_{CACE})^2 \geq \mathrm{var}(\hat{\tau}_B)$$

□

Lemma TA-2.2 (Approximate Variance of the Principal Score Weighted Estimator) *The approximate variance of the principal score weighted estimator is:*

$$\begin{aligned} var(\hat{\tau}_{PSW}) \approx & \frac{1}{np^2}\left(\frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{\mu}_{c,1})^2 + \frac{1}{2n-1}\sum_{i=1}^{N}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_{c,0})^2\right) \\ & - \frac{1}{(2n-1)p^2}\sum_{i=1}^{N}(C_iY_i(1)-\hat{e}_c(X_i)Y_i(0)-\tau_{PSW})^2 \end{aligned}$$

Proof: By construction:

$$\sum_{i=1}^{N} C_iT_i = \sum_{i=1}^{N}(1-T_i)\hat{e}_c(X_i)$$

As such, we can write the principal score weighted estimator as a ratio estimator:

$$\begin{aligned} \hat{\tau}_{PSW} &= \frac{\sum_{i=1}^{N} C_iT_iY_i}{\sum_{i=1}^{N} C_iT_i} - \frac{\sum_{i=1}^{N}\hat{e}_c(X_i)\cdot(1-T_i)Y_i}{\sum_{i=1}^{N}\hat{e}_c(X_i)(1-T_i)} \\ &= \frac{\sum_{i=1}^{N}\big(C_iT_iY_i - \hat{e}_c(X_i)(1-T_i)Y_i\big)}{\sum_{i=1}^{N} T_iC_i} \\ &= \frac{\frac{1}{N}\sum_{i=1}^{N}\big(C_iT_iY_i - \hat{e}_c(X_i)(1-T_i)Y_i\big)}{\frac{1}{N}\sum_{i=1}^{N} T_iC_i} \end{aligned}$$

We denote the numerator as $\hat{\mu}_{PSW}$. Applying Delta Method to the principal score weighted estimator, the approximate variance of the principal score weighted estimator is as follows:

$$\mathrm{var}(\hat{\tau}_{PSW}) \approx \frac{\mathrm{var}(\hat{\mu}_{PSW})}{p^2}, \tag{9}$$

where $\hat{\mu}_{PSW} := \frac{1}{N}\sum_{i=1}^{N}\big(C_iT_iY_i - \hat{e}_c(X_i)(1-T_i)Y_i\big)$. Following Imbens and Rubin (2015),

$$\begin{aligned} \mathrm{var}(\hat{\mu}_{PSW}) =& \mathrm{var}\left(\frac{1}{N}\sum_{i=1}^{N} C_iT_iY_i - \hat{e}_c(X_i)(1-T_i)Y_i\right) \\ =& \frac{1}{n}\left(\frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{\mu}_{c,1})^2 + \frac{1}{2n-1}\sum_{i=1}^{N}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_{c,0})^2\right) \\ & - \frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{e}_c(X_i)Y_i(0)-\tau_{PSW})^2 \end{aligned}$$

As such, the approximate variance of the principal score weighted estimator is:

$$\begin{aligned}\text{var}(\hat{\tau}_{PSW}) \approx & \frac{1}{np^2}\left(\frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{\mu}_{c,1})^2 + \frac{1}{2n-1}\sum_{i=1}^{N}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_{c,0})^2\right) \\ & - \frac{1}{(2n-1)p^2}\sum_{i=1}^{N}(C_iY_i(1)-\hat{e}_c(X_i)Y_i(0)-\tau_{PSW})^2\end{aligned}$$

□

TA-2.2 Efficiency Comparisons

TA-2.2.1 Proof of Theorem TA-1.1

The difference in variance between a blocked IV estimator and the block difference-in-means estimator is approximately:

$$\text{var}(\hat{\tau}_B^{IV}) - \text{var}(\hat{\tau}_B) = \frac{1}{(np)^2}\left(\sum_{b=1}^{B}\big(1-P(C_b=1)\big)\text{var}(\hat{\tau}_b) - \sum_{b=1}^{B}P(C_b=1)(1-P(C_b=1))\mathbb{E}(\hat{\tau}_b)^2\right)$$

