Check for updates

Received: 8 April 2023 | Revised: 28 August 2023 | Accepted: 17 October 2023

DOI: 10.1002/jae.3025

WILEY Journal of APPLIED ECONOMETRICS

RESEARCH ARTICLE

Revisiting the analysis of matched-pair and stratified experiments in the presence of attrition

Yuehao Bai[1] | **Meng Hsuan Hsieh**[2] | **Jizhou Liu**[3] | **Max Tabord-Meehan**[4]

[1]Department of Economics, University of Southern California, Los Angeles, California, USA

[2]Ross School of Business, University of Michigan, Ann Arbor, Michigan, USA

[3]Booth School of Business, University of Chicago, Chicago, Illinois, USA

[4]Department of Economics, University of Chicago, Chicago, Illinois, USA

Correspondence
Yuehao Bai, Department of Economics, University of Southern California, Los Angeles, CA, USA.
Email: yuehao.bai@usc.edu

Funding information
We thank Rachel Glennerster, Hongchang Guo, David McKenzie, Azeem Shaikh, Alex Torgovitsky, Ed Vytlacil, and three anonymous referees for helpful comments. We also thank Lorenzo Casaburi and Tristan Reed for helpful comments and for sharing their data for one of our empirical applications. The fourth author acknowledges support from NSF Grant SES-2149408.

Abstract

In this paper, we revisit some common recommendations regarding the analysis of matched-pair and stratified experimental designs in the presence of attrition. Our main objective is to clarify a number of well-known claims about the practice of dropping pairs with an attrited unit when analyzing matched-pair designs. Contradictory advice appears in the literature about whether or not dropping pairs is beneficial or harmful, and stratifying into larger groups has been recommended as a resolution to the issue. To address these claims, we derive the estimands obtained from the difference-in-means estimator in a matched-pair design both when the observations from pairs with an attrited unit are retained and when they are dropped. We find limited evidence to support the claims that dropping pairs helps recover the average treatment effect, but we find that it may potentially help in recovering a convex-weighted average of conditional average treatment effects. We report similar findings for stratified designs when studying the estimands obtained from a regression of outcomes on treatment with and without strata fixed effects.

KEYWORDS

attrition, fixed effects, matched pairs, randomized controlled trial, stratified randomization

1 | INTRODUCTION

In this paper, we revisit some common recommendations regarding the analysis of matched-pair and stratified experimental designs in the presence of attrition. Here, we define attrition to mean that we do not observe outcomes for some subset of the experimental units. This situation may arise, for instance, if subjects refuse to participate in the experiment's endline survey or if researchers lose track of subjects prior to observing their experimental outcomes.

Our main objective is to clarify a number of well-known claims about the practice of dropping pairs with an attrited unit in matched-pair designs. Specifically, when one unit in a pair is lost, several contradictory suggestions have been made in the literature about whether or not experimenters should drop the remaining unit in their analyses.[1] For instance,

[1]Appendix A.7 in the Supporting Information contains relevant excerpts from the referenced sources.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Journal of Applied Econometrics* published by John Wiley & Sons Ltd.

King et al. (2007) and Bruhn and McKenzie (2009) assert that a key advantage of matched-pair designs is that dropping pairs with an attrited unit may protect against attrition bias when attrition is a function of the matching variables. In contrast, Glennerster and Takavarasha (2013) claim that dropping pairs may *increase* attrition bias and point out that the widespread practice of including pair fixed effects in a regression of outcomes on treatment is equivalent to computing the difference-in-means estimator after dropping pairs. Accordingly, they go on to suggest that experimenters should instead stratify the units into larger groups if there is risk of attrition. Donner and Klar (2000) assert that dropping pairs with an attrited unit is a *requirement* in analyses of matched-pair designs with attrition and characterize this as a weakness of matched-pair designs. As a result, they also recommend stratifying units into larger groups.

To address these claims, we first derive the estimands obtained from the difference-in-means estimator in a matched-pair design both when the observations from pairs with an attrited unit are retained and when they are dropped. We find that the estimand produced when retaining the units is simply the difference in the mean outcomes conditional on not attriting. In contrast, the estimand produced when dropping the units is a complicated function of the mean outcomes and attrition probabilities conditional on the matching variables. Using this result, we show that dropping pairs does not recover the average treatment effect when attrition is a function of the matching variables and instead recovers a convex-weighted average[2] of conditional average treatment effects. Moreover, we argue that natural conditions under which this convex-weighted average further collapses to the average treatment effect are in fact stronger than the condition that attrition is independent of experimental outcomes. From these results, we conclude that although dropping pairs may potentially help in recovering a convex-weighted average of conditional average treatment effects, we find limited evidence to support the claims that dropping pairs in a matched-pair design helps protect against attrition bias more generally.

Next, to address the claims that the issues surrounding whether or not to drop pairs with an attrited unit can be resolved by instead stratifying the experiment into larger groups, we repeat the above exercise in the context of a stratified randomized experiment where the strata are made up of a large number of observations. To mirror the analysis carried out for matched pairs, we study the estimands obtained from a regression of outcomes on treatment with and without strata fixed effects. We find analogous results: The estimand produced when omitting strata fixed effects is once again the difference in the mean outcomes conditional on not attriting, and the estimand produced when including strata fixed effects is a function of the mean outcomes and attrition probabilities conditional on the strata labels with very similar properties to what was obtained for matched pairs. From these results, we conclude that we do not find compelling evidence to support the idea that stratifying into larger groups resolves the issues surrounding attrition that we explore in this paper.

Including pair fixed effects when conducting inference via linear regression is a widely adopted practice (see, for instance, the recommendations in Bruhn & McKenzie, 2009) and is numerically equivalent to dropping pairs with an attrited unit. As a consequence, inference considerations sometimes drive the discussion of whether or not to drop pairs (see, e.g., Chapter 4, footnote 32 in Glennerster & Takavarasha, 2013). However, in our view, this should not play a primary role when deciding whether or not to drop pairs for three reasons. First, as we show in this paper, including versus excluding pair fixed effects produces estimands with distinct interpretations in the presence of attrition. Second, as argued in Bai et al. (2022) and Bugni et al. (2018) (in settings without attrition), including pair/strata fixed effects is not a requirement for conducting valid inference on the ATE in matched-pair/stratified experiments, and there is no clear benefit obtained from doing so in general. Third, there are no formal results which justify the use of conventional robust standard errors in the presence of attrition (with or without fixed effects), and we conjecture that alternative inference procedures should be developed in this case (see Remark 3.1 for a preliminary discussion). For these reasons, in this paper, our primary focus is on studying the interpretation of the resulting estimands.

