Evaluating and Improving Targeting Policies with Field Experiments Using Counterfactual Policy Logging

Duncan Simester[*], Artem Timoshenko[*], and Spyros I. Zoumpoulis[†]

[*]Marketing, MIT Sloan School of Management, Massachusetts Institute of Technology
[†]Decision Sciences, INSEAD

September 7, 2018

Appendices

Contents

A STZ Study: Outcomes for the Experimental Conditions Associated with the Three Actions

Table 1: STZ Study: Outcomes for the Experimental Conditions Associated with the Three Actions

	Recommended Action		**Sample Size**			**Average Profit**		
	Policy 1	Policy 2	Action a	Action b	Action c	Action a	Action b	Action c
Row	s	t	$N_a^{1:s,2:t}$	$N_b^{1:s,2:t}$	$N_c^{1:s,2:t}$			
1	a	a	595	2,562	824	−\$0.330	−\$0.004	\$1.004
2	a	b	13,038	11,495	16,086	\$0.097	\$0.016	\$0.249
3	a	c	20,229	8,148	16,164	−\$0.131	\$0.008	\$0.293
4	b	a	15,602	12,847	18,784	\$1.847	\$1.452	\$1.288
5	b	b	198,416	215,824	195,885	\$1.445	\$1.610	\$1.194
6	b	c	51,421	45,211	65,311	\$1.220	\$1.231	\$0.824
7	c	a	1,239	2,098	70	−\$0.505	−\$0.519	\$0.000
8	c	b	19,551	19,972	16,215	−\$0.263	−\$0.052	\$0.112
9	c	c	37,562	42,616	41,597	−\$0.302	−\$0.277	\$0.143
	Total		357,653	360,773	370,936	\$1.005	\$1.131	\$0.888

The table reports outcomes from an experiment reported by Simester, Timoshenko, and Zoumpoulis (2018). It reports the sample size and average profit in the three experimental conditions associated with actions a, b, and c. The outcomes are reported for each customer segment, where the customer segments are constructed using the actions recommended by Policies 1 and 2. The shading identifies the outcomes used to evaluate Policy 1. The profits are all multiplied by a common random number.

B A Formal Analysis of Efficiency

We contemplate an experimental design in which each customer i is assigned to one of T policies, or not assigned to any of the policies, by randomly selecting customers from a super-population of size N. Letting W_i be a random variable indicating what policy customer i is assigned to, we have $W_i \in \{1, 2, \ldots, T, 0\}$, where $W_i = 0$ denotes the outcome when customer i is not sampled to any of the policies (and therefore does not participate in the experiment). This random selection and assignment to the policies is the source of variation in the system. We use $N_1, \ldots, N_T, N_0$ to denote the number of customers assigned to each of $\{1, 2, \ldots, T, 0\}$, so that $N_1 + \ldots + N_T + N_0 = N$.

We assume a completely randomized experiment:

Assumption 1 (Random sampling and assignment without replacement)**.** Conditional on $N_1, \ldots, N_T, N_0$, the vector $\mathbf{W} \in \mathbf{R}^N$ has multinomial distribution, with

$$\Pr(\mathbf{W} = \mathbf{w} \mid N_1, \ldots, N_T, N_0) = \begin{cases} \frac{N_1! \cdots N_T! N_0!}{N!} & \text{for all } \mathbf{w} \text{ with} \\ \quad \sum_{i=1}^{N} \mathbb{1}_{w_i=1} = N_1, \ldots, \sum_{i=1}^{N} \mathbb{1}_{w_i=T} = N_T, \sum_{i=1}^{N} \mathbb{1}_{w_i=0} = N_0, \\ 0 & \text{otherwise.} \end{cases}$$

Customers are described by a set of covariates. We think of the policies as mappings from the space of covariates to a selection of an action. For example, in the context of prospecting new customers for a retailer, the actions are the types of promotional offers the retailer considers mailing to different households according to their covariates, including the control treatment of no mail.