Proof: The approximate variance of a blocked IV estimator is written as:

$$\text{var}(\hat{\tau}_B^{IV}) \approx \frac{1}{(np)^2}\sum_{b=1}^{B}\text{var}(\hat{\tau}_b), \tag{10}$$

which is the scaled sum of the within-block variances, and once again, because we have blocks of size 2, $B = n$. We can then apply the results from Lemma TA-2.1 and take the difference between Equation (10) and (5):

$$\begin{aligned}\text{var}(\hat{\tau}_B^{IV}) - \text{var}(\hat{\tau}_B) &= \frac{1}{(np)^2}\sum_{b=1}^{n}\text{var}(\hat{\tau}_b) - \left(\frac{1}{(np)^2}\sum_{b=1}^{n}\big(P(C_b=1)\text{var}(\hat{\tau}_b) - P(C_b=1)(1-P(C_b=1))\mathbb{E}(\hat{\tau}_b)^2\big)\right) \\ &= \frac{1}{(np)^2}\left(\sum_{b=1}^{n}\big(1-P(C_b=1)\big)\text{var}(\hat{\tau}_b) - \sum_{b=1}^{n}P(C_b=1)(1-P(C_b=1))\mathbb{E}(\hat{\tau}_b)^2\right) \\ &= \frac{1}{(np)^2}\left(\sum_{b=1}^{B}\big(1-P(C_b=1)\big)\text{var}(\hat{\tau}_b) - \sum_{b=1}^{B}P(C_b=1)(1-P(C_b=1))\mathbb{E}(\hat{\tau}_b)^2\right)\end{aligned}$$

□

TA-2.2.2 Proof of Theorem TA-1.2

The difference in variance between a blocked IV estimator and the block difference-in-means estimator is approximately:

$$\begin{aligned}\text{var}(\hat{\tau}_B^{IV}) - \text{var}(\hat{\tau}_B^{PSW}) \approx \frac{1}{(np)^2}\sum_{b=1}^{B}\Bigg(&(1-P(C_b=1))\text{var}_b(Y_i(1)) + (1-\mathbb{E}_b(e_c(X_i)^2))\text{var}_b(Y_i(0)) \\ &- \big(\mathbb{E}_b(Y(1))^2\text{var}(C_b) - \mathbb{E}_b(Y_i(0))^2\text{var}(e_c(X_i))\big)\Bigg) + \xi_1,\end{aligned}$$

where ξ_1 represents the difference between the unobservable, individual-level treatment effect.

Proof: To begin, we write the variance of the blocked PSW estimator:

$$\text{var}(\hat{\tau}_B^{PSW}) = \frac{1}{(np)^2} \sum_{b=1}^{B} \text{var}(\hat{\tau}_{PSW}^{(b)})$$

Thus, comparing it with the blocked IV estimator, we see that the primary differences in efficiency will be driven by the difference in the within-block variance from weighting the observations. To make these gains more clear, we can decompose the variance of the blocked PSW estimator as the sum of the treatment units that are revealed to be compliers, the sum of the re-weighted control units, and an unobservable individual-level treatment effect term.

$$\begin{aligned}
\text{var}(\hat{\tau}_B^{PSW}) =& \frac{1}{(np)^2} \sum_{b=1}^{B} \text{var}(\hat{\tau}_{PSW}^{(b)}) \\
=& \frac{1}{(np)^2} \sum_{b=1}^{B} (\text{var}_b(C_i Y_i(1)) + \text{var}_b(e_c(X_i) Y_i(0))) + \underbrace{\frac{1}{(np)^2} \sum_{b=1}^{B} \sum_{i=1}^{2} (\tau_{PSW,i}^{(b)} - \tau_{PSW}^{(b)})^2}_{\xi_2} \\
=& \frac{1}{(np)^2} \sum_{b=1}^{B} P(C_b = 1)\text{var}_b(Y_i(1)) + \mathbb{E}_b(Y_i(1))^2 \text{var}_b(C_i) + \\
& \frac{1}{(np)^2} \sum_{b=1}^{B} \mathbb{E}_b(e_c(X_i)^2)\text{var}_b(Y_i(0)) + \mathbb{E}_b(Y_i(0))^2 \text{var}_b(e_c(X_i)) + \xi_2
\end{aligned}$$