Finally, we explore the empirical relevance of our results using experimental data collected in Groh and McKenzie (2016) and data collected from a systematic survey of all papers published in the American Economic Review (AER) and American Economic Journal: Applied Economics (AEJ: Applied) from 2020 to 2022 which conduct matched-pair or stratified experiments in the presence of attrition. Using these datasets, we find that there can be noticeable differences between the point estimates obtained from dropping or retaining pairs with an attrited unit (or including/omitting stratum fixed effects), even when attrition is comparatively low. For instance, using the data in Groh and McKenzie (2016), we find an average absolute percentage difference of 13.82% in point estimates across a collection of outcomes even with an average attrition rate of only 1.4%.

[2]Here and throughout the paper, we define a convex-weighted average to be a weighted average whose coefficients are non-negative and sum to one.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Our paper is related to a large literature on the analysis of randomized experiments with attrition. Most of this literature focuses on developing methods to recover the average treatment effect, often by either modeling the missing data process (Heckman, 1979; Rubin, 2004), inverse probability weighting (Little & Rubin, 2019; Wooldridge, 2002), bounding (Behaghel et al., 2015; Horowitz & Manski, 2000; Lee, 2009), or testing for the presence of attrition bias (Ghanem et al., 2021). Instead, the focus of our paper is on studying the behavior of commonly used estimators in the analysis of matched-pair and stratified experiments. To our knowledge, the paper most similar to ours is Fukumoto (2022), who conducts finite population and super-population analyses of the bias and variance of the difference-in-means estimator in matched-pair designs with and without dropping pairs. However, his super-population analysis maintains a sampling framework where the observations are drawn together as pairs, whereas we consider a sampling framework where observations are drawn as individuals and then subsequently paired according to their covariates. As a consequence, his results and ours are not directly comparable (we note that every empirical application we consider in Section 4 describes a specific procedure by which they stratified their sample using available covariates and thus does not feature a sample constructed from pre-formed strata as modeled in Fukumoto, 2022). Moreover, Fukumoto (2022) exclusively focuses on the setting of matched-pair designs and thus does not derive results for stratified randomized experiments.

The rest of the paper is structured as follows. In Section 2, we describe our setup and introduce the main assumptions we consider on the attrition process. Section 3 presents the main results. In Section 4, we present an empirical illustration. Finally, we conclude in Section 5 with some recommendations for empirical practice.

2 | SETUP AND NOTATION

Let Y_i^* denote the realized outcome of interest for the ith unit in the absence of attrition, $D_i \in \{0,1\}$ denote treatment status for the ith unit, and X_i denote the observed, baseline covariates for the ith unit. Further denote by $Y_i(1)$ the potential outcome of the ith unit if treated and by $Y_i(0)$ the potential outcome if not treated. As usual, the realized outcome is related to the potential outcomes and treatment status by the relationship

$$Y_i^* = Y_i(1)D_i + Y_i(0)(1 - D_i). \tag{1}$$

We consider a framework that allows for the possibility that units collected in the baseline survey may drop out (attrit) after treatment is assigned. In particular, let $R_i \in \{0,1\}$ be an indicator where $R_i = 1$ indicates the ith unit is present in the endline survey (i.e., has *not* attrited) and $R_i = 0$ indicates otherwise. Let $R_i(1)$ denote the potential attrition decision of the ith unit if treated and $R_i(0)$ denote the potential attrition decision of the ith unit if not treated. As was the case for the realized outcome, the realized attrition decision is related to the potential attrition decisions and treatment status by the relationship

$$R_i = R_i(1)D_i + R_i(0)(1 - D_i). \tag{2}$$

With these definitions in hand, we define the observed outcome to be

$$Y_i = Y_i^* R_i = Y_i(1)R_i(1)D_i + Y_i(0)R_i(0)(1 - D_i). \tag{3}$$

We note that the observed outcome is undefined if individual i is not observed in the endline survey, and so we set it arbitrarily to zero in Equation (3).

We assume that we observe a sample $\{(Y_i, R_i, D_i, X_i) : 1 \le i \le n\}$, obtained from i.i.d. random variables $\{W_i : 1 \le i \le n\}$ where $W_i = (Y_i(1), Y_i(0), R_i(1), R_i(0), X_i)$. As a result, the distribution of the observed data is determined by (1)–(3), $\{W_i : 1 \le i \le n\}$, and the mechanism for determining treatment assignment (which we specify in Sections 3.1 and 3.2). We maintain the following assumption on $\{W_i : 1 \le i \le n\}$ throughout the entirety of the paper:

Assumption 2.1.

(a) $E[|Y_i(d)|] < \infty$ for $d \in \{0,1\}$.
(b) $E[R_i(d)] > 0$ for $d \in \{0,1\}$.

Assumption 2.1(a) imposes mild restrictions on the moments of the potential outcomes. Assumption 2.1(b) rules out situations where the probability of attrition is one for either treatment status.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Our parameter of interest is the average treatment effect, denoted as

$$\theta = E[Y_i(1) - Y_i(0)]. \tag{4}$$

Without further assumptions on the nature of attrition, θ is not point-identified from the observed data. As a consequence, in this paper, we first study the estimands produced by commonly used estimators in the analysis of matched-pair and stratified randomized experiments and then document if and when these estimands collapse to θ under well-known, albeit strong, assumptions on the attrition process; see Remark 3.2 for further discussion. The first assumption we consider is that attrition is independent of the potential outcomes:

Assumption 2.2.

$$(Y_i(1), Y_i(0)) \perp\!\!\!\perp (R_i(1), R_i(0)).$$

Under Assumption 2.2, the average treatment effect θ is point-identified in a classical randomized experiment by simply comparing the mean outcomes under treatment and control for the non-attritors (see, for instance, Gerber & Green, 2012). The next assumption we consider is that attrition is independent of potential outcomes conditional on some set of observable characteristics:

Assumption 2.3. For some set of observable characteristics C_i,

$$(Y_i(1), Y_i(0)) \perp\!\!\!\perp (R_i(1), R_i(0))|C_i.$$

Although Assumption 2.2 does not necessarily imply Assumption 2.3 or vice versa, it is often argued that Assumption 2.3 may be easier to defend in practice (Gerber & Green, 2012; Hirano et al., 2001; Little & Rubin, 2019; Moffit et al., 1999). Under Assumption 2.3, θ is point-identified in a classical randomized experiment by first identifying the average treatment effect conditional on each value $C = c$ and then averaging these conditional treatment effects across C. Note that Assumption 2.3 generalizes the assumption discussed in the introduction that attrition is a function of observable characteristics. The final assumption we consider is that attrition is independent of observable characteristics:

