STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:4857–4873
DOI: 10.1002/sim.3337

Variance identification and efficiency analysis in randomized experiments under the matched-pair design

Kosuke Imai*, †

Department of Politics, Princeton University, Princeton, NJ 08544, U.S.A.

SUMMARY

In his 1923 landmark article, Neyman introduced randomization-based inference to estimate average treatment effects from experiments under the completely randomized design. Under this framework, Neyman considered the statistical estimation of the sample average treatment effect and derived the variance of the standard estimator using the treatment assignment mechanism as the sole basis of inference. In this paper, I extend Neyman's analysis to randomized experiments under the matched-pair design where experimental units are paired based on their pre-treatment characteristics and the randomization of treatment is subsequently conducted within each matched pair. I study the variance identification for the standard estimator of average treatment effects and analyze the relative efficiency of the matched-pair design over the completely randomized design. I also show how to empirically evaluate the relative efficiency of the two designs using experimental data obtained under the matched-pair design. My randomization-based analysis differs from previous studies in that it avoids modeling and other assumptions as much as possible. Finally, the analytical results are illustrated with numerical and empirical examples. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: causal inference; average treatment effect; randomization inference; paired comparison

1. INTRODUCTION

Despite the sharp disagreements on some issues [1], Neyman and Fisher agreed with each other on the use of the randomized treatment assignment mechanism as the sole basis of statistical inference in the statistical analysis of randomized experiments. Fisher introduced the method of permutation inference where the sharp null hypothesis of zero *unit treatment effect* was postulated and a distribution-free hypothesis test was developed [2, 3]. Since then, this method has been extensively studied [4]. Similarly, in his 1923 landmark article, which was not translated into English until 1990,

---

*Correspondence to: Kosuke Imai, Department of Politics, Princeton University, Princeton, NJ 08544, U.S.A.
†E-mail: kimai@Princeton.Edu, http://imai.princeton.edu

Contract/grant sponsor: National Science Foundation; contract/grant number: SES-0550873
Contract/grant sponsor: Princeton University Committee on Research in the Humanities and Social Sciences

Received 20 July 2007
Accepted 18 April 2008

Neyman proposed randomization-based inference in experiments and considered the statistical estimation of the *sample average treatment effect* (SATE) under the completely randomized design. Under this setting, Neyman [5] derived the variance of the standard estimator using the treatment assignment mechanism as the only source of randomness.

In this paper, after briefly reviewing Neyman's original analysis (Section 2), I extend his analysis to randomized experiments under the matched-pair design where experimental units are paired based on their pre-treatment characteristics and the randomization of treatment is subsequently conducted within each matched pair (Section 3). I derive the variances of the standard estimator when the estimand is either the sample or *population average treatment effect* (PATE). I then use these analytical results to study the relative efficiency of the matched-pair design over the completely randomized design. Unlike previous studies on the topic (see Section 3.1), I avoid modeling and other assumptions and base my inference solely on the randomization of treatment and the random sampling of units. My analysis points out the important differences between the matched-pair design and the randomized-block design, which appear to be neglected in the literature. I also show that the randomization-based approach clarifies some of the important questions raised in the statistical literature and identifies a hidden and yet unrealistic assumption that is made for the efficiency analysis in a widely used textbook. Finally, Section 4 illustrates the results of this paper with numerical and empirical examples, and Section 5 concludes.

2. NEYMAN'S ANALYSIS OF THE COMPLETELY RANDOMIZED DESIGN

In this section, I briefly review Neyman's [5] analysis of randomized experiments under the completely randomized design, which motivates my analysis of the matched-pair design. Suppose that there exist $2n$ units. Consider a binary treatment variable T_i, which equals 1 if unit i receives the treatment and 0 otherwise. Under the completely randomized design, n units are randomly selected to receive the treatment while the remaining n units are assigned to the control group. Using the potential outcomes notation that is commonly used in the literature of causal inference [6], I use $Y_i(t)$ to represent the potential outcomes under the treatment ($t=1$) and control ($t=0$) for each i. Then, the observed outcome, Y_i, is equal to $Y_i = T_i Y_i(1) + (1-T_i)Y_i(0)$.

Neyman considers the SATE as the estimand:

$$\tau_c \equiv \frac{1}{2n}\sum_{i=1}^{2n}(Y_i(1)-Y_i(0))$$

and derives the statistical properties of the standard estimator under the completely randomized design. This estimator is unbiased and is defined by

$$\hat{\tau}_c \equiv \frac{1}{n}\sum_{i=1}^{2n}\{T_i Y_i - (1-T_i)Y_i\} \tag{1}$$

The estimator corresponds to the mean difference in the observed outcome variable between the two groups, and hence is often referred to as the 'difference-in-means' estimator.

Like Fisher [3], Neyman bases his analysis on the treatment assignment mechanism alone, which is under the control of the experimenter, while assuming that all potential outcomes are fixed (though possibly unknown) quantities. Under this randomization-based framework, the variance of the estimator is derived conditional on the set of all potential outcomes, which I denote

DOI: 10.1002/sim

by $\mathcal{O}_c \equiv \{Y_i(0), Y_i(1)\}_{i=1}^{2n}$. Thus, the variability of the estimator can be attributed solely to the randomization of the treatment. Neyman shows that under the completely randomized design, this randomization-based variance of $\hat{\tau}_c$ is given by

$$\mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c) = \frac{1}{n}\left\{\mathrm{var}(Y_i(1)) + \mathrm{var}(Y_i(0)) - \frac{\mathrm{var}(Y_i(1) - Y_i(0))}{2}\right\} \tag{2}$$

where $\mathrm{var}(\cdot)$ represents the sample variance. Note that within this randomization framework, $\mathrm{var}(Y_i(1))$ and $\mathrm{var}(Y_i(0))$ can be estimated without bias by $s_1 \equiv \sum_{i=1}^{2n}\{T_i Y_i - \sum_{i'=1}^{2n} T_{i'} Y_{i'}/n\}^2/(n-1)$ and $s_0 \equiv \sum_{i=1}^{2n}\{(1-T_i)Y_i - \sum_{i'=1}^{2n}(1-T_{i'})Y_{i'}/n\}^2/(n-1)$, respectively. Thus, equation (2) implies that the standard variance estimator, $\hat{\sigma}_c \equiv (s_1 + s_0)/n$, overestimates $\mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c)$ on average unless the constant additive treatment effect assumption, i.e. $Y_i(1) - Y_i(0) = a$ for some constant a, holds so that $\mathrm{var}(Y_i(1) - Y_i(0)) = 0$.

The expression given in equation (2) has additional important implications [7]. First, note that the variance in equation (2) can be rewritten as

$$\mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c) = \frac{1}{2n}\{\mathrm{var}(Y_i(1)) + \mathrm{var}(Y_i(0)) - 2\,\mathrm{cov}(Y_i(1), Y_i(0))\} \tag{3}$$

where $\mathrm{cov}(\cdot,\cdot)$ represents the sample covariance. Since $Y_i(1)$ and $Y_i(0)$ are never jointly observed, correlation between them is not identifiable from the observed data. Consequently, $\mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c)$ cannot be identified. After observing this result, Neyman [5, p. 471] states:

> For the time being, we will conclude that since it is impossible to calculate directly an estimate of r [the correlation between $Y_i(1)$ and $Y_i(0)$], it is necessary to take $r = 1$; the method of comparing varieties or fertilizers by way of comparing average yields from several parallel plots has to be considered inaccurate.