We can similarly write the blocked IV estimator as a function of these three components:

$$\begin{aligned}
\text{var}(\hat{\tau}_B^{IV}) &= \frac{1}{(np)^2} \sum_{b=1}^{B} \text{var}(\hat{\tau}_b) \\
&= \frac{1}{(np)^2} \left(\sum_{b=1}^{B} \text{var}_b(Y_i(1)) + \text{var}_b(Y_i(0)) \right) + \xi_3,
\end{aligned}$$

where $\xi_3 := \frac{1}{(np)^2} \sum_{b=1}^{B} (\tau_i - \tau_b)^2$.

Then, the results of the theorem follow directly by subtracting the two expressions, and defining $\xi_1 := \xi_3 - \xi_2$.

□

TA-2.2.3 Proof of Theorem TA-1.3

Assuming that the principal scores are estimated on the same set of covariates X_i as the covariates used to estimate blocks, the approximate reduction in variance from using blocked difference-in-means over principal score weighting is written as:

$$\begin{aligned}\text{var}(\hat{\tau}_{PSW}) - \text{var}(\hat{\tau}_B) &\approx \underbrace{\frac{1}{2n-1}\frac{1}{p^2}\left(\text{var}_B(\hat{\mu}_b(1)+\hat{\mu}_b(0)) - \frac{1}{n}\sum_{b=1}^{B}\frac{n-1}{n}\text{var}(\hat{\tau}_b^{PSW})\right)}_{(a)\text{ Reduction in variance from blocking}} \\ &\underbrace{-\frac{1}{(np)^2}\sum_{b=1}^{B}\left((\mathbb{E}_b(e_c(X_i)) - \mathbb{E}_b(e_c(X_i)^2))(\text{var}_b(Y_i(0)) + \mathbb{E}_b(Y_i(0))^2\right) + \zeta}_{(b)\text{ Reduction in variance from weighting control units}},\end{aligned} \tag{11}$$

where $\hat{\mu}_b(1)$ and $\hat{\mu}_b(0)$ represent the average complier outcome of the treatment units and average weighted outcome of the control units (respectively), for a block b. The subscript b denotes that the quantity is calculated with respect to the block b. $\text{var}_B(X_b) = \frac{1}{n}\sum_{b=1}^{B}(X_b - \frac{1}{n}\sum_{k=1}^{n}X_k)^2$, $\text{var}(\hat{\tau}^{(b)})$ represents the within-block variation of block b, and ζ is a function of the unobservable individual-level treatment effect terms.

Proof: To begin, we re-write the approximate variance of the block difference-in-means estimator in terms of the treatment and control outcomes:

$$\begin{aligned}\text{var}(\hat{\tau}_B) &\approx \frac{1}{np^2}\frac{1}{n}\left(\sum_{b=1}^{B}P(C_b=1)\text{var}_b(\hat{\tau}_b) + \mathbb{E}_b(\hat{\tau}_b)^2\text{var}(C_b)\right) \\ &= \underbrace{\frac{1}{np^2}\frac{1}{n}\left(\sum_{b=1}^{n}P(C_b=1)\text{var}_b(Y_i(1)) + \mathbb{E}_b(Y_i(1))^2\text{var}(C_b)\right)}_{(1)} + \\ &\quad \underbrace{\frac{1}{np^2}\frac{1}{n}\left(\sum_{b=1}^{n}P(C_b=1)\text{var}_b(Y_i(0)) + \mathbb{E}_b(Y_i(0))^2\text{var}(C_b)\right)}_{(2)} \\ &\quad \underbrace{-\frac{1}{(np)^2}\left(\sum_{b=1}^{n}\left(\sum_{i=1}^{2}P(C_b=1)(\tau_i-\tau_b)^2 + 2\mathbb{E}_b(Y_i(1)Y_i(0))\text{var}(C_b)\right)\right)}_{(3)},\end{aligned} \tag{12}$$

where the subscript b denotes that the quantity is calculated with respect to the block b.