Assumption 2.4. For some set of observable characteristics C_i,

$$C_i \perp\!\!\!\perp (R_i(1), R_i(0)).$$

A useful observation for the discussion which follows is that although Assumptions 2.2 and 2.3 are not nested, Assumptions 2.3 and 2.4 do in fact imply Assumption 2.2. To see this, consider the following derivation:

$$\begin{aligned}
&P\{(Y_i(1), Y_i(0)) \in U_1, (R_i(1), R_i(0)) \in U_2\} \\
&= E\left[E[I\{(Y_i(1), Y_i(0)) \in U_1\} I\{(R_i(1), R_i(0)) \in U_2\}|C_i]\right] \\
&= E\left[E[I\{(Y_i(1), Y_i(0)) \in U_1|C_i] E[I\{(R_i(1), R_i(0)) \in U_2\}|C_i]\right] \\
&= E\left[E[I\{(Y_i(1), Y_i(0)) \in U_1|C_i] E[I\{(R_i(1), R_i(0)) \in U_2\}]\right] \\
&= E\left[I\{(Y_i(1), Y_i(0)) \in U_1\}\right] E[I\{(R_i(1), R_i(0)) \in U_2\}] \\
&= P\{(Y_i(1), Y_i(0)) \in U_1\} P\{(R_i(1), R_i(0)) \in U_2\},
\end{aligned}$$

where the first equality follows from the law of iterated expectations, the second equality from Assumption 2.3, the third from Assumption 2.4, and the fourth from the law of iterated expectations once again.

3 | MAIN RESULTS

3.1 | Matched-pair designs with attrition

In this section, we study the estimands produced by the difference-in-means estimator in a matched-pair design when the observations from pairs with an attrited unit are retained and when they are dropped. Before defining the estimators,

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

we provide a formal description of the treatment assignment mechanism. To simplify the exposition, we assume that n is even for the remainder of Section 3.1. For any random variable indexed by i, for example, D_i, we denote by $D^{(n)}$ the random vector $(D_1, D_2, \ldots, D_n)$. Let $\pi = \pi_n(X^{(n)})$ be a permutation of $\{1, \ldots, n\}$, potentially dependent on $X^{(n)}$. The $n/2$ matched pairs are then represented by the sets

$$\left\{ \{\pi(2j-1), \pi(2j)\} : 1 \le j \le \frac{n}{2} \right\}.$$

In other words, pairs are formed by arranging observations in the order $\{\pi(1), \pi(2), \ldots, \pi(n)\}$ according to the permutation π and then forming pairs from the adjacent units as $\{\pi(1), \pi(2)\}$, $\{\pi(3), \pi(4)\}$, and so on. Next, given such a π, we assume treatment status is assigned as follows:

Assumption 3.1. Treatment status is assigned so that

$$(Y^{(n)}(1), Y^{(n)}(0), R^{(n)}(1), R^{(n)}(0)) \perp\!\!\!\perp D^{(n)} | X^{(n)},$$

and conditional on $X^{(n)}$, $(D_{\pi(2j-1)}, D_{\pi(2j)})$, $1 \le j \le n/2$ are i.i.d. and each uniformly distributed over $\{(0,1),(1,0)\}$.

To summarize, the assignment mechanism first forms pairs of units (according to π) and then assigns both treatments exactly once in each pair at random. The first estimator we consider is the standard difference-in-means estimator computed on non-attritors:

$$\hat{\theta}_n = \frac{\sum_{1 \le i \le n} Y_i R_i D_i}{\sum_{1 \le i \le n} R_i D_i} - \frac{\sum_{1 \le i \le n} Y_i R_i (1 - D_i)}{\sum_{1 \le i \le n} R_i (1 - D_i)}. \tag{5}$$

Note that $\hat{\theta}_n$ may be obtained as the estimator of the coefficient on D_i in an ordinary least squares regression of Y_i on a constant and D_i, computed on the non-attritors. The second estimator we consider is the difference-in-means estimator computed by first dropping any observations belonging to a pair with an attritor:

$$\hat{\theta}_n^{\text{drop}} = \left(\sum_{1 \le j \le n/2} R_{\pi(2j-1)} R_{\pi(2j)} \right)^{-1} \times \left(\sum_{1 \le j \le n/2} R_{\pi(2j-1)} R_{\pi(2j)} (Y_{\pi(2j-1)} - Y_{\pi(2j)})(D_{\pi(2j-1)} - D_{\pi(2j)}) \right).$$

Note that $\hat{\theta}_n^{\text{drop}}$ corresponds to the estimator recommended in Bruhn and McKenzie (2009) and King et al. (2007). We emphasize that in the absence of attrition, $\hat{\theta}_n$ and $\hat{\theta}_n^{\text{drop}}$ are numerically equivalent.

As a consequence of the Frisch–Waugh–Lovell theorem, $\hat{\theta}_n^{\text{drop}}$ can equivalently be obtained as the ordinary least squares estimator of the coefficient on D_i in the linear regression of Y_i on D_i and pair fixed effects computed on the non-attritors (i.e., individuals with $R_i = 1$)[3]:

$$Y_i = \theta^{\text{drop}} D_i + \sum_{1 \le j \le n/2} \delta_j I\{i \in \{\pi(2j-1), \pi(2j)\}\} + \epsilon_i \quad \text{(for individuals with } R_i = 1\text{)}. \tag{6}$$

Similar regression specifications are extremely common in the analysis of matched-pair experiments. See, for example, Ashraf et al. (2006), Angrist and Lavy (2009), Crepon et al. (2015), Bruhn et al. (2016), and Fryer (2018).

We impose the following assumption in addition to Assumption 2.1:

Assumption 3.2.

(a) $E[R_i(d)|X_i = x]$ is Lipschitz in x for $d \in \{0,1\}$.
(b) $E[Y_i(d)R_i(d)|X_i = x]$ is Lipschitz in x for $d \in \{0,1\}$.

Assumptions 3.2(a)–(b) are smoothness requirements that ensure that units that are "close" in terms of their baseline covariates are also "close" in terms of their potential attrition indicators and potential outcomes on average. Similar smoothness requirements are also imposed in Bai et al. (2022) and Bai (2022).