Equation (3) together with the covariance inequality implies the following bounds of $\mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c)$:

$$\frac{1}{2n}\left\{\sqrt{\mathrm{var}(Y_i(1))} - \sqrt{\mathrm{var}(Y_i(0))}\right\}^2 \leqslant \mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c) \leqslant \frac{1}{2n}\left\{\sqrt{\mathrm{var}(Y_i(1))} + \sqrt{\mathrm{var}(Y_i(0))}\right\}^2 \tag{4}$$

The variance equals the upper and lower bounds when the sample correlation between $Y_i(1)$ and $Y_i(0)$ is 1 and -1, respectively. Although the variance itself $\mathbb{V}(\hat{\tau}_c \mid \mathcal{O})$ is not identified, these bounds can be consistently estimated from the observed data. Note that the upper bound in equation (4) is informative because it is no greater than the expected value of the standard variance estimator, i.e. $\mathbb{E}(\hat{\sigma}_c \mid \mathcal{O}_c) = \{\mathrm{var}(Y_i(1)) + \mathrm{var}(Y_i(0))\}/n$. In contrast, the lower bound is unlikely to be informative as it equals zero when $\mathrm{var}(Y_i(1)) = \mathrm{var}(Y_i(0))$.

Second, the standard variance estimator estimates the variance of $\hat{\tau}_c$ without bias when the estimand is the PATE. The definition of the PATE is given by $\mathbb{E}(Y_i(1) - Y_i(0))$, where the expectation is defined with respect to a population of interest. Under Neyman's framework, the population analysis can be considered as inference based on the two-step procedure: simple random sampling of $2n$ units from a population of infinite size followed by the complete randomization of the treatment.

Formally, let $\mathbb{V}(\hat{\tau}_c)$ represent the population variance of $\hat{\tau}_c$ under the completely randomized design and notice the following relationship:

$$\mathbb{V}(\hat{\tau}_c) = \mathbb{E}\{\mathbb{V}(\hat{\tau}_c \mid \mathcal{O}_c)\} + \mathbb{V}\{\mathbb{E}(\hat{\tau}_c \mid \mathcal{O}_c)\} = \frac{1}{n}\{\mathbb{V}(Y_i(1)) + \mathbb{V}(Y_i(0))\} \tag{5}$$

where the second equality follows from $\mathbb{E}(\hat{\tau}_c \mid \mathscr{O}_c) = \tau_c$ and equation (2). The law of iterated expectations implies that s_1 and s_0 estimate $\mathbb{V}(Y_i(1))$ and $\mathbb{V}(Y_i(0))$ without bias, respectively. Thus, the standard variance estimator is unbiased for $\mathbb{V}(\hat{\tau}_c)$.

In sum, from Neyman's randomization perspective, the variance of the standard difference-in-means estimator is not identified when the SATE is the estimand. In this case, the standard variance estimator is likely to be biased upwards. However, the bounds on this variance are derived and can be consistently estimated from the observed data. When the PATE is the estimand, on the other hand, the standard variance estimator is unbiased.

3. RANDOMIZATION-BASED ANALYSIS OF THE MATCHED-PAIR DESIGN

3.1. Overview of analytical results and previous studies

In this section, I extend the randomization-based inference of Neyman [5] reviewed above to the analysis of randomized experiments under the matched-pair design. First, I derive the variance of the standard estimator from the randomization-based perspective and show that the standard variance estimator is likely to be biased upwards when the SATE is the estimand. Similar to the case of the completely randomized design, the standard variance estimator is shown to be unbiased for the true variance only when the within-pair average treatment effect is constant across different pairs. Moreover, the standard variance estimator is unbiased, when the estimand is the PATE rather than the SATE. To my knowledge, this paper presents the first randomization-based analysis of the matched-pair design without making modeling assumptions. For example, Kempthorne [8] conducts a randomization analysis of generalized randomized block design (which include the matched-pair design as a special case) but entertains an additivity assumption. Wilk [9] considers a similar analysis under weaker assumptions but for the matched-pair design block-treatment interactions are assumed to be zero.

Second, using these results, I study the statistical efficiency of the matched-pair design relative to the completely randomized design. In particular, I show a couple of ways in which the two designs can be compared and derive the conditions under which the matched-pair design yields more efficient estimates than the completely randomized design. There exist some studies on the related topics. For example, Cochran [10] studies the efficiency of pairing in the context of linear regression. Billewicz [11] conducts simulation studies and observes that some of the findings are at odds with Cochran's [10] results (see also Youkeles [12] who conducts a numerical analysis using the sign test). Chase [13] shows that random matching does not result in the loss of asymptotic power over the standard normal theory test for the two proportions. Miettinen [14] also presents an asymptotic power analysis under particular modeling assumptions. More recently, Freedman *et al.* [15] and Martin *et al.* [16] study the efficiency of the matched-pair design in the context of community intervention trials using a model similar to the one used in Cochran [10]. Here, I contribute to this literature by applying Neyman's randomization-based approach to conduct the efficiency analysis of the matched-pair design. Unlike previous studies, my analysis is nonparametric and does not involve asymptotic approximations.

3.2. Variance identification

Suppose that there exist $2n$ units and n matched-pairs are formed based on the observed pre-treatment characteristics. Under the matched-pair design, the binary treatment is randomly assigned to one of the two units within each pair. An indicator variable, Z_j, is randomized by the experimenter with equal probability $\Pr(Z_j)=\frac{1}{2}$ and determines which unit receives the treatment within the jth matched-pair, where $j=1,2,\ldots,n$. If $Z_j=1$, then the first unit in the jth pair receives the treatment while the second unit does not. In contrast, $Z_j=0$ means that the second unit within the jth pair receives the treatment, whereas the first unit is assigned to the control group. Thus, if we use T_{ij} to indicate the value of the binary treatment variable for the ith unit in the jth matched pair, $T_{1j}=Z_j$ and $T_{2j}=1-Z_j$ hold for all j. Under the matched-pair design, Z_j and Z_k are assumed to be randomized independently of each other for any $j\neq k$. Finally, let $Y_{ij}(t)$ represent the potential outcomes under the treatment $(t=1)$ and control $(t=0)$ conditions for the ith unit in the jth matched pair. Then, Y_{ij} can be used to represent the observed outcome variable for each unit and $Y_{ij}=T_{ij}Y_{ij}(1)+(1-T_{ij})Y_{ij}(0)$.

Under the matched-pair design, the SATE can be expressed as

$$\tau_m \equiv \frac{1}{2n}\sum_{j=1}^{n}\sum_{i=1}^{2}(Y_{ij}(1)-Y_{ij}(0)) \tag{6}$$

The standard estimator under the matched-pair design is the sample average of within-pair differences, which is essentially the same as the difference-in-means estimator under the completely randomized design, i.e. $\hat{\tau}_c$ in equation (1). The estimator is given by

$$\begin{aligned}\hat{\tau}_m &\equiv \frac{1}{n}\sum_{j=1}^{n}\{Z_j(Y_{1j}-Y_{2j})+(1-Z_j)(Y_{2j}-Y_{1j})\}\\ &= \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{2}\{T_{ij}Y_{ij}-(1-T_{ij})Y_{ij}\}\end{aligned} \tag{7}$$

It is well known that this estimator is unbiased for the SATE. To show this familiar fact from the randomization perspective, rewrite $\hat{\tau}_m$ using potential outcomes as $\hat{\tau}_m=(1/n)\sum_{j=1}^{n}\{Z_j(Y_{1j}(1)-Y_{2j}(0))+(1-Z_j)(Y_{2j}(1)-Y_{1j}(0))\}$. Then, since $\mathbb{E}(Z_j\mid\mathcal{O}_m)=\frac{1}{2}$ for all j, it is immediate that $\mathbb{E}(\hat{\tau}_m\mid\mathcal{O}_m)=\tau_m$, where $\mathcal{O}_m$ represents the set of all potential outcomes under the matched-pair design, i.e. $\mathcal{O}_m\equiv\{Y_{1j}(1),Y_{1j}(0),Y_{2j}(1),Y_{2j}(0)\}_{j=1}^{n}$.