From Lemma TA-2.2, the approximate variance of the principal score weighted estimator is:

$$\begin{aligned}\text{var}(\hat{\tau}_{PSW}) &\approx \frac{1}{np^2}\left(\underbrace{\frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{\mu}_{c,1})^2}_{(1)} + \underbrace{\frac{1}{2n-1}\sum_{i=1}^{N}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_{c,0})^2}_{(2)}\right) \\ &\quad \underbrace{-\frac{1}{(2n-1)\cdot p^2}\sum_{i=1}^{N}(\hat{\tau}_i^{PSW} - \tau_{PSW})^2}_{(3)},\end{aligned} \tag{13}$$

where $\hat{\tau}_i^{PSW} := C_iY_i(1) - \hat{e}_c(X_i)Y_i(0)$. The remainder of the proof follows by comparing each of the components in Equation 12 and Equation 13. We will start with the treatment outcomes. First, note that we can re-write the variance of the treatment outcomes in the principal score weighted estimator (i.e., Equation 13(1)) as a function of the within and between block variance:

$$\begin{aligned}\frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{\mu}_{c,1})^2 &= \frac{1}{2n-1}\sum_{b=1}^{n}\sum_{i=1}^{2}(Y_i(1)C_i-\hat{\mu}_b(1)+\hat{\mu}_b(1)-\hat{\mu}_{c,1})^2\\ &= \frac{1}{2n-1}\sum_{b=1}^{n}\text{var}_b(C_iY_i(1))+\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(1)-\hat{\mu}_{c,1})^2\end{aligned}$$

The variance across block b can be further decomposed:

$$\begin{aligned}\text{var}_b(C_iY_i(1)) &\equiv \text{var}(C_iY_i(1) \mid i\in\mathcal{B}(b))\\ &= \mathbb{E}(C_i \mid i\in\mathcal{B}(b))\cdot\text{var}_b(Y_i(1))+\mathbb{E}(Y_i(1)\mid i\in\mathcal{B}(b))^2P(C_i=1\mid i\in\mathcal{B}(b))(1-P(C_i\mid i\in\mathcal{B}(b))\end{aligned}$$

Because $\mathbb{E}(C_i \mid i\in\mathcal{B}(b)) = P(C_b=1)$, we can rewrite the above:

$$= P(C_b=1)\text{var}_b(Y_i(1))+\mathbb{E}_b(Y_i(1))^2P(C_b=1)(1-P(C_b=1)).$$

As such, if we just compare the variance of the treatment leg with that of the treated units in the block difference-in-means estimator (i.e., Equation 12-(1) and Equation 13-(1)), we obtain the following:

$$\begin{aligned}&\frac{1}{2n-1}\sum_{i=1}^{N}(C_iY_i(1)-\hat{\mu}_{c,1})^2-\frac{1}{n}\sum_{b=1}^{n}P(C_b=1)\text{var}_b(Y_i(1))+\mathbb{E}_b(Y_i(1))^2\text{var}(C_b)\\ &\quad=\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(1)-\hat{\mu}_{c,1})^2\\ &\quad\quad-\frac{1}{n^2}\frac{n-1}{(2n-1)}\sum_{b=1}^{n}P(C_b=1)\text{var}_b(Y_i(1))+\mathbb{E}_b(Y_i(1))^2\text{var}(C_b)\\ &\quad\equiv\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(1)-\hat{\mu}_{c,1})^2-\frac{1}{n^2}\frac{n-1}{(2n-1)}\sum_{b=1}^{n}\text{var}(\hat{\mu}_b(1)) \qquad (14)\end{aligned}$$