Finally, we require that the matched-pair design is such that the units in each pair are "close" in terms of their baseline covariates in the following sense:

[3]See Appendix A.6 in the Supporting Information for a derivation of this fact.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Assumption 3.3. The pairs used in determining treatment status satisfy

$$\frac{1}{n}\sum_{1\le j\le n}\|X_{\pi(2j-1)}-X_{\pi(2j)}\|\xrightarrow{P}0.$$

See Bai et al. (2022) for sufficient conditions for Assumption 3.3. In particular, if $\dim(X_i)=1$, then Assumption 3.3 is satisfied if $E[|X_i|]<\infty$, and we construct pairs by simply ordering the units from smallest to largest according to X_i and then pairing adjacent units. For the case $\dim(X_i)>1$, Bai et al. (2022) provide sufficient conditions under which Assumption 3.3 is satisfied when using the popular R package `nbpMatching`. Using appropriate laws of large numbers developed in Bai et al. (2022), we now establish the following result:

Theorem 3.1. *Suppose the data satisfy Assumptions 2.1 and 3.2 and the treatment assignment mechanism satisfies Assumptions 3.1 and 3.3. Then, as* $n\to\infty$, $\hat{\theta}_n\xrightarrow{P}\theta^{obs}$, *where*

$$\theta^{obs}=\frac{E[R_i(1)Y_i(1)]}{E[R_i(1)]}-\frac{E[R_i(0)Y_i(0)]}{E[R_i(0)]}=E[Y_i(1)|R_i(1)=1]-E[Y_i(0)|R_i(0)=1],$$

and $\hat{\theta}_n^{drop}\xrightarrow{P}\theta^{drop}$, *where*

$$\theta^{drop}=E[\tau^{obs}(X_i)\rho(X_i)],$$

with

$$\tau^{obs}(x)=E[Y_i(1)|R_i(1)=1,X_i=x]-E[Y_i(0)|R_i(0)=1,X_i=x],\ \rho(x)=\frac{E[R_i(0)|X_i=x]E[R_i(1)|X_i=x]}{E[E[R_i(0)|X_i]E[R_i(1)|X_i]]}.$$

Theorem 3.1 shows that the estimand produced by the difference-in-means estimator, θ^{obs}, is simply the difference in the mean outcomes conditional on not attriting (under the additional assumption that $R_i(1)=R_i(0)$, this could be interpreted as the average treatment effect for units that do not attrit: see Remark 3.3 for details). It follows immediately that under Assumption 2.2, $\theta^{\text{obs}}=\theta$, and thus, under this assumption, we recover the average treatment effect.

On the other hand, the estimand produced by first dropping units belonging to a pair with an attritor, θ^{drop}, is a complicated function of the mean outcomes and attrition probabilities conditional on the matching variables. First, note that unlike θ^{obs}, θ^{drop} does not collapse to θ under Assumption 2.2. Moreover, θ^{drop} does not collapse to θ under Assumption 2.3 with $C_i=X_i$ either. Instead, under Assumption 2.3 with $C_i=X_i$, $\tau^{\text{obs}}(x)=\tau(x)$ where $\tau(x)=E[Y_i(1)-Y_i(0)|X_i=x]$, so that

$$\theta^{\text{drop}}=E[\tau(X_i)\rho(X_i)],$$

that is, θ^{drop} may be written as a convex-weighted average of the conditional average treatment effects $\tau(x)$. In some special cases, this convex-weighted average has a simple and transparent interpretation: Consider, for example, a setting where X_i is a binary variable, and suppose that attrition is such that units with $X_i=1$ always appear in the endline survey, so that $R_i(1)=R_i(0)=1$ if $X_i=1$, but units with $X_i=0$ appear only if they are treated, so that $R_i(1)=1$ and $R_i(0)=0$ if $X_i=0$. Then,

$$\rho(1)=\frac{1}{P\{X_i=1\}},$$

and $\rho(0)=0$. We thus have that in this case,

$$\theta^{\text{drop}}=E[Y_i(1)-Y_i(0)|X_i=1],$$

which is the average treatment effect for those units with $X_i=1$. In contrast, θ^{obs} does not lend itself to a straightforward causal interpretation in this example (however, in Remark 3.3, we provide a favorable interpretation of θ^{obs} under Assumption 2.3 and the additional assumption that $R_i(1)=R_i(0)$).

In general, straightforward algebra shows that $\rho(x)=1$ if and only if

$$E[R_i(0)|X_i=x]=\frac{E[E[R_i(0)|X_i]E[R_i(1)|X_i]]}{E[R_i(1)|X_i=x]}. \tag{7}$$

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In words, $\rho(x) = 1$ if and only if the conditional probability of attrition under treatment is inversely proportional to the conditional probability of attrition under control. A natural assumption which guarantees (7) for all x is Assumption 2.4 with $C_i = X_i$, so that attrition is independent of the matching variables X_i. Finally, we note that under Assumption 2.4 with $C_i = X_i$, it follows that $\theta^{\text{drop}} = \theta^{\text{obs}}$. As a result, $\theta^{\text{drop}} = \theta$ under Assumptions 2.2 and 2.4. We summarize the above discussion in the following corollary:

Corollary 3.1.

(a) Under Assumption 2.2, $\theta^{obs} = \theta$.
(b) Under Assumption 2.3 with $C_i = X_i$, $\theta^{drop} = E[\tau(X_i)\rho(X_i)]$.
(c) Under Assumption 2.4 with $C_i = X_i$, $\theta^{drop} = \theta^{obs}$.
(d) Under Assumption 2.4 with $C_i = X_i$ and either Assumption 2.2 or 2.3 with $C_i = X_i$, $\theta^{drop} = \theta$.

We conclude this section by noting that as explained in the derivation following the statement of Assumption 2.4, Assumptions 2.3 and 2.4 *imply* Assumption 2.2. In other words, we see that the sufficient conditions provided in Corollary 3.1 under which $\theta^{\text{drop}} = \theta$ are in fact *stronger* than the conditions required for $\theta^{\text{obs}} = \theta$. We thus find limited evidence to support the claims that dropping pairs in a matched-pair design helps in reducing attrition bias. However, we emphasize that dropping pairs may potentially help in recovering a convex-weighted average of conditional average treatment effects.