I now consider the identification of the variance of the standard estimator, $\hat{\tau}_m$, under the matched-pair design from the randomization perspective. Since Z_j is assumed to be independent across matched pairs, the variance of $\hat{\tau}_m$ equals

$$\mathbb{V}(\hat{\tau}_m\mid\mathcal{O}_m)=\frac{1}{4n^2}\sum_{j=1}^{n}\{(Y_{1j}(1)-Y_{2j}(0))-(Y_{2j}(1)-Y_{1j}(0))\}^2 \tag{8}$$

whereas the standard variance estimator, $\hat{\sigma}_m$, is based on the sample variance of within-pair differences:

$$\hat{\sigma}_m\equiv\frac{1}{n(n-1)}\sum_{j=1}^{n}\{Z_j(Y_{1j}-Y_{2j})+(1-Z_j)(Y_{2j}-Y_{1j})-\hat{\tau}_m\}^2 \tag{9}$$

The following proposition extends the results of Neyman [5] summarized in Section 2 to the matched-pair design. In particular, the proposition shows that when the SATE is the estimand, the standard variance estimator tends to overestimate the variance on average. Moreover, the true variance is not identifiable from the observed data but its bounds can be derived.

Proposition 1 (*Identification of the variance*)
Suppose that the estimand is the SATE, τ_m defined in equation (6), and the estimator is given by $\hat{\tau}_m$ in equation (7). Then, in randomized experiments under the matched-pair design, the following results hold:

1. *Bias of the standard variance estimator*: The standard variance estimator, $\hat{\sigma}_m$, defined in equation (9), tends to overestimate the true variance on average, and its bias equals

$$\mathbb{E}(\hat{\sigma}_m \mid \mathcal{O}_m) - \mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m) = \frac{1}{n}\text{var}\left\{\frac{1}{2}\sum_{i=1}^{2}(Y_{ij}(1) - Y_{ij}(0))\right\}$$

   where var$\{\cdot\}$ represents the sample variance across pairs.

2. *Bounds of the variance*: While $\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)$ is not identifiable, its bounds are given by

$$\begin{aligned}\frac{1}{4n^2}&\left\{\left(\sum_{j=1}^{n} D_j(1)^2\right)^{1/2} - \left(\sum_{j=1}^{n} D_j(0)^2\right)^{1/2}\right\}^2\\ &\leqslant \mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)\\ &\leqslant \frac{1}{4n^2}\left\{\left(\sum_{j=1}^{n} D_j(1)^2\right)^{1/2} + \left(\sum_{j=1}^{n} D_j(0)^2\right)^{1/2}\right\}^2\end{aligned}$$

   where $D_j(1) = Y_{1j}(1) - Y_{2j}(0)$ and $D_j(0) = Y_{2j}(1) - Y_{1j}(0)$.

A proof is given in Appendix A.1. The variance equals the lower (upper) bound when $D_j(1) = aD_j(0)$ for all j and some constant $a \leqslant 0$ ($a \geqslant 0$). Proposition 1 has several implications. First, the magnitude of the bias decreases as the variance of the within-pair SATE across pairs decreases. When there is a considerable degree of heterogeneity (an important motivation for the matched-pair design), then the bias is likely to be large. Indeed, the standard variance estimator is unbiased if and only if the variance of the within-pair SATE is zero; for example, this occurs when the constant additive treatment effect assumption holds, i.e. $Y_{ij}(1) - Y_{ij}(0) = \alpha$ for all (i, j) and some constant α. This is similar to the case of the completely randomized design in Section 2, where the standard variance estimator is unbiased if and only if the constant additive treatment effect assumption holds. Unfortunately, such an assumption is unlikely to hold especially in social science experiments where the heterogeneity among experimental subjects tends to be large.

Second, as is the case for the completely randomized design, the true variance is not identifiable from the data. The expectation of the standard variance estimator is an upper bound but this bound is in general not sharp and can be improved. Indeed, a tighter upper bound is given in Proposition 1. Similar to the case of the completely randomized design, the lower bound may not be very informative as it equals zero when $\sum_{j=1}^{n} D_j(1)^2 = \sum_{j=1}^{n} D_j(0)^2$.

Finally, as it is the case under the completely randomized design, although the standard variance estimator is biased when the SATE is the estimand, it is an unbiased estimator of the variance when the PATE is the estimand. Here, as in Section 2, I prove this fact formally by assuming the two-step procedure: an experimenter obtains a simple random sample of n matched pairs from an infinite population, and then conducts the simple randomization of the treatment within each pair. I further assume that the order within each pair is randomized so that the population distribution of $(Y_{1j}(1), Y_{1j}(0))$ is identical to that of $(Y_{2j}(1), Y_{2j}(0))$. Then, under these assumptions, the population variance of $\hat{\tau}_m$ can be derived as follows:

$$\begin{aligned}\mathbb{V}(\hat{\tau}_m) &= \mathbb{E}\{\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)\} + \mathbb{V}\left(\frac{1}{2n}\sum_{j=1}^{n}\sum_{i=1}^{2}(Y_{ij}(1) - Y_{ij}(0))\right)\\ &= \mathbb{E}\{\mathbb{E}(\hat{\sigma}_m \mid \mathcal{O}_m)\} = \mathbb{E}(\hat{\sigma}_m) \end{aligned} \tag{10}$$

where the last line follows from Proposition 1 and the fact that the sample variance is an unbiased estimate of the population variance. Note that using equation (A1) in Appendix A.1 and the law of iterated expectations, $\mathbb{E}(\hat{\sigma}_m)$ can be written as

$$\begin{aligned}\mathbb{E}(\hat{\sigma}_m) &= \frac{1}{2n}\mathbb{E}\{(Y_{1j}(1) - Y_{2j}(0))^2 + (Y_{2j}(1) - Y_{1j}(0))^2\}\\ &\quad - \frac{1}{4n}\mathbb{E}\{(Y_{1j}(1) + Y_{2j}(1) - Y_{1j}(0) - Y_{2j}(0))(Y_{1k}(1) + Y_{2k}(1) - Y_{1k}(0) - Y_{2k}(0))\}\\ &= \frac{1}{n}[\mathbb{E}(Y_{ij}(1) - Y_{i'j}(0))^2 - \{\mathbb{E}(Y_{ij}(1) - Y_{i'j}(0))\}^2]\\ &= \frac{1}{n}\mathbb{V}(Y_{ij}(1) - Y_{i'j}(0)) \end{aligned} \tag{11}$$

where $i \neq i'$, $j \neq k$, and $\mathbb{E}(Y_{ij}(1) - Y_{i'j}(0)) = \mathbb{E}(Y_{1j}(1) - Y_{2j}(0)) = \mathbb{E}(Y_{2j}(1) - Y_{1j}(0))$ (this follows from the assumption that the order within each pair is randomized), and the second equality follows from the random sampling of pairs, e.g. $\mathbb{E}\{(Y_{1j}(1) - Y_{2j}(0))(Y_{2k}(1) - Y_{1k}(0))\} = \{\mathbb{E}(Y_{ij}(1) - Y_{i'j}(0))\}^2$.