Similarly, for the control units, we may re-write the variance of the control group for the principal score weighted estimator (i.e., Equation 13-(2)) as:

$$\begin{aligned}\frac{1}{2n-1}\sum_{i=1}^{N}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_{c,0})^2 &= \frac{1}{2n-1}\sum_{b=1}^{n}\sum_{i=1}^{2}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_b(0)+\hat{\mu}_b(0)-\hat{\mu}_{c,0})^2\\ &= \frac{1}{2n-1}\sum_{b=1}^{n}\text{var}_b(\hat{e}_c(X_i)Y_i(0))+\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,0})^2\end{aligned}$$

Decomposing the within block variance term:

$$\begin{aligned}\text{var}_b(\hat{e}_c(X_i)Y_i(0)) &\equiv \text{var}(\hat{e}_c(X_i)Y_i(0)\mid i\in\mathcal{B}(b))\\ &= \mathbb{E}_b(\hat{e}_c(X_i)^2)\cdot\text{var}_b(Y_i(0))+\mathbb{E}(Y_i(0)\mid i\in\mathcal{B}(b))^2\text{var}_b(\hat{e}_c(X_i))\end{aligned}$$

Because we are assuming that the blocks are being formed on the same set of covariates X_i that are included in the principal score weights, $\mathbb{E}(C_b) = \mathbb{E}(\hat{e}_c(X_i)\mid i\in\mathcal{B}(b)) \equiv \mathbb{E}_b(\hat{e}_c(X_i))$.

Therefore, comparing just the control leg of both estimators (i.e., Equation 13-(2) and Equation 12-(2)):

$$\frac{1}{2n-1}\sum_{i=1}^{N}(\hat{e}_c(X_i)Y_i(0)-\hat{\mu}_{c,0})^2-\frac{1}{n}\sum_{b=1}^{n}P(C_b=1)\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\text{var}(C_b)$$

$$=\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,1})^2+\frac{1}{2n-1}\cdot\frac{1}{n}\sum_{b=1}^{n}\mathbb{E}_b(\hat{e}_c(X_i)^2)\cdot\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\text{var}_b(\hat{e}_c(X_i))$$

$$-\frac{1}{n^2}\sum_{b=1}^{n}\big(\mathbb{E}(C_b)\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\cdot\text{var}_b(Y_i(0))\big)$$

$$=\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,0})^2+\frac{1}{(2n-1)n^2}\sum_{b=1}^{n}\big(n\mathbb{E}_b(\hat{e}_c(X_i)^2)-(2n-1)\mathbb{E}(C_b)\big)\text{var}_b(Y_i(0))+$$

$$\big(n\text{var}_b(\hat{e}_c(X_i))-(2n-1)\text{var}(C_b)\big)\mathbb{E}_b(Y_i(0))^2$$

Adding and subtracting a $(2n-1)\cdot\mathbb{E}_b(\hat{e}_c(X_i))$ term to the first line, and a $(2n-1)\text{var}_b(\hat{e}_c(X_i))$ to the second line:

$$=\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,1})^2-\frac{1}{n^2}\frac{n-1}{(2n-1)}\sum_{b=1}^{n}\mathbb{E}_b(e_c(X_i)^2)\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\text{var}_b(\hat{e}_c(X_i))$$

$$+\frac{1}{n^2}\frac{(2n-1)}{(2n-1)}\sum_{b=1}^{n}\big((\mathbb{E}_b(e_c(X_i)^2)-\mathbb{E}(C_b))\text{var}_b(Y_i(0))+(\text{var}(e_c(X_i)-\text{var}(C_b))\mathbb{E}_b(Y_i(0))^2\big)$$

$$=\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,1})^2-\frac{1}{n^2}\frac{n-1}{(2n-1)}\sum_{b=1}^{n}\mathbb{E}_b(e_c(X_i)^2)\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\text{var}_b(\hat{e}_c(X_i))$$

$$-\frac{1}{n^2}\sum_{b=1}^{n}\big((\mathbb{E}_b(e_c(X_i))-\mathbb{E}_b(e_c(X_i)^2))(\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\big)$$

$$=\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,1})^2-\frac{1}{n^2}\frac{n-1}{(2n-1)}\sum_{b=1}^{n}\text{var}(\hat{\mu}_b(0))$$

$$-\frac{1}{n^2}\sum_{b=1}^{n}\big((\mathbb{E}_b(e_c(X_i))-\mathbb{E}_b(e_c(X_i)^2))(\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\big) \tag{15}$$