Remark 3.1. In Appendix A.3 in the Supporting Information, we develop the requisite distributional results to use $\hat{\theta}_n$ for inference about θ^{obs}. In contrast, the large sample distribution of $\hat{\theta}_n^{\text{drop}}$ seems non-trivial to characterize and may in fact feature an asymptotic bias in general. For this reason, we leave an in-depth study of the limiting distribution of $\hat{\theta}_n^{\text{drop}}$ to future work.

Remark 3.2. We note that in the absence of additional assumptions like Assumptions 2.2–2.4, we are not able to conclude that either θ^{obs} or θ^{drop} is less biased for θ relative to the other, and in fact, it is possible to construct data-generating processes where either estimand is closer to the true average treatment effect. We present a concrete construction of such a set of DGPs in Appendix A.4 in the Supporting Information.

Remark 3.3. From Theorem 3.1, we also observe that in the absence of additional assumptions like Assumptions 2.2–2.4, neither θ^{obs} nor θ^{drop} can be interpreted as a "treatment effect parameter" (in the sense that neither parameter can be interpreted as an average treatment effect for some subset of individuals or more generally as a weighted average of treatment effects). This is because the subgroup of units who attrit under treatment ($R_i(1) = 0$) may not correspond to the subgroup of units who attrit under control ($R_i(0) = 0$). Under the additional assumption that $R_i(1) = R_i(0)$, so that these subgroups coincide, we obtain

$$\theta^{\text{obs}} = E[Y_i(1) - Y_i(0) | R_i = 1],$$

and then, θ^{obs} could be understood as the average treatment effect for units who do not attrit. Imposing the same assumption for θ^{drop}, we obtain that

$$\theta^{\text{drop}} = \frac{E[(Y_i(1) - Y_i(0))\, R_i E[R_i|X_i]]}{E[R_i E[R_i|X_i]]},$$

and then, θ^{drop} could be understood as a "probability of attrition"-weighted average of individual-level treatment effects for the non-attritors. If we additionally impose Assumption 2.3, we alternatively obtain

$$\theta^{\text{obs}} = \frac{E[\tau(X_i)E[R_i|X_i]]}{P(R_i = 1)},\ \theta^{\text{drop}} = \frac{E[\tau(X_i)E[R_i|X_i]^2]}{E[E[R_i|X_i]^2]}.$$

In this case, *both* parameters can be interpreted as convex-weighted averages of conditional average treatment effects, with the main difference being that θ^{obs} is weighted using the conditional attrition rate $E[R_i|X_i]$, whereas θ^{drop} "doubles down" by weighting using the squared conditional attrition rate $E[R_i|X_i]^2$.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Remark 3.4. Regardless of whether or not a practitioner finds the interpretation of θ^{drop} more or less attractive than the interpretation of θ^{obs}, it is crucial to note that even in the absence of attrition, inferences produced using robust standard errors obtained from a regression with pair fixed effects are generally conservative, but in some cases may in fact be *invalid*, in the sense that the limiting rejection probability could be strictly larger than the nominal level. See Bai et al. (2022) and de Chaisemartin and Ramirez-Cuellar (2020) for details.

3.2 | STRATIFIED DESIGNS WITH ATTRITION

In this section, we repeat the exercise presented in Section 3.1 but in the context of stratified designs. Before describing the estimators, we provide a description of the class of treatment assignment mechanisms we consider. In words, our results accommodate any treatment assignment mechanism which first partitions the covariate space into a finite number of "large" strata and then performs treatment assignment independently across strata so as to achieve "balance" within each stratum. Formally, let $S : \text{supp}(X_i) \to \mathcal{S}$ be a function which maps the support of the covariates into a finite set $\mathcal{S}$ of strata labels. For $1 \leq i \leq n$, let $S_i = S(X_i)$ denote the strata label of individual i. For $s \in \mathcal{S}$, let

$$D_n(s) = \sum_{1 \leq i \leq n} (D_i - \nu) I\{S_i = s\},$$

where $\nu \in (0, 1)$ denotes the "target" proportion of units to assign to treatment in each stratum. Intuitively, $D_n(s)$ measures the amount of imbalance in stratum s relative to the target proportion ν. Our requirements on the treatment assignment mechanism can then be summarized as follows:

Assumption 3.4. The treatment assignment mechanism is such that

(a) $W^{(n)} \perp\!\!\!\perp D^{(n)} | S^{(n)}$.
(b) $\frac{D_n(s)}{n} \xrightarrow{P} 0$ for every $s \in \mathcal{S}$.

Assumption 3.4(a) simply requires that treatment assignment be exogenous conditional on the strata labels. Assumption 3.4(b) formalizes the requirement that the assignment mechanism performs treatment assignment so as to achieve "balance" within strata. Assumption 3.4(b) is a relatively mild assumption which is satisfied by most stratified randomization procedures employed in field experiments: See Bugni et al. (2018) for examples.

As before, the first estimator we consider is the standard difference-in-means estimator computed on non-attritors $\hat{\theta}_n$. The second estimator we consider, denoted $\hat{\theta}_n^{\text{sfe}}$, is the estimator obtained as the estimator of the coefficient on D_i in an ordinary least squares regression of Y_i on D_i and strata fixed effects computed on the non-attritors:

$$Y_i = \theta^{\text{sfe}} D_i + \sum_{s \in \mathcal{S}} \delta_s I\{S_i = s\} + \epsilon_i \quad \text{(for individuals with } R_i = 1\text{)}.$$

Similar regression specifications are extremely common in the analysis of stratified randomized experiments. See, for example, Bruhn and McKenzie (2009), Duflo et al. (2015), Glennerster and Takavarasha (2013), de Mel et al. (2019), and Callen et al. (2020). Using appropriate laws of large numbers developed in Bugni et al. (2018), we now establish the following result:

Theorem 3.2. *Suppose the data satisfy Assumptions 2.1 and the treatment assignment mechanism satisfies Assumption 3.4, Then, as* $n \to \infty$, $\hat{\theta}_n \xrightarrow{P} \theta^{obs}$, *where*

$$\theta^{obs} = \frac{E[R_i(1)Y_i(1)]}{E[R_i(1)]} - \frac{E[R_i(0)Y_i(0)]}{E[R_i(0)]} = E[Y_i(1)|R_i(1) = 1] - E[Y_i(0)|R_i(0) = 1],$$

and $\hat{\theta}_n^{sfe} \xrightarrow{P} \theta^{sfe}$, *where*

$$\theta^{sfe} = \left(E\left[\frac{E[R_i(1)|S_i]E[R_i(0)|S_i]}{\nu E[R_i(1)|S_i] + (1-\nu)E[R_i(0)|S_i]} \right] \right)^{-1} \times E\left[\frac{E[R_i(1)Y_i(1)|S_i]E[R_i(0)|S_i] - E[R_i(0)Y_i(0)|S_i]E[R_i(1)|S_i]}{\nu E[R_i(1)|S_i] + (1-\nu)E[R_i(0)|S_i]} \right].$$