3.3. *Efficiency analysis*

Using the variance expressions obtained above, I next evaluate the statistical efficiency of the matched-pair design relative to the completely randomized design. I compare the efficiency of the matched-pair design with that of the completely randomized design in three different ways. First, I consider the relative efficiency of the matched-pair design for the estimation of the SATE. Next, the comparison of variances is made for the estimation of the PATE assuming that the completely randomized design uses the simple random sampling of units, whereas the matched-pair design uses the simple random sampling of matched pairs. Finally, the comparison of population variances is made under the assumption that the simple random sampling of matched units is made for both designs. Unlike the existing approaches, Neyman's randomization-based analysis clarifies the distinctions among these three types of comparisons.

The following proposition presents the result of the efficiency analysis when the SATE is the estimand.

Proposition 2 (*Efficiency comparison for the SATE estimation*)
Suppose that the SATE is the estimand. Let $\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)$ represent the variance of the standard estimator, $\hat{\tau}_m$, under the matched-pair design. Let $\mathbb{V}(\hat{\tau}_c^* \mid \mathcal{O}_m)$ denote the variance of the

standard estimator, $\hat{\tau}_c^* \equiv (1/2n)\sum_{i=1}^n \sum_{j=1}^2 \{T_{ij}Y_{ij} - (1-T_{ij})Y_{ij}\}$, under the completely randomized design, where T_{ij} is completely randomized so that n units receive the treatment. Then, the relative efficiency of the matched-pair design over the completely randomized design is characterized by

$$\begin{aligned}&\mathbb{V}(\hat{\tau}_c^* \mid \mathcal{O}_m) - \mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)\\ &\quad = \frac{n-1}{n(2n-1)}\left\{\operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j}) - \frac{1}{4}(\overline{Y}_1 - \overline{Y}_2)^2 + \frac{1}{4n}\operatorname{var}(\widetilde{Y}_{1j} - \widetilde{Y}_{2j})\right\}\end{aligned}$$

where $\widetilde{Y}_{ij} = Y_{ij}(1) + Y_{ij}(0)$ and $\overline{Y}_i = \sum_{j=1}^n \widetilde{Y}_{ij}/n$.

A proof is given in Appendix A.2. The next proposition summarizes the results of the efficiency analysis when the PATE is the estimand. In particular, the comparisons are made under different assumptions about sampling procedures.

Proposition 3 (*Efficiency comparison for the PATE estimation*)
Suppose that the PATE is the estimand. Let $\mathbb{V}(\hat{\tau}_m)$ and $\mathbb{V}(\hat{\tau}_c^*)$ denote the variances of the standard estimator under the matched pair and completely randomized designs, respectively, when the simple random sampling of matched pairs is assumed. Let $\mathbb{V}(\hat{\tau}_c)$ represent the variances of the standard estimator under the completely randomized design when the simple random sampling of units is assumed. Then, the relative efficiency of the matched-pair design over the completely randomized design is characterized by

$$\mathbb{V}(\hat{\tau}_c^*) - \mathbb{V}(\hat{\tau}_m) = \frac{n-1}{n(2n-1)}\operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j})$$

$$\mathbb{V}(\hat{\tau}_c) - \mathbb{V}(\hat{\tau}_m) = \frac{2}{n}\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0))$$

where $\operatorname{cov}(\cdot,\cdot)$ is the population covariance, $\widetilde{Y}_{ij} = Y_{ij}(1) + Y_{ij}(0)$ and $i \neq i'$.

A proof is in Appendix A.3. These two propositions have several important implications. First, when the PATE is the estimand, under the simple random sampling of matched pairs, the matched-pair design yields more efficient estimates than the completely randomized design when $\widetilde{Y}_{1j}$ and $\widetilde{Y}_{2j}$ are positively correlated. Unfortunately, this correlation is not identifiable from the observed data since $Y_{ij}(1)$ and $Y_{ij}(0)$ are never jointly observed for the ith unit in the jth pair. When the SATE is the estimand, two additional terms determine the relative efficiency together with the correlation between $\widetilde{Y}_{1j}$ and $\widetilde{Y}_{2j}$. However, as n increases, $\operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j})$ dominates $(\overline{Y}_1 - \overline{Y}_2)/4$, which tends to 0 by the law or large numbers, as well as $\operatorname{var}(\widetilde{Y}_{1j} - \widetilde{Y}_{2j})/4n$.

Second, the matched-pair design under the simple random sampling of matched pairs is compared with the completely randomized design under the simple random sampling of units. In this case, the difference between the two variances, i.e. $\mathbb{V}(\hat{\tau}_c)$ and $\mathbb{V}(\hat{\tau}_m)$, equals $\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0))$, which is identifiable from the observed data under the simple random sampling of matched pairs. Indeed, one can estimate $\mathbb{V}(\hat{\tau}_c)$ from the matched sample without bias (see equation (11) and

Appendix A.4):

$$\mathbb{V}(\hat{\tau}_c) = \mathbb{E}(\hat{\sigma}_m) + \frac{2}{n}\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0)) = \frac{1}{n}\mathbb{E}(s_1^* + s_0^*)$$

where s_1^* and s_0^* be the sample variances of the observed outcomes for the treatment and control groups, i.e.

$$s_1^* \equiv \frac{1}{n-1}\sum_{j=1}^{n}\left[Z_j Y_{1j} + (1-Z_j)Y_{2j} - \frac{1}{n}\sum_{j'=1}^{n}\{Z_{j'}Y_{1j'} + (1-Z_{j'})Y_{2j'}\}\right]^2$$

$$s_0^* \equiv \frac{1}{n-1}\sum_{j=1}^{n}\left[(1-Z_j)Y_{1j} + Z_j Y_{2j} - \frac{1}{n}\sum_{j'=1}^{n}\{(1-Z_{j'})Y_{1j'} + Z_{j'}Y_{2j'}\}\right]^2$$

Thus, depending on the sampling procedures under which the comparison is made, the efficiency analysis leads to different conclusions. In both cases, however, random matching neither reduces nor increases the variance relative to the completely randomized design. This result is consistent with the asymptotic results obtained by Chase [13].

Third, it is of interest to compare the above efficiency analysis with that of Imai *et al.* [17] who examine the relative efficiency of the stratified design over the completely randomized design. The stratified design refers to the procedure where the complete randomization of the treatment is conducted within a group of more than two observations. Imai *et al.* [17] show that when the estimand is the PATE, the variance of the standard difference-in-means estimator is no greater under the stratified design than under the completely randomized design. They also prove that when the estimand is the SATE, the same result holds asymptotically. Thus, Propositions 2 and 3 show that there is an important difference between these two seemingly closely related designs. While the adaptation of the stratified design over the completely randomized design can never hurt researchers in terms of statistical efficiency, the relative efficiency of the matched-pair design depends on whether matching induces positive or negative correlations regarding potential outcomes within each pair. This distinction appears to be neglected in the literature where the efficiency gain of the two designs are often discussed altogether (see Section 1 of Greevy *et al.* [18]).

Finally, the comparison between $\mathbb{V}(\hat{\tau}_c^*)$ and $\mathbb{V}(\hat{\tau}_m)$ may be of interest to researchers who are analyzing a pilot study in order to determine which design to employ in the main study. Snedecor and Cochran [19] conduct such a comparison and provide a formula (p. 100), which they argue estimates $\mathbb{V}(\hat{\tau}_c^*)$ without bias from the observed matched sample:

$$\hat{\sigma}_c^* \equiv \frac{2(n-1)}{n(2n-1)}(s_1^* + s_0^*) + \frac{\hat{\sigma}_m}{2n-1} \tag{12}$$

However, Proposition 3 implies that while $\mathbb{V}(\hat{\tau}_c)$ can be identified from the observed matched sample, $\mathbb{V}(\hat{\tau}_c^*)$ is not identifiable. Thus, $\hat{\sigma}_c^*$ is generally a biased estimate of $\mathbb{V}(\hat{\tau}_c^*)$. The following proposition derives the expressions for the bias of $\hat{\sigma}_c^*$ and the bounds on $\mathbb{V}(\hat{\tau}_c^*)$.