We consider a finite set of possible actions. We follow the potential outcomes framework: for each customer i and action s, we define the outcome $Y_i(s)$ that would have occurred, had customer i been treated with action s. (With some abuse of notation, and for notational simplicity, we also define $Y_i(j)$ to be the outcome that would have occured, had customer i been treated with the action assigned to her by policy j.) We define $Y_i(s)$ regardless of whether customer i is actually treated with action s or not. Only one of the potential outcomes is realized and observed for customer i, which we denote by Y_i^{obs}. We do not observe what would have happened had customer i been exposed to other actions. The uncertainty therefore does not only come from random sampling from the super-population, but also from the unobserved potential outcomes, in the spirit of Abadie et al. (2014).

We first describe how to evaluate a single policy under both the standard and proposed approaches. We then describe how to compare two policies. We summarize the resulting mean and variance calculations under each approach in Table 2.

Evaluating a Single Policy Using the Standard Approach

We label the policy of interest as policy 1. Our target (estimand) when evaluating policy 1 is the population-level measure

$$y_1 = \frac{1}{N}\sum_{i=1}^{N} Y_i(1).$$

We do not observe the outcomes for the entire population under each policy. Therefore, the standard approach is to estimate y_1 by

$$\hat{y}_1 = \frac{1}{N_1}\sum_{i=1}^{N} \mathbb{1}_{W_i=1} Y_i^{obs}.$$

It is straightforward to establish that $\hat{y}_1$ is an unbiased estimator of y_1. As we show in Appendix B.1, the variance of this estimator is given by

$$\mathrm{Var}_W(\hat{y}_1) = \frac{S_1^2}{N_1}\left(1 - \frac{N_1}{N}\right),$$

where

$$S_1^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i(1) - \frac{1}{N}\sum_{j=1}^{N} Y_j(1)\right)^2. \tag{1}$$

We estimate $\mathrm{Var}(\hat{y}_1)$ by

$$\widehat{\mathrm{Var}}_W(\hat{y}_1) = \frac{s_1^2}{N_1}\left(1 - \frac{N_1}{N}\right),$$

where

$$s_1^2 = \frac{1}{N_1 - 1}\sum_{i:W_i=1}\left(Y_i^{obs} - \hat{y}_1\right)^2 \tag{2}$$

is an unbiased estimator for S_1^2 (Imbens and Rubin, 2015, Chapter 6.5, Appendix A).

We next discuss how to estimate y_1 under the proposed approach.

Evaluating a Single Policy Using the Proposed Approach

Even though we are evaluating a single policy, the proposed approach recommends first segmenting customers using the recommended actions from at least two policies. For this illustration we will continue to evaluate policy 1, but will use the recommended actions from both policies 1 and 2. We must first introduce some additional notation to identify segments of customers, where the segments are constructed using the recommended actions from each policy.

Define $g(1:s, 2:t)$ as the segment of customers in the super-population to whom policy 1 would assign action s and policy 2 would assign action t. Moreover, $N^{1:s,2:t}$ is the number of customers in segment $g(1:s,2:t)$, and $N_1^{1:s,2:t}$ is the number of customers in segment $g(1:s,2:t)$ that are randomly assigned to receive policy 1.

Our goal is unchanged. We are after the population-level measure

$$y_1 = \frac{1}{N}\sum_{i=1}^{N} Y_i(1),$$

which can also be written as

$$y_1 = \frac{\sum_{s,t} \sum_{i \in g(1:s,2:t)} Y_i(s)}{\sum_{s,t} N^{1:s,2:t}} = \frac{\sum_{s,t} N^{1:s,2:t} \cdot y_1^{1:s,2:t}}{N}, \tag{3}$$

where $y_1^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} Y_i(s)}{N^{1:s,2:t}}$. Equation (3), that is, calculates a weighted average of the outcomes in each segment, where the segments are as many as the square of the number of possible actions, and where the weights are the proportions of each of the segments in the super-population.