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The conclusions we draw from Theorem 3.2 closely mirror those of Theorem 3.1. In this case, under Assumption 2.3 with $C_i = S_i$,

$$\theta^{\text{sfe}} = E\left[\tau(S_i)\lambda(S_i)\right],$$

where $\tau(s) = E[Y_i(1) - Y_i(0)|S_i = s]$ and

$$\lambda(s) = \left(E\left[\frac{E[R_i(1)|S_i]E[R_i(0)|S_i]}{\nu E[R_i(1)|S_i] + (1-\nu)E[R_i(0)|S_i]}\right]\right)^{-1} \times \frac{E[R_i(1)|S_i = s]E[R_i(0)|S_i = s]}{\nu E[R_i(1)|S_i = s] + (1-\nu)E[R_i(0)|S_i = s]},$$

so that θ^{sfe} is also a convex-weighted average of the strata-level treatment effects $\tau(s)$, although the weights $\lambda(s)$ are arguably more complicated to interpret than the weights $\rho(x)$ defined in Section 3.1. Straightforward algebra shows that $\lambda(s) = 1$ if and only if

$$E[R_i(1)|S_i = s] = \frac{E[R_i(0)|S_i = s](1-\nu)\Lambda}{E[R_i(0)|S_i = s] - \Lambda\nu}, \tag{8}$$

where $\Lambda = E\left[\frac{E[R_i(1)|S_i]E[R_i(0)|S_i]}{\nu E[R_i(1)|S_i]+(1-\nu)E[R_i(0)|S_i]}\right]$. Conditions under which this holds seem difficult to articulate in words, but once again, a natural assumption that guarantees (8) for every $s \in \mathcal{S}$ is that Assumption 2.4 is satisfied with $C_i = S_i$. We summarize these observations in the following corollary:

Corollary 3.2.

(a) Under Assumption 2.2, $\theta^{obs} = \theta$.
(b) Under Assumption 2.3 with $C_i = S_i$, $\theta^{sfe} = E[\tau(S_i)\lambda(S_i)]$.
(c) Under Assumption 2.4 with $C_i = S_i$, $\theta^{sfe} = \theta^{obs}$.
(d) Under Assumption 2.4 with $C_i = S_i$ and either Assumption 2.2 or Assumption 2.3 with $C_i = S_i$, $\theta^{sfe} = \theta$.

We conclude this section by stating that given how closely the results presented in Section 3.2 mirror those in Section 3.1, we do not find compelling evidence to support the idea that stratifying into larger groups resolves the issues surrounding attrition that we explore in this paper.

4 | EMPIRICAL ILLUSTRATIONS

4.1 | Re-analysis of Groh and McKenzie (2016)

In this section, we illustrate the potential empirical relevance of deciding whether or not to drop pairs with an attrited unit using the experimental data collected in Groh and McKenzie (2016), which implemented a matched-pair design in the presence of attrition. The regression specifications in the paper contain pair fixed effects, which, as explained in Section 3.1, is mechanically equivalent to dropping pairs with an attrited unit when regressing outcomes on a constant and treatment.

Groh and McKenzie (2016) study the effect of insuring microenterprises (clients) against macroeconomic instability and political uncertainty in post-revolution Egypt. A baseline survey was completed for 2961 clients, who were then randomly assigned to treatment (1481 individuals) and control (1480 individuals) using a matched-pair design.[4] In Table 1, we reproduce the intention-to-treat estimates from tab. 7 of their paper, which presents estimated treatment effects on profits, revenues, employees, and household consumption. "Original" corresponds to the estimates obtained from running the regression specifications in the original paper which include pair fixed effects, and $\hat{\theta}_n$ corresponds to estimates obtained from running an identical regression specification without pair fixed effects (we note that we were able success-

[4]Per the authors, they "created matched pairs […] to minimize the Mahalanobis distance between the values of 13 variables that [they] hypothesized may determine loan take-up and investment decisions." The final assignment contained *one* stratum with 16 individuals, each belonging to a different branch office. We follow the authors' methodology in *keeping* this stratum when we conduct our analysis in Table 1. We drop these when we perform additional analyses in Table 2.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Summary of estimates obtained from empirical application: Groh and McKenzie (2016).

	Profits	High profit	Revenue	High revenue	Number employees	Any worker	Owner's hours	Monthly consumption
Original	−59.702	−0.009	−737.199	−0.020	−0.024	0.008	−0.655	−7.551
$\hat{\theta}_n$	−38.642	−0.007	−692.818	−0.020	−0.023	0.007	−0.773	−7.337
Attrition (%)	2.086	2.086	2.153	2.153	1.783	1.783	1.480	0.000

Note: For each outcome listed in tab. 7 of Groh and McKenzie (2016), we report (a) the original estimates obtained in paper ("Original"), (b) the estimate on treatment status without pair fixed effects ($\hat{\theta}_n$), and (c) the attrition rate in % by outcome, defined as [number of individuals with missing outcome/total number of individuals]. The regression specifications here include baseline covariates; see Table 2 for analogous results without baseline covariates included.

TABLE 2 Summary of additional estimates obtained from empirical application: Groh and McKenzie (2016).

	Profits	High profit	Revenue	High revenue	Number employees	Any worker	Owner's hours	Monthly consumption
Original	−91.197	−0.011	−967.967	−0.024	−0.032	0.004	−0.561	−3.600
$\hat{\theta}_n$	−80.058	−0.009	−888.608	−0.023	−0.026	0.005	−0.481	−3.600
Attrition (%)	1.755	1.755	1.824	1.824	1.411	1.411	1.514	0.000

Note: For each outcome regression specification listed in tab. 7 of Groh and McKenzie (2016), we report (a) the original estimates obtained in paper ("Original"), (b) the estimate on treatment status without pair fixed effects ($\hat{\theta}_n$), and (c) the attrition rate in % by outcome, defined as [number of individuals with missing outcome/total number of individuals]. The regression specifications here exclude baseline covariates from the authors' original work.

fully reproduce all of the reported estimates from the paper). We find an average absolute percentage difference of 13.82%[5] for the point estimates of these effects, with the largest differences appearing for profits and revenue.