Proposition 4 (*Identification of* $\mathbb{V}(\hat{\tau}_c^*)$)
Suppose that the PATE is the estimand. Let $\mathbb{V}(\hat{\tau}_c^*)$ be the variance of the standard estimator, $\hat{\tau}_c^* \equiv (1/2n)\sum_{i=1}^{n}\sum_{j=1}^{2}\{T_{ij}Y_{ij} - (1-T_{ij})Y_{ij}\}$, under the completely randomized design when the

simple random sampling of matched pairs is assumed.

1. *Bias of* $\hat{\sigma}_c^*$: The bias of $\hat{\sigma}_c^*$ is given by

$$\begin{aligned}&\mathbb{V}(\hat{\tau}_c^*)-\mathbb{E}(\hat{\sigma}_c^*)\\&\quad=\frac{n-1}{n(2n-1)}\{\mathrm{cov}(Y_{1j}(1),Y_{2j}(1))+\mathrm{cov}(Y_{1j}(0),Y_{2j}(0))-2\,\mathrm{cov}(Y_{ij}(1),Y_{i'j}(0))\}\end{aligned}$$

2. *Bounds of* $\mathbb{V}(\hat{\tau}_c^*)$: $\mathbb{V}(\hat{\tau}_c^*)$ is not identifiable but its bounds are given by

$$\frac{n}{2n-1}\mathbb{E}(\hat{\sigma}_m)\leqslant\mathbb{V}(\hat{\tau}_c^*)\leqslant\frac{3n-2}{2n-1}\mathbb{E}(\hat{\sigma}_m)+\frac{4(n-1)}{n(2n-1)}\mathrm{cov}(Y_{ij}(1),Y_{i'j}(0))$$

A proof is given in Appendix A.4. The proposition shows that the bias of $\hat{\sigma}_c^*$ is not always zero, and that the direction of bias is unknown. One possible condition under which the bias is zero is given by $\mathrm{cov}(Y_{1j}(t),Y_{2j}(t))=\mathrm{cov}(Y_{ij}(1),Y_{i'j}(0))$ for $t=0,1$, but this need not hold in general. Although Snedecor and Cochran [19] introduce $\hat{\sigma}_c^*$ under the assumption of equal variances, i.e. $\mathbb{V}(Y_{ij}(1))=\mathbb{V}(Y_{ij}(0))$, this additional assumption does not alter the conclusion that $\hat{\sigma}_c^*$ is generally biased.

Furthermore, both the upper and the lower bounds of $\mathbb{V}(\hat{\tau}_c^*)$ can be estimated without bias where the sample covariance between Y_{1j} and Y_{2j} is an unbiased estimator of its population counterpart, $\mathrm{cov}(Y_{ij}(1),Y_{i'j}(0))$. Moreover, the lower bound is unlikely to be informative. Thus, one possible assumption to entertain is to assume that the correlation between $Y_{1j}(t)$ and $Y_{2j}(t)$ is greater than or equal to 0 for $t=0,1$. The assumption may be reasonable if matching is based on pre-treatment covariates that are known to be good predictors of the outcome and hence a positive correlation is likely to result. Under this assumption, the lower bound becomes

$$\mathbb{E}(\hat{\sigma}_m)+\frac{2(n-1)}{n(2n-1)}\mathrm{cov}(Y_{ij}(1),Y_{i'j}(0)) \tag{13}$$

while the upper bound remains the same. This lower bound is typically much greater than the lower bound given in Proposition 4 (see also Section 4.2).

3.4. *Practice of 'breaking the matches'*

The results given in the previous section have some implications for the practice of 'breaking the matches' where researchers analyze the matched-pair data as if matching had not occurred. Prior studies have investigated the statistical properties of such practice largely via simulation studies based on parametric models [20, 21]. From Neyman's nonparametric randomization perspective that is advocated in this paper, such practice that ignores the key aspect of experimental design cannot yield valid inferences. In particular, the variance estimator used when breaking the matches will be biased, and hence the resulting confidence intervals and statistical tests will also be biased.

4. ILLUSTRATIONS

In this section, the analytical results derived in the previous section are illustrated with numerical and empirical examples.

DOI: 10.1002/sim

4.1. Numerical examples

I first illustrate the fact that the matched-pair design can reduce statistical efficiency when compared with the completely randomized experiment (Proposition 2) using a simple numerical example given in Table I. In this example, there are four observations, all of which have zero treatment effect. All the potential outcomes are assumed to be known so that the exact variance can be calculated for the standard estimator of the SATE.

For the completely randomized design, there exist a total of six permutations of the treatment assignment and the same number of estimates of the SATE, i.e. $\{-2,-1,0,0,1,2\}$. Since each of these values has the equal probability of being realized, the variance of the standard estimator equals $\frac{5}{3}$. Next, consider a matched-pair design where unit a is paired with d and unit b is paired with c. This is clearly a 'poor' design since it eliminates the two permutations that yield the accurate estimate 0 for the SATE (the possible values the standard estimator can take are now $\{-2,-1,1,2\}$). Thus, under this design, the variance is $\frac{5}{2}$, which is larger than the one under the completely randomized design. It is easily confirmed that the difference in variance between the two designs equals the result based on the formula given in Proposition 2 by noting that $\operatorname{cov}(\widetilde{Y}_{1j},\widetilde{Y}_{1j})=-2$, $\overline{Y}_1-\overline{Y}_2=-4$, and $\operatorname{var}(\widetilde{Y}_{1j}-\widetilde{Y}_{1j})=8$. In contrast, consider the matching of similar units where unit a is paired with b and unit c is paired with d. In this case, the permutations that yield the least accurate estimate, i.e. -2 and 2, are eliminated. Then, the variance under this matched-pair design is $\frac{2}{3}$, which is smaller than that under the completely randomized design.

The second numerical example illustrates the efficiency results for the estimation of the PATE. Consider the following population model and let the sample size be 8, i.e. $n=4$:

$$\begin{pmatrix} Y_{1j}(1) \\ Y_{1j}(0) \\ Y_{2j}(1) \\ Y_{2j}(0) \end{pmatrix} \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left[\begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & -0.25 & -0.25 & 0.5 \\ -0.25 & 1 & 0.5 & -0.25 \\ -0.25 & 0.5 & 1 & -0.25 \\ 0.5 & -0.25 & -0.25 & 1 \end{pmatrix}\right] \tag{14}$$

Table I. An illustrative numerical example: $Y_i(1)$ and $Y_i(0)$ represent the potential outcomes under the treatment and control conditions, respectively.

	Potential outcomes	
Units	$Y(1)$	$Y(0)$
a	1	1
b	2	2
c	3	3
d	4	4

Pairing unit a with b and unit c with d yields a more efficient estimate than the completely randomized design, whereas pairing unit a with d and unit b with c results in the loss of relative efficiency.