Under the proposed approach we estimate y_1 by calculating the outcome in each segment and then aggregating across segments:

$$\hat{y}_1 = \frac{\sum_{s,t} N^{1:s,2:t} \cdot \hat{y}_1^{1:s,2:t}}{N},$$

with

$$\hat{y}_1^{1:s,2:t} = \begin{cases} \frac{\sum_{i \in g(1:s,2:t)} \mathbb{1}_{W_i=1} Y_i^{obs}}{N_1^{1:s,2:t}} & \text{if } s \neq t, \\ \frac{\sum_{i \in g(1:s,2:s)} \left(\mathbb{1}_{W_i=1} + \mathbb{1}_{W_i=2}\right) Y_i^{obs}}{N_1^{1:s,2:s} + N_2^{1:s,2:s}} & \text{if } s = t, \end{cases}$$

where the first case is for segments for which the two policies recommend different actions, while the second case is for segments for which the two policies recommend the same action.

Pooling customers across the two conditions in segments for which the two policies recommend the same action, irrespective of which of the two policies they were assigned to, is one of the core ideas of our paper. The efficiency gain of pooling is twofold. First, the sample size used to measure the outcome for customers for whom the two policies would recommend the same action increases. Second, as we illustrate later on in this section, the difference in the performance of the two policies in the segments for which the two policies recommend the same action is estimated to be zero, as it truly is. This eliminates the random error in these segments when comparing the performance of the two policies.

It is straightforward to establish that $\hat{y}_1$ is an unbiased estimator of y_1 (see Appendix B.2). We can also write expressions for the variance of the estimator $\hat{y}_1$ within each segment. To do so we again distinguish between segments in which the two policies recommend the same or different actions.

For segments for which the two policies recommend different actions (i.e., $s \neq t$), we explain in Appendix B.3 that the variance is given by

$$\text{Var}_W\left(\hat{y}_1^{1:s,2:t}\right) = \frac{\left(S_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} \left(1 - \frac{N_1^{1:s,2:t}}{N^{1:s,2:t}}\right),$$

where

$$\left(S_1^{1:s,2:t}\right)^2 = \frac{1}{N^{1:s,2:t}-1} \sum_{i \in g(1:s,2:t)} \left(Y_i(s) - \frac{1}{N^{1:s,2:t}} \sum_{j \in g(1:s,2:t)} Y_j(s) \right)^2 . \tag{4}$$

We estimate $\mathrm{Var}_W\left(\hat{y}_1^{1:s,2:t}\right)$ by

$$\widehat{\mathrm{Var}}_W\left(\hat{y}_1^{1:s,2:t}\right) = \frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} \left(1 - \frac{N_1^{1:s,2:t}}{N^{1:s,2:t}}\right),$$

where

$$\left(s_1^{1:s,2:t}\right)^2 = \frac{1}{N_1^{1:s,2:t}-1} \sum_{i \in g(1:s,2:t), W_i=1} \left(Y_i^{obs} - \hat{y}_1^{1:s,2:t}\right)^2 \tag{5}$$

is an unbiased estimator for $\left(S_1^{1:s,2:t}\right)^2$ (Imbens and Rubin, 2015, Chapter 9.5).

We next consider segments in which the two policies recommend the same action (i.e., $s = t$). As we show in Appendix B.3, the variance in these segments is given by

$$\mathrm{Var}_W\left(\hat{y}_1^{1:s,2:s}\right) = \frac{\left(S^{1:s,2:s}\right)^2}{N_1^{1:s,2:s} + N_2^{1:s,2:s}} \left(1 - \frac{N_1^{1:s,2:s} + N_2^{1:s,2:s}}{N^{1:s,2:s}}\right),$$

where

$$\left(S^{1:s,2:s}\right)^2 = \frac{1}{N^{1:s,2:s}-1} \sum_{i \in g(1:s,2:s)} \left(Y_i(s) - \frac{1}{N^{1:s,2:s}} \sum_{j \in g(1:s,2:s)} Y_j(s) \right)^2 .$$