One caveat to the findings in Table 1 is that the setting does not map exactly into our theoretical results: First, both regressions control for baseline covariates, and second, the final assignment contained *one* stratum with 16 individuals, each belonging to a different branch office. Given this, in Table 2, we report the intention-to-treat estimates without baseline covariates and without this additional stratum. In this case, we find an average absolute percentage difference of 15.61% for the point estimates of the effects. We emphasize that we consider these difference particularly salient given that attrition is quite low (on average 1.4% across the outcomes) and that in the absence of attrition, these estimates would be *numerically identical*, as illustrated from the estimates of the effect of treatment for monthly consumption.

4.2 | Re-analysis of recent publications in AER and AEJ: Applied

Next, we perform a similar exercise using the data from a systematic survey of all papers published in the AER and the AEJ: Applied from 2020 to 2022 which conducted matched-pair or stratified randomized experiments in the presence of attrition. Our survey identified seven such papers: Abebe et al. (2021), Attanasio et al. (2020), Carter et al. (2021), Casaburi and Reed (2022), Dhar et al. (2022), Hjort et al. (2021), and Romero et al. (2020). For each paper, we collected a set of "relevant" regression specifications[6] and reproduced these regressions with and without pair/stratum fixed effects (we note that we were able to successfully reproduce all of the reported estimates from each paper). In Figure 1, we report the average absolute percentage change (computed as $\left(\frac{|\text{Alternative}-\text{Original}|}{|\text{Original}|}\right)\times 100$, where "Original" corresponds to the point estimate computed in the paper and "Alternative" corresponds to the estimate computed from the alternative specification with or without fixed effects) across all specifications for each paper. Similar to our findings for Groh and McKenzie (2016), we find that there can be noticeable differences in the point estimates with and without fixed effects (although we emphasize that we do not claim that these differences are necessarily statistically significant).

[5]Here, the absolute percentage difference is computed as $\left(\frac{|\text{Original}-\hat{\theta}_n|}{|\text{Original}|}\right)\times 100$.

[6]We note that in some papers such as Attanasio et al. (2020) and Casaburi and Reed (2022), the primary results were not necessarily the output of a linear regression, and so in these cases, we selected a collection of preliminary regression analyses. In other papers such as Hjort et al. (2021), the primary results were LATE estimates obtained via IV regression, and so in these cases, we report the intention-to-treat analyses. Specific selection details for each paper are outlined in Appendix A.5 in the Supporting Information.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

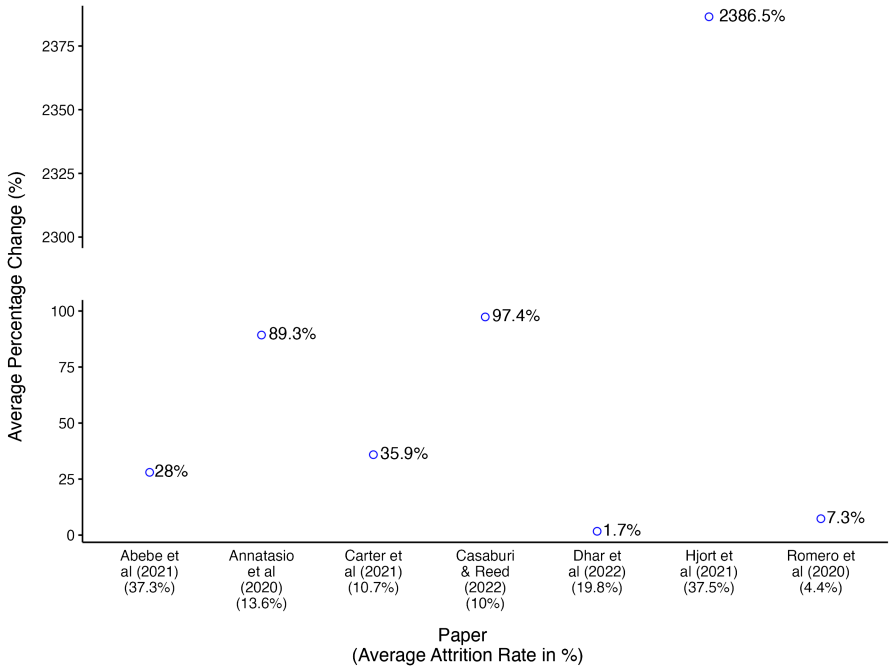

FIGURE 1 Average absolute percentage difference for "Original" versus "Alternative" point estimates. Average attrition rate, defined as [number of individuals with missing outcome/total number of individuals], is reported in parentheses below each author label.

5 | RECOMMENDATIONS FOR EMPIRICAL PRACTICE

We conclude with some recommendations for empirical practice based on our theoretical results. Our main takeaway is that choosing whether or not to include pair/strata fixed effects when attrition is a concern can make a substantive difference to empirical findings and to the interpretation of the resulting estimand. In our view, unless practitioners are interested in recovering the convex-weighted averages produced by θ^{drop} and θ^{sfe} under a conditional independence assumption (Assumption 2.3), primary analyses should be based on regressions *without* pair/strata fixed effects: The resulting estimand θ^{obs} has a simple interpretation in the absence of any assumptions and collapses to the average treatment effect under arguably weaker assumptions than θ^{drop} and θ^{sfe}. A secondary benefit of θ^{obs} is that under the additional assumption that $R_i(1) = R_i(0)$, θ^{obs} *also* enjoys an interpretation as a convex-weighted average under Assumption 2.3, with weights which may be more desirable than those appearing in θ^{drop} or θ^{sfe} in that they do not "double-up" on attrition: See Remark 3.3 for details.

OPEN RESEARCH BADGES

This article has been awarded Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. Data is available at DOI: 10.15456/jae.2023324.2137435210.

DATA AVAILABILITY STATEMENT

The data and code used in this paper, including links to the original data files, have been made available at the Journal of Applied Econometrics' repository: https://journaldata.zbw.eu/dataset/revisiting-the-analysis-of-matched-pair-and-stratified-experiments-in-the-presence-of-attrition.

ORCID

Meng Hsuan Hsieh https://orcid.org/0000-0002-8382-5352

REFERENCES

Abebe, G., Caria, A. S., & Ortiz-Ospina, E. (2021). The selection of talent: Experimental and structural evidence from Ethiopia. *American Economic Review*, *111*(6), 1757–1806.