Then, using the results in Section 3.2, $\mathbb{V}(\hat{\tau}_m)=\frac{1}{4}$ and $\mathbb{V}(\hat{\tau}_c)=\frac{1}{2}$. The latter result can also be verified through a Monte Carlo simulation procedure where one first draws a large number of samples from the model in equation (14) and then obtains a simple random sample of size n from this larger pool of draws. As shown in Proposition 3, the difference between $\mathbb{V}(\hat{\tau}_m)$ and $\mathbb{V}(\hat{\tau}_c)$ equals the $2\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0))/n$. Moreover, since $\operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j})=\frac{1}{2}$ in this example, $\mathbb{V}(\hat{\tau}_c^*)$ can be calculated using the same proposition, and the resulting value is $\frac{17}{56}\approx 0.3$. Finally, as shown in Proposition 4, the estimator of $\mathbb{V}(\hat{\tau}_c^*)$ proposed by Snedecor and Cochran [19], i.e. $\hat{\sigma}_c^*$ in equation (12), is biased. In this example, the bias equals $-\frac{9}{56}\approx -0.16$ where $\mathbb{E}(\hat{\sigma}_c^*)=\frac{13}{28}\approx 0.46$.

4.2. An empirical example

I next illustrate the theoretical results derived in this paper with an empirical example. I use the data from a public health randomized field experiment where health clusters in Mexico were randomly assigned to the treatment and control groups. In the treatment clusters, residents are encouraged to sign up the universal health insurance program called *Segro Popular de Salud* (SPS), whereas in the control group no such encouragement was given. The experiment was designed using the matched-pair design where clusters were paired based on the pre-treatment characteristics and cluster sizes. For the details of the design, see King *et al.* [22]. Researchers are interested in the causal effects of the SPS program on household heath-care expenditure, utilization of medical services, and individual health outcomes. Here, we conduct a cluster level analysis and focus on the intention-to-treat effect on the proportion of households suffering catastrophic health expenditures, i.e. out-of-pocket health-care expenditures totaling more than 30 per cent of a household's annual disposable income. An individual-level analysis of the same variable appears in Imai *et al.* [23].

The matched-pair design was implemented with 50 pairs of health clusters. The ITT (Intention-to-Treat) effect is estimated as $\hat{\tau}_m=1.27$ percentage point increase with the standard error of 0.69 (i.e. $\hat{\sigma}_m=0.48$). Now, the population variance under the completely randomized design (with the simple random sampling of units), i.e. $\mathbb{V}(\hat{\tau}_c)$, is estimated to be 0.91, which is almost twice large as the estimated variance under the matched-pair design. In fact, the estimated within-pair correlation coefficient is as high as 0.48. Furthermore, under the simple random sampling of matched pairs, the variance bounds under the completely randomized design are estimated to be $[0.24, 1.15]$ without any assumption. However, if one entertains the assumption of positive correlation between $Y_{1j}(t)$ and $Y_{2j}(t)$ for $t=0,1$, which appears reasonable in this case given the fact that matching was based on pre-test scores, the lower bound is estimated to be 0.69, which is about 1.5 times higher than the estimated variance obtained under the matched-pair design (see equation (13)). Thus, I conclude that in this experiment, the matched-pair design was highly effective reducing the estimation variance relative to the completely randomized design.

5. CONCLUDING REMARKS

In this paper, I study the variance identification and relative efficiency of the matched-pair design in the statistical analysis of randomized experiments by applying and extending the randomization-based framework of Neyman [5]. By doing so, I derive the conditions under which the matched-pair design yields more efficient estimates of the average treatment effects than the completely randomized design. These analytical results are illustrated with numerical and empirical examples. Researchers may wish to apply the variance formulae (or the bounds formulae when the variance is not identified) of this paper to their pilot study in order to determine whether the matched-pair design should be used in the main experiment. Furthermore, this paper also demonstrates the advantages of the potential outcomes framework, which dates back to Neyman [5] and Fisher [3], and is now widely used in the literature of causal inference. My analysis clarifies important questions concerning the statistical efficiency of the matched-pair design in the experimental design literature and reveals an implicit and yet unrealistic assumption for the efficiency analysis in a classic textbook. Finally, the analytical strategies of this paper can be extended to cluster-randomized trials where the matched-pair design is frequently employed. Such an extension has been undertaken by Imai *et al.* [23].

APPENDIX A

A.1. Proof of Proposition 1

Let $D_j(1)=Y_{1j}(1)-Y_{2j}(0)$ and $D_j(0)=Y_{2j}(1)-Y_{1j}(0)$. Using the potential outcome notation and taking the expectation with respect to Z_j yield

$$\begin{aligned}
&n(n-1)\mathbb{E}(\hat{\sigma}_m \mid \mathcal{O}_m)\\
&=\frac{1}{2}\sum_{j=1}^{n}\sum_{z=0}^{1}D_j(z)^2-\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathbb{E}[\{Z_jD_j(1)+(1-Z_j)D_j(0)\}\{Z_kD_k(1)+(1-Z_k)D_k(0)\}]\\
&=\frac{1}{2n}\left\{(n-1)\sum_{j=1}^{n}\sum_{z=0}^{1}D_j(z)^2-\frac{1}{2}\sum_{j=1}^{n}\sum_{k\neq j}(D_j(0)+D_j(1))(D_k(0)+D_k(1))\right\} \qquad \text{(A1)}
\end{aligned}$$

where the second equality follows from the fact that $\mathbb{E}(Z_j)=\mathbb{E}(Z_j^2)=\frac{1}{2}$, $\mathbb{E}\{Z_j(1-Z_j)\}=0$, and $\mathbb{E}(Z_jZ_k)=\mathbb{E}\{Z_j(1-Z_k)\}=\frac{1}{4}$ for all $j\neq k$. Thus, we have

$$\begin{aligned}
&\mathbb{E}(\hat{\sigma}_m \mid \mathcal{O}_m)-\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)\\
&=\frac{1}{4n^2}\left\{\sum_{j=1}^{n}(D_j(1)+D_j(0))^2-\frac{1}{n-1}\sum_{j=1}^{n}\sum_{k\neq j}(D_j(0)+D_j(1))(D_k(0)+D_k(1))\right\}\\
&=\frac{n}{4n^2(n-1)}\left\{\sum_{j=1}^{n}(D_j(1)+D_j(0))^2-\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(D_j(0)+D_j(1))(D_k(0)+D_k(1))\right\}\\
&=\frac{1}{4n}\mathrm{var}(D_j(0)+D_j(1))
\end{aligned}$$

From this equality, the desired bias expression immediately follows.

The above result implies

$$\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m) = \mathbb{E}(\hat{\sigma}_m \mid \mathcal{O}_m) - \frac{1}{4n}\text{var}\left\{\sum_{i=1}^{2}(Y_{ij}(1) - Y_{ij}(0))\right\}$$

However, $\text{var}\{\sum_{i=1}^{2}(Y_{ij}(1) - Y_{ij}(0))\}$ is not identifiable since $Y_{ij}(1)$ and $Y_{ij}(0)$ are never jointly observed. To derive the bounds of $\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m)$, use equation (8) and write

$$\mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m) = \frac{1}{4n^2}\sum_{j=1}^{n}\{D_j(1)^2 + D_j(0)^2 - 2D_j(1)D_j(0)\}$$

where the mean of the product $D_j(1)D_j(0)$ is not identifiable since they are never jointly observed. Thus, applying the Cauchy–Schwartz inequality to $\sum_{j=1}^{n} D_j(1)D_j(0)$, the desired expressions for the lower and upper bounds follow.