We estimate $\mathrm{Var}_W\left(\hat{y}_1^{1:s,2:s}\right)$ by

$$\widehat{\mathrm{Var}}_W\left(\hat{y}_1^{1:s,2:s}\right) = \frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s} + N_2^{1:s,2:s}} \left(1 - \frac{N_1^{1:s,2:s} + N_2^{1:s,2:s}}{N^{1:s,2:s}}\right),$$

where

$$\left(s^{1:s,2:s}\right)^2 = \frac{1}{N_1^{1:s,2:s} + N_2^{1:s,2:s} - 1} \sum_{i \in g(1:s,2:s), W_i \in \{1,2\}} \left(Y_i^{obs} - \hat{y}_1^{1:s,2:s}\right)^2$$

is an unbiased estimator for $\left(S^{1:s,2:s}\right)^2$.

Notice the benefit from pooling across both experimental conditions. In particular, the variance is reduced because the sample includes data from the experimental condition associated with both policy 1 and policy 2. This is an efficiency advantage over the standard approach, which does not use any data from the experimental condition associated with policy 2 when evaluating policy 1.

Overall, the variance of estimator $\hat{y}_1$ across all segments is

$$\mathrm{Var}_W(\hat{y}_1) = \sum_{s,t} \left(\frac{N^{1:s,2:t}}{N}\right)^2 \mathrm{Var}_W\left(\hat{y}_1^{1:s,2:t}\right),$$

which we estimate by

$$\widehat{\mathrm{Var}}_W(\hat{y}_1) = \sum_{s,t} \left(\frac{N^{1:s,2:t}}{N} \right)^2 \widehat{\mathrm{Var}}_W \left(\hat{y}_1^{1:s,2:t} \right).$$

Next we describe estimators for estimating the difference in two policies.

Comparing Two Policies Using the Standard Approach

When comparing two policies, the standard approach is to first calculate the mean outcome for each policy and then calculate the difference in these means. For this illustration we will calculate this difference as the outcome for policy 1 minus the outcome for policy 2.

The target estimand is the population-level measure

$$y_{1-2} = y_1 - y_2 = \frac{1}{N} \sum_{i=1}^{N} \Big(Y_i(1) - Y_i(2) \Big).$$

Because we do not observe the outcomes for the entire population under each policy, we estimate y_{1-2} using the unbiased estimator

$$\hat{y}_{1-2} = \hat{y}_1 - \hat{y}_2 = \frac{1}{N_1} \sum_{i=1}^{N} \mathbb{1}_{W_i=1} Y_i^{obs} - \frac{1}{N_2} \sum_{i=1}^{N} \mathbb{1}_{W_i=2} Y_i^{obs}.$$

As shown in Appendix B.4, the variance of this estimator is given by

$$\mathrm{Var}_W \left(\hat{y}_{1-2} \right) = \frac{S_1^2}{N_1} + \frac{S_2^2}{N_2} - \frac{S_{1,2}^2}{N}, \tag{6}$$

where variances S_1^2, S_2^2 are given by Equation (1), and where the third term includes the variance of the customer-level differences between policies:

$$S_{1,2}^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(Y_i(1) - Y_i(2) - \frac{1}{N} \sum_{j=1}^{N} \Big(Y_j(1) - Y_j(2) \Big) \right)^2.$$

We note that, perhaps counterintuitively, the variance $\mathrm{Var}_W \left(\hat{y}_{1-2} \right)$ is larger when the treatment effect is constant across customers. This happens because complete randomization ensures that the possible values of the estimator $\hat{y}_{1-2}$ will be further apart, in the case of a constant (or less heterogeneous) treatment effect across customers.