Finally, we decompose the unobservable individual-level treatment effect term in Equation 13-(3):

$$\frac{1}{2np^2}\sum_{i=1}^{N}(\tau_{PSW,i}-\tau_{PSW})=\frac{1}{2np^2}\frac{1}{2n-1}\left(\sum_{b=1}^{n}\sum_{i=1}^{2}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2+2\sum_{b=1}^{n}(\tau_{PSW}^{(b)}-\tau_{PSW})^2\right),$$

where $\tau_b^{PSW}=\hat{\mu}_b(1)-\hat{\mu}_b(0)$. Comparing it to the unobservable individual-level treatment effect term in Equation 12-(3):

$$\frac{1}{(np)^2}\left(\sum_{b=1}^{n}\left(\sum_{i=1}^{2}P(C_b=1)(\tau_i-\tau_b)^2+2\mathbb{E}_b(Y_i(1)Y_i(0))\text{var}(C_b)\right)\right)-\frac{1}{2np^2}\sum_{i=1}^{N}(\tau_i^{PSW}-\tau_{PSW})^2$$
$$=\frac{1}{(np)^2}\left(\sum_{b=1}^{n}\left(\sum_{i=1}^{2}P(C_b=1)(\tau_i-\tau_b)^2+2\mathbb{E}_b(Y_i(1)Y_i(0))\text{var}(C_b)\right)\right)$$
$$-\frac{1}{2np^2}\frac{1}{2n-1}\left(\sum_{b=1}^{n}\sum_{i=1}^{2}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2+2\sum_{b=1}^{n}(\tau_{PSW}^{(b)}-\tau_{PSW})^2\right)$$
$$=\underbrace{\frac{1}{(np)^2}\left(\sum_{b=1}^{n}\left(\sum_{i=1}^{2}P(C_b=1)(\tau_i-\tau_b)^2+2\mathbb{E}_b(Y_i(1)Y_i(0))\text{var}(C_b)\right)\right)-\frac{1}{(np)^2}\sum_{b=1}^{n}\sum_{i=1}^{n}(\tau_{PSW,i}^{(b)}+\tau_{PSW}^{(b)})^2+}_{:=\zeta}$$
$$\frac{1}{(np)^2}\sum_{b=1}^{n}\sum_{i=1}^{n}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2-\frac{1}{2np^2}\frac{1}{2n-1}\left(\sum_{b=1}^{n}\sum_{i=1}^{2}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2+2\sum_{b=1}^{n}(\tau_{PSW}^{(b)}-\tau_{PSW})^2\right)$$
$$=\zeta+\frac{1}{(np)^2}\sum_{b=1}^{n}\sum_{i=1}^{n}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2-\frac{1}{2np^2}\frac{1}{2n-1}\left(\sum_{b=1}^{n}\sum_{i=1}^{2}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2+2\sum_{b=1}^{n}(\tau_{PSW}^{(b)}-\tau_{PSW})^2\right)$$
$$=\zeta+\frac{1}{np^2}\cdot\left(\frac{1}{n}-\frac{1}{2}\cdot\frac{1}{2n-1}\right)\sum_{b=1}^{n}\sum_{i=1}^{n}(\tau_{PSW,i}^{(b)}-\tau_{PSW}^{(b)})^2-\frac{1}{2np^2}\frac{2}{2n-1}\sum_{b=1}^{n}(\tau_{PSW}^{(b)}-\tau_{PSW})^2 \tag{16}$$