A.2. Proof of Proposition 2

Using equation (2), write the variance under the completely randomized design as

$$\begin{aligned}
\mathbb{V}(\hat{\tau}_c^* \mid \mathcal{O}_m) &= \frac{1}{2n(2n-1)}\sum_{j=1}^{n}\sum_{i=1}^{2}\left(\widetilde{Y}_{ij} - \frac{1}{2n}\sum_{j'=1}^{n}\sum_{i'=1}^{2}\widetilde{Y}_{i'j'}\right)^2 \\
&= \frac{1}{2n(2n-1)}\left\{\sum_{j=1}^{n}\sum_{i=1}^{2}\widetilde{Y}_{ij}^2 - \frac{1}{2n}\sum_{j=1}^{n}\sum_{k=1}^{n}(\widetilde{Y}_{1j} + \widetilde{Y}_{2j})(\widetilde{Y}_{1k} + \widetilde{Y}_{2k})\right\} \\
&= \frac{1}{4n^2}\left[\sum_{j=1}^{n}\sum_{i=1}^{2}\widetilde{Y}_{ij}^2 - \frac{1}{2n-1}\left\{\sum_{j=1}^{n}\sum_{k\neq j}(\widetilde{Y}_{1j} + \widetilde{Y}_{2j})(\widetilde{Y}_{1k} + \widetilde{Y}_{2k}) + 2\sum_{j=1}^{n}\widetilde{Y}_{1j}\widetilde{Y}_{2j}\right\}\right]
\end{aligned}$$

Then, together with equation (8), we have

$$\begin{aligned}
&\mathbb{V}(\hat{\tau}_c^* \mid \mathcal{O}_m) - \mathbb{V}(\hat{\tau}_m \mid \mathcal{O}_m) \\
&\quad = \frac{1}{4n^2}\left[2\sum_{j=1}^{n}\widetilde{Y}_{1j}\widetilde{Y}_{2j} - \frac{1}{2n-1}\left\{\sum_{j=1}^{n}\sum_{k\neq j}(\widetilde{Y}_{1j} + \widetilde{Y}_{2j})(\widetilde{Y}_{1k} + \widetilde{Y}_{2k}) + 2\sum_{j=1}^{n}\widetilde{Y}_{1j}\widetilde{Y}_{2j}\right\}\right] \\
&\quad = \frac{n-1}{4n(2n-1)}\text{var}(\widetilde{Y}_{1j} + \widetilde{Y}_{2j}) \\
&\qquad + \frac{1}{4n^2(2n-1)}\left\{\sum_{j=1}^{n}\sum_{i=1}^{2}\widetilde{Y}_{ij}^2 - n\sum_{j=1}^{n}\left(\sum_{i=1}^{2}\widetilde{Y}_{ij}\right)^2 + 2(2n-1)\sum_{j=1}^{n}\widetilde{Y}_{1j}\widetilde{Y}_{2j}\right\} \\
&\quad = \frac{n-1}{4n(2n-1)}\left\{\text{var}(\widetilde{Y}_{1j} + \widetilde{Y}_{2j}) - \frac{1}{n}\sum_{j=1}^{n}(\widetilde{Y}_{1j} - \widetilde{Y}_{2j})^2\right\} \qquad \text{(A2)}
\end{aligned}$$

Using this result, the desired expression can be derived from equation (A2) as follows:

$$
\begin{aligned}
&\operatorname{var}(\widetilde{Y}_{1j}+\widetilde{Y}_{2j})-\frac{1}{n}\sum_{j=1}^{n}(\widetilde{Y}_{1j}-\widetilde{Y}_{2j})^2\\
&=2\operatorname{cov}(\widetilde{Y}_{1j},\widetilde{Y}_{2j})+\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i=1}^{2}\widetilde{Y}_{ij}^2-\left(\frac{1}{n-1}-\frac{1}{n}\right)\left(\sum_{i=1}^{2}\overline{Y}_i^2\right)+\frac{2}{n}\sum_{j=1}^{n}\widetilde{Y}_{1j}\widetilde{Y}_{2j}\\
&=\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i=1}^{2}\widetilde{Y}_{ij}^2-\frac{n}{n-1}(\overline{Y}_1-\overline{Y}_2)^2+\left(\frac{2}{n}+\frac{2}{n-1}\right)\sum_{j=1}^{n}\widetilde{Y}_{1j}\widetilde{Y}_{2j}\\
&\quad-\left(\frac{2n}{n-1}+\frac{2n}{n-1}\right)\overline{Y}_1\overline{Y}_2\\
&=4\operatorname{cov}(\widetilde{Y}_{1j},\widetilde{Y}_{2j})+\frac{1}{n(n-1)}\sum_{j=1}^{n}(\widetilde{Y}_{1j}-\widetilde{Y}_{2j})^2-\frac{n}{n-1}(\overline{Y}_1-\overline{Y}_2)^2\\
&=4\operatorname{cov}(\widetilde{Y}_{1j},\widetilde{Y}_{2j})-(\overline{Y}_1-\overline{Y}_2)^2+\frac{1}{n}\operatorname{var}(\widetilde{Y}_{1j}-\widetilde{Y}_{2j})
\end{aligned}
$$

This completes the proof.

A.3. Proof of Proposition 3

Equation (11) and the fact that the order within each matched pair is randomized imply $\mathbb{V}(\hat{\tau}_m)=\{\mathbb{V}(Y_{ij}(1))+\mathbb{V}(Y_{ij}(0))-2\operatorname{cov}(Y_{ij}(1),Y_{i'j}(0))\}/n$. Since $\mathbb{V}(Y_{ij}(t))=\mathbb{V}(Y_i(t))$ for $t=0,1$, the result follows from the expression in equation (2). To prove the second equality, I use the expression in equation (A2) to show that

$$
\begin{aligned}
\mathbb{V}(\hat{\tau}_c^*)-\mathbb{V}(\hat{\tau}_m)&=\mathbb{E}(\mathbb{V}(\hat{\tau}_c^*\,|\,\mathcal{O}_m)-\mathbb{V}(\hat{\tau}_m\,|\,\mathcal{O}_m))\\
&=\frac{n-1}{4n(2n-1)}\{\mathbb{V}(\widetilde{Y}_{1j}+\widetilde{Y}_{2j})-E(\widetilde{Y}_{1j}-\widetilde{Y}_{2j})^2\}\\
&=\frac{n-1}{4n(2n-1)}[-\{\mathbb{E}(\widetilde{Y}_{1j})\}^2-\{\mathbb{E}(\widetilde{Y}_{2j})\}^2+4\mathbb{E}(\widetilde{Y}_{1j}\widetilde{Y}_{2j})-2\mathbb{E}(\widetilde{Y}_{1j})\mathbb{E}(\widetilde{Y}_{2j})]\\
&=\frac{n-1}{n(2n-1)}\left[\operatorname{cov}(\widetilde{Y}_{1j},\widetilde{Y}_{2j})-\frac{1}{4}\{\mathbb{E}(\widetilde{Y}_{1j}-\widetilde{Y}_{2j})\}^2\right]\\
&=\frac{n-1}{n(2n-1)}\operatorname{cov}(\widetilde{Y}_{1j},\widetilde{Y}_{2j})
\end{aligned}
$$

where the last equality follows from the fact that the order of units within each pair is assumed to be randomized, i.e. the population distributions of $(Y_{1j}(1),Y_{1j}(0))$ and $(Y_{2j}(1),Y_{2j}(0))$ are identical.