We estimate S_1^2, S_2^2 with s_1^2, s_2^2 respectively, as defined in Equation (2). The term $S_{1,2}^2$ is in general impossible to estimate empirically because we do not observe the outcome of both policies 1 and 2 for the same customer. It is common practice to thus approximate the variance $\mathrm{Var}_W \left(\hat{y}_{1-2} \right)$ using the first two terms and ignoring the third term of Equation (6), with the Neyman estimator

(Neyman, 1934)

$$\mathrm{Var}^{\mathrm{Neyman}}_{1-2} = \frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}.$$

As the term we ignored is non-negative, this Neyman estimator of the variance in the estimate of the difference between the two policies is generally upwardly biased, leading to conservative confidence intervals (i.e., with actual coverage at least as large as their nominal coverage).[1]

Comparing Two Policies Using the Proposed Approach

We complete this formal description by showing how to compare two policies under the proposed approach. Our target estimand is unchanged:

$$y_{1-2} = y_1 - y_2 = \frac{1}{N}\sum_{i=1}^{N}\Big(Y_i(1) - Y_i(2)\Big),$$

which, accounting for the segmentation of the customers according to the recommended actions, can also be written as

$$y_{1-2} = \frac{\sum_{s,t}\sum_{i\in g(1:s,2:t)}\Big(Y_i(s) - Y_i(t)\Big)}{\sum_{s,t} N^{1:s,2:t}}.$$

One of the main contributions of the paper is to recognize that in segments in which the two policies recommend the same action, the difference in the performance of the two policies is equal to zero. We can thus write

$$y_{1-2} = \frac{\sum_{\substack{s,t\\ s\neq t}}\sum_{i\in g(1:s,2:t)}\Big(Y_i(s) - Y_i(t)\Big)}{\sum_{s,t} N^{1:s,2:t}} = \frac{\sum_{\substack{s,t\\ s\neq t}} N^{1:s,2:t}\Big(y_1^{1:s,2:t} - y_2^{1:s,2:t}\Big)}{N},$$

where, for $s \neq t$,

$$y_1^{1:s,2:t} = \frac{\sum_{i\in g(1:s,2:t)} Y_i(s)}{N^{1:s,2:t}}, \qquad y_2^{1:s,2:t} = \frac{\sum_{i\in g(1:s,2:t)} Y_i(t)}{N^{1:s,2:t}}.$$

We estimate y_{1-2} with the unbiased estimator

$$\hat{y}_{1-2} = \frac{\sum_{s\neq t} N^{1:s,2:t}\Big(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t}\Big)}{N},$$

[1] A positive lower bound on the population variance of the customer-level treatment effect, $S^2_{1,2}$, would result in an estimator of $\mathrm{Var}_W\left(\hat{y}_{1-2}\right)$ that is less (upwardly) biased than the Neyman estimator we propose. In practice, such a (non-zero) lower bound is in general not easy to provide, and as a result researchers in practice use the Neyman estimator that we also propose (Athey and Imbens, 2017). Other than the fact that it provides a simple bound irrespective of the heterogeneity in the treatment effect, there is another reason for the popularity of the Neyman estimator (Imbens and Rubin, 2015): irrespective of the heterogeneity in the individual treatment effect, the Neyman estimator is always unbiased for the sampling variance of the estimator of the infinite super-population, as opposed to the finite sample, average treatment effect.

where, for $s \neq t$,

$$\hat{y}_1^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} \mathbb{1}_{W_i=1} Y_i^{obs}}{N_1^{1:s,2:t}}, \qquad \hat{y}_2^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} \mathbb{1}_{W_i=2} Y_i^{obs}}{N_2^{1:s,2:t}}.$$

In particular, the *true* difference in the outcome between the two policies is zero in segments for which the two policies recommend the same action. Remember that because of pooling customers across the two conditions in these segments, the *observed* difference in the outcomes between the two policies in these segments is also zero. In contrast, under the standard approach, random error would have been introduced, and the observed outcomes in the two conditions would have been different, although we know the true outcomes should be identical in segments where the policies recommend the same action.