Thus, we can combine Equations 14, 15, and 16:

$$\text{var}(\hat{\tau}_{PSW})-\text{var}(\hat{\tau}_B)$$
$$\approx\frac{1}{np^2}\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(0)-\hat{\mu}_{c,0})^2+\frac{1}{np^2}\frac{1}{n}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\mu}_b(1)-\hat{\mu}_{c,1})^2-\frac{1}{2np^2}\frac{2}{2n-1}\sum_{b=1}^{n}(\hat{\tau}_b^{PSW}-\hat{\tau}_{PSW})^2$$
$$-\frac{1}{n^2}\frac{n-1}{(2n-1)}\left(\sum_{b=1}^{n}\text{var}(\hat{\mu}_b(0))+\text{var}(\hat{\mu}_b(1))\right)+\frac{1}{np^2}\cdot\left(\frac{1}{n}-\frac{1}{2}\cdot\frac{1}{2n-1}\right)\sum_{b=1}^{n}\sum_{i=1}^{n}(\hat{\tau}_{PSW,i}^{(b)}-\hat{\tau}_{PSW}^{(b)})^2$$
$$-\frac{1}{n^2}\sum_{b=1}^{n}\left((\mathbb{E}_b(e_c(X_i))-\mathbb{E}_b(e_c(X_i)^2))(\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\right)+\zeta$$
$$=\frac{1}{2n-1}\frac{1}{p^2}\left(\text{var}_B(\hat{\mu}_b(1)+\hat{\mu}_b(0))-\frac{1}{n}\sum_{b=1}^{n}\frac{n-1}{n}\text{var}(\hat{\tau}_{PSW}^{(b)})\right)$$
$$-\frac{1}{(np)^2}\cdot\frac{1}{n}\sum_{b=1}^{n}\left((\mathbb{E}_b(e_c(X_i))-\mathbb{E}_b(e_c(X_i)^2))(\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\right)+\zeta$$
$$=\frac{1}{2n-1}\frac{1}{p^2}\left(\text{var}_B(\hat{\mu}_b(1)+\hat{\mu}_b(0))-\frac{1}{n}\sum_{b=1}^{B}\frac{n-1}{n}\text{var}(\hat{\tau}_{PSW}^{(b)})\right)$$
$$-\frac{1}{(np)^2}\cdot\frac{1}{n}\sum_{b=1}^{B}\left((\mathbb{E}_b(e_c(X_i))-\mathbb{E}_b(e_c(X_i)^2))(\text{var}_b(Y_i(0))+\mathbb{E}_b(Y_i(0))^2\right)+\zeta,$$

which is the result of Theorem TA-1.3. □

References

Bound, John, David Jaeger, and Regina Baker (1995). "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak". In: *Journal of the American statistical association* 90.430, pp. 443–450.

Feller, Avi, Fabrizia Mealli, and Luke Miratrix (2017). "Principal Score Methods: Assumptions, Extensions, and Practical Considerations". In: *Journal of Educational and Behavioral Statistics* 42.6, pp. 726–758.

Imai, Kosuke (2008). "Variance identification and efficiency analysis in randomized experiments under the matched-pair design". In: *Statistics in Medicine* 27.24, pp. 4857–4873.

Imbens, Guido and Donald Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. ISBN: 9781316094396.

Miratrix, Luke, Jasjeet Sekhon, and Bin Yu (2013). "Adjusting Treatment Effect Estimates by Post-stratification in Randomized Experiments". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 75.2, pp. 369–396.

Nickerson, David (2005). "Scalable Protocols Offer Efficient Design for Field Experiments". In: *Political Analysis* 13.3, pp. 233–252.

Pashley, Nicole and Luke Miratrix (2021a). "Block What You Can, Except When You Shouldn't". In: *Journal of Educational and Behavioral Statistics*, p. 107699862110272. ISSN: 1935-1054. DOI: 10.3102/10769986211027240. URL: http://dx.doi.org/10.3102/10769986211027240.

— (2021b). "Insights on variance estimation for blocked and matched pairs designs". In: *Journal of Educational and Behavioral Statistics* 46.3, pp. 271–296.

Porcher, Raphaël et al. (2016). "Performance of Principal Scores to Estimate the Marginal Compliers Causal Effect of an Intervention". In: *Statistics in Medicine* 35.5, pp. 752–767.

Stuart, Elizabeth and Booil Jo (2015). "Assessing the Sensitivity of Methods for Estimating Principal Causal Effects". In: *Statistical Methods in Medical Research* 24.6, pp. 657–674.