Angrist, J., & Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American Economic Review*, *99*(4), 1384–1414. https://www.aeaweb.org/articles?id=10.1257/aer.99.4.1384

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Ashraf, N., Karlan, D., & Yin, W. (2006). Deposit collectors. *Advances in Economic Analysis & Policy*, *5*(2), 121–144.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2020). Estimating the production function for human capital: Results from a randomized controlled trial in Colombia. *American Economic Review*, *110*(1), 48–85.

Bai, Y. (2022). Optimality of matched-pair designs in randomized controlled trials. *American Economic Review*, *112*(12), 3911–3940. https://doi.org/10.1257/aer.20201856

Bai, Y., Liu, J., & Tabord-Meehan, M. (2022). Inference for matched tuples and fully blocked factorial designs. arXiv preprint arXiv:2206.04157.

Bai, Y., Romano, J. P., & Shaikh, A. M. (2022). Inference in experiments with matched pairs. *Journal of the American Statistical Association*, *117*(540), 1726–1737. https://doi.org/10.1080/01621459.2021.1883437

Behaghel, L., Crépon, B., Gurgand, M., & Le Barbanchon, T. (2015). Please call again: Correcting nonresponse bias in treatment effect models. *Review of Economics and Statistics*, *97*(5), 1070–1080.

Bruhn, M., Leo, L. S., Legovini, A., Marchetti, R., & Zia, B. (2016). The impact of high school financial education: Evidence from a large-scale evaluation in Brazil. *American Economic Journal: Applied Economics*, *8*(4), 256–295. https://www.aeaweb.org/articles?id=10.1257/app.20150149

Bruhn, M., & McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, *1*(4), 200–232. https://www.aeaweb.org/articles?id=10.1257/app.1.4.200

Bugni, F. A., Canay, I. A., & Shaikh, A. M. (2018). Inference under covariate-adaptive randomization. *Journal of the American Statistical Association*, *113*(524), 1784–1796. https://doi.org/10.1080/01621459.2017.1375934

Callen, M., Gulzar, S., Hasanain, A., Khan, M. Y., & Rezaee, A. (2020). Data and policy decisions: Experimental evidence from Pakistan. *Journal of Development Economics*, *146*, 102523.

Carter, M., Laajaj, R., & Yang, D. (2021). Subsidies and the African green revolution: Direct effects and social network spillovers of randomized input subsidies in Mozambique. *American Economic Journal: Applied Economics*, *13*(2), 206–229.

Casaburi, L., & Reed, T. (2022). Using individual-level randomized treatment to learn about market structure. *American Economic Journal: Applied Economics*, *14*(4), 58–90.

Crpon, B., Devoto, F., Duflo, E., & Parient, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics*, *7*(1), 123–150. https://www.aeaweb.org/articles?id=10.1257/app.20130535

de Chaisemartin, C., & Ramirez-Cuellar, J. (2020). At what level should one cluster standard errors in paired experiments, and in stratified experiments with small strata?: National Bureau of Economic Research.

de Mel, S., McKenzie, D., & Woodruff, C. (2019). Labor drops: Experimental evidence on the return to additional labor in microenterprises. *American Economic Journal: Applied Economics*, *11*(1), 202–235. http://www.aeaweb.org/articles?id=10.1257/app.20170497

Dhar, D., Jain, T., & Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American Economic Review*, *112*(3), 899–927.

Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*. Wiley.

Duflo, E., Dupas, P., & Kremer, M. (2015). Education, HIV, and early fertility: Experimental evidence from Kenya. *American Economic Review*, *105*(9), 2757–97.

Fryer, R. (2018). The "pupil" factory: Specialization and the production of human capital in schools. *American Economic Review*, *108*(3), 616–656. https://www.aeaweb.org/articles?id=10.1257/aer.20161495

Fukumoto, K. (2022). Nonignorable attrition in pairwise randomized experiments. *Political Analysis*, *30*(1), 132–141.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*: W.W. Norton.

Ghanem, D., Hirshleifer, S., & Ortiz-Becerra, K. (2021). Testing attrition bias in field experiments.

Glennerster, R., & Takavarasha, K. (2013). *Running randomized evaluations: A practical guide*: Princeton University Press.

Groh, M., & McKenzie, D. (2016). Macroinsurance for microenterprises: A randomized experiment in post-revolution Egypt. *Journal of Development Economics*, *118*, 13–25.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society*, *1979*, 153–161.

Hirano, K., Imbens, G. W., Ridder, G., & Rubin, D. B. (2001). Combining panel data sets with attrition and refreshment samples. *Econometrica*, *69*(6), 1645–1659. https://doi.org/10.1111/1468-0262.00260

Hjort, J., Moreira, D., Rao, G., & Santini, J. F. (2021). How research affects policy: Experimental evidence from 2,150 Brazilian municipalities. *American Economic Review*, *111*(5), 1442–1480.

Horowitz, J. L., & Manski, C. F. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association*, *95*(449), 77–84. http://www.jstor.org/stable/2669526

King, G., Gakidou, E., Ravishankar, N., Moore, R. T., Lakin, J., Vargas, M., Téllez-Rojo, M. M., Hernández Ávila, J. E., Ávila, M. H., & Llamas, H. H. (2007). A "politically robust" experimental design for public policy evaluation, with application to the Mexican universal health insurance program. *Journal of Policy Analysis and Management*, *26*(3), 479–506.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, *76*(3), 1071–1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

Little, RJA, & Rubin, D. B. (2019). *Statistical analysis with missing data*, Vol. *793*: John Wiley & Sons.

Moffit, R., Fitzgerald, J., & Gottschalk, P. (1999). Sample attrition in panel data: The role of selection on observables. *Annales d'Economie et de Statistique*, *1999*, 129–152.

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Romero, M., Sandefur, J., & Sandholtz, W. A. (2020). Outsourcing education: Experimental evidence from Liberia. *American Economic Review*, *110*(2), 364–400.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*, Vol. *81*: John Wiley & Sons.

Wooldridge, J. M. (2002). Inverse probability weighted *M*-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal*, *1*(2), 117–139.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of the article.

How to cite this article: Bai Y., Hsieh M. H., Liu J., & Tabord-Meehan M. (2023). Revisiting the analysis of matched-pair and stratified experiments in the presence of attrition. *Journal of Applied Econometrics*, 1–13 https://doi.org/10.1002/jae.3025

10991255, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jae.3025, Wiley Online Library on [15/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License