A.4. Proof of Proposition 4

First, I derive the expression of $\mathbb{E}(s_1^*)$ by calculating $\mathbb{E}(s_1^* \mid \mathcal{O}_m)$ and then using the law of iterated expectation. Since $Z_j^2 = Z_j$, $(1-Z_j)^2 = 1-Z_j$, and $Z_j(1-Z_j)=0$, rewrite s_1^* as

$$\begin{aligned} s_1^* = {} & \frac{1}{n}\sum_{j=1}^{n}\{Z_j Y_{1j}(1)^2 + (1-Z_j)Y_{2j}(1)^2\} \\ & - \frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{k\neq j}\{Z_j Y_{1j}(1) + (1-Z_j)Y_{2j}(1)\}\{Z_k Y_{1k}(1) + (1-Z_k)Y_{2k}(1)\} \end{aligned}$$

Using the fact that $\mathbb{E}(Z_j) = \frac{1}{2}$ and $\mathbb{E}(Z_j Z_k) = \mathbb{E}\{Z_j(1-Z_k)\} = \frac{1}{4}$ for all $j \neq k$, I obtain

$$\begin{aligned} \mathbb{E}(s_1^* \mid \mathcal{O}_m) = {} & \frac{1}{2n}\sum_{j=1}^{n}(Y_{1j}(1)^2 + Y_{2j}(1)^2) \\ & - \frac{1}{4n(n-1)}\sum_{j=1}^{n}\sum_{k\neq j}(Y_{1j}(1) + Y_{2j}(1))(Y_{1k}(1) + Y_{2k}(1)) \end{aligned}$$

Thus, the independence across matched pairs and the law of iterated expectations give $\mathbb{E}(s_1^*) = \mathbb{E}(Y_{ij}(1)^2) - \{\mathbb{E}(Y_{ij}(1))\}^2 = \mathbb{V}(Y_{ij}(1))$. Similarly, we have $\mathbb{E}(s_0^*) = \mathbb{E}(Y_{ij}(0)^2) - \{\mathbb{E}(Y_{ij}(0))\}^2 = \mathbb{V}(Y_{ij}(0))$. This implies that

$$\begin{aligned} \mathbb{V}(\hat{\tau}_c^*) - \mathbb{E}(\hat{\sigma}_c^*) &= \frac{2(n-1)}{n(2n-1)}\left[n\mathbb{E}(\hat{\sigma}_m) + \frac{1}{2}\operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j}) - \sum_{t=0}^{1}\{\mathbb{E}(Y_{ij}(t)^2) - \{\mathbb{E}(Y_{ij}(t))\}^2\}\right] \\ &= \frac{2(n-1)}{n(2n-1)}\left\{\frac{1}{2}\operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j}) - 2\operatorname{cov}(Y_{1j}(1), Y_{i'j}(0))\right\} \\ &= \frac{n-1}{n(2n-1)}\{\operatorname{cov}(Y_{1j}(1), Y_{2j}(1)) + \operatorname{cov}(Y_{1j}(0), Y_{2j}(0)) - 2\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0))\} \end{aligned}$$

To derive the bounds on $\mathbb{V}(\hat{\tau}_c^*)$, the Cauchy–Schwarz inequality implies that

$$-\mathbb{V}(Y_{ij}(1) - Y_{i'j}(0)) \leqslant \operatorname{cov}(\widetilde{Y}_{1j}, \widetilde{Y}_{2j}) \leqslant \mathbb{V}(Y_{ij}(1) - Y_{i'j}(0)) + 4\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0))$$

where $i \neq i'$. This together with the result in Proposition 3 implies

$$\left(1 - \frac{n-1}{2n-1}\right)\mathbb{V}(\hat{\tau}_m) \leqslant \mathbb{V}(\hat{\tau}_c^*) \leqslant \left(1 + \frac{n-1}{2n-1}\right)\mathbb{V}(\hat{\tau}_m) + \frac{4(n-1)}{n(2n-1)}\operatorname{cov}(Y_{ij}(1), Y_{i'j}(0))$$

Then, the desired expressions for the bounds immediately follow.

ACKNOWLEDGEMENTS

I thank a referee for the helpful comments. I acknowledge the financial support from the National Science Foundation (SES-0550873) and Princeton University Committee on Research in the Humanities and Social Sciences.

REFERENCES

1. Neyman J, Iwaszkiewicz K, Kolodziejczyk S. Statistical problems in agricultural experimentation (with Discussion). *Supplement of Journal of the Royal Statistical Society* 1935; **2**:107–180.
2. Fisher RA. The arrangement of field experiments. *Journal of Ministry of Agriculture* 1926; **33**:503–513.
3. Fisher RA. *The Design of Experiments*. Oliver and Boyd: London, 1935.
4. Rosenbaum PR. *Observational Studies* (2nd edn). Springer: New York, 2002.
5. Neyman J. On the application of probability theory to agricultural experiments: essay on principles, Section 9 (translated in 1990). *Statistical Science* 1923; **5**:465–480.
6. Holland PW. Statistics and causal inference (with Discussion). *Journal of the American Statistical Association* 1986; **81**:945–960.
7. Rubin DB. Comments on 'On the application of probability theory to agricultural experiments. Essay on principles. Section 9' by J. Splawa-Neyman translated from the Polish and edited by D. M. Dabrowska and T. P. Speed. *Statistical Science* 1990; **5**:472–480.
8. Kempthorne O. The randomization theory of experimental inference. *Journal of the American Statistical Association* 1955; **50**:946–967.
9. Wilk MB. Randomization analysis of a generalized randomized block design. *Biometrika* 1955; **42**:70–79.
10. Cochran WG. Matching in analytical studies. *American Journal of Public Health* 1953; **43**:684–691.
11. Billewicz WZ. The efficiency of matched samples: an empirical investigation. *Biometrics* 1965; **21**:623–644.
12. Youkeles LH. Loss of power through ineffective pairing of observations in small two-treatment all-or-none experiments. *Biometrics* 1963; **19**:175–180.
13. Chase GR. On the efficiency of matched pairs in Bernoulli trials. *Biometrika* 1968; **55**:365–369.
14. Miettinen OS. The matched pairs design in the case of all-or-none responses. *Biometrics* 1968; **24**:339–352.
15. Freedman LS, Green SB, Byar DP. Assessing the gain in efficiency due to matching in a community intervention study. *Statistics in Medicine* 1990; **9**:943–952.
16. Martin DC, Diehr P, Perrin EB, Koepsell TD. The effect of matching on the power of randomized community intervention studies. *Statistics in Medicine* 1993; **12**:329–338.
17. Imai K, King G, Stuart EA. Misunderstandings among experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society*, *Series A* 2008; **171**:481–502.
18. Greevy R, Lu B, Silber JH, Rosenbaum P. Optimal multivariate matching before randomization. *Biostatistics* 2004; **5**:263–275.
19. Snedecor GW, Cochran WG. *Statistical Methods* (8th edn). Iowa State University Press: Ames, IA, 1989.
20. Lynn HS, McCulloch CE. When does it pay to break the matches for analysis of a matched-pair design? *Biometrics* 1992; **48**:397–490.
21. Diehr P, Martin DC, Koespell T, Cheadle A. Breaking the matches in a paired t-test for community interventions when the number of pairs is small. *Statistics in Medicine* 1995; **14**:1491–1504.
22. King G, Gakidou E, Ravishankar N, Moore RT, Lakin J, Vargas M, Téllez-Rojo MM, Ávila JEH, Ávila MH, Llamas HH. A 'politically robust' experimental design for public policy evaluation, with application to the Mexican universal health insurance program. *Journal of Policy Analysis and Management* 2007; **26**:479–506.
23. Imai K, King G, Nall C. The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation. *Technical Report*, Department of Politics, Princeton University, 2007.

DOI: 10.1002/sim