In segments for which the two policies recommend different actions ($s \neq t$), the variance of the difference between the two policies is given by[2]

$$\mathrm{Var}_W\left(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t}\right) = \frac{\left(S_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} + \frac{\left(S_2^{1:s,2:t}\right)^2}{N_2^{1:s,2:t}} - \frac{\left(S_{1,2}^{1:s,2:t}\right)^2}{N^{1:s,2:t}},$$

where variances $\left(S_1^{1:s,2:t}\right)^2, \left(S_2^{1:s,2:t}\right)^2$ are given by Equation (4), and the variance of the customer-level differences between policies is

$$\left(S_{1,2}^{1:s,2:t}\right)^2 = \frac{1}{N^{1:s,2:t} - 1} \sum_{i \in g(1:s,2:t)} \left(Y_i(s) - Y_i(t) - \frac{1}{N^{1:s,2:t}} \sum_{j \in g(1:s,2:t)} \left(Y_j(s) - Y_j(t)\right)\right)^2. \tag{7}$$

We estimate $\left(S_1^{1:s,2:t}\right)^2, \left(S_2^{1:s,2:t}\right)^2$ with $\left(s_1^{1:s,2:t}\right)^2, \left(s_2^{1:s,2:t}\right)^2$ respectively, as defined in Equation (5). The term $\left(S_{1,2}^{1:s,2:t}\right)^2$ is in general impossible to estimate empirically because we never observe the outcome of both actions s, t for the same customer. This is the same issue that arises with the standard approach. We again use the Neyman variance estimator

$$\mathrm{Var}_{g(1:s,2:t)}^{\mathrm{Neyman}} = \frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} + \frac{\left(s_2^{1:s,2:t}\right)^2}{N_2^{1:s,2:t}},$$

which is generally upwardly biased.

Overall, the variance of estimator $\hat{y}_{1-2}$ across all segments is

$$\mathrm{Var}_W\left(\hat{y}_{1-2}\right) = \sum_{\substack{s,t \\ s \neq t}} \left(\frac{N^{1:s,2:t}}{N}\right)^2 \mathrm{Var}_W\left(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t}\right),$$

[2]The derivation for the variance of the difference in each segment is similar to the derivation of Equation (6) for the overall variance under the standard approach.

which we estimate with

$$\text{Var}_{1-2}^{\text{Neyman}} = \sum_{\substack{s,t \\ s \neq t}} \left(\frac{N^{1:s,2:t}}{N} \right)^2 \text{Var}_{g(1:s,2:t)}^{\text{Neyman}}.$$

This is generally upwardly biased and, as a result, confidence intervals will be conservative.

Summary

As we mentioned at the start of this section, in Table 2 we summarize the estimators and variances for the performance of a single policy and comparison of two policies under the standard and proposed approaches. We note that under both the standard approach and the proposed approach, the calculations for the variances include an adjustment for the size of the super-population. We also include this adjustment in our calculations in the example in the next section. However, the adjustment asymptotes to zero as the super-population grows large and so for large enough populations it is often ignored.

Table 2: Evaluating and Comparing Policies: Estimators and Variances

	Standard Approach	**Proposed Approach**
Single Policy Evaluation		
Estimator $\hat{y}_{P_1}$	$\frac{1}{N_1} \sum_{i=1}^{N} \mathbb{1}_{W_i=1} Y_i^{obs}$	$\frac{\sum_{s,t} N^{1:s,2:t} \cdot \hat{y}_1^{1:s,2:t}}{N}$
Variance $\widehat{\text{Var}}_W(\hat{y}_1)$	$\frac{s_1^2}{N_1}\left(1 - \frac{N_1}{N}\right)$	$\sum_{\substack{s,t \\ s \neq t}} \left(\frac{N^{1:s,2:t}}{N}\right)^2 \frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} \left(1 - \frac{N_1^{1:s,2:t}}{N^{1:s,2:t}}\right) + \sum_s \left(\frac{N^{1:s,2:s}}{N}\right)^2 \frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s} + N_2^{1:s,2:s}} \cdot \left(1 - \frac{N_1^{1:s,2:s} + N_2^{1:s,2:s}}{N^{1:s,2:s}}\right)$
Comparison of Two Policies		
Estimator $\hat{y}_{1-2}$	$\frac{1}{N_1} \sum_{i=1}^{N} \mathbb{1}_{W_i=1} Y_i^{obs} - \frac{1}{N_2} \sum_{i=1}^{N} \mathbb{1}_{W_i=2} Y_i^{obs}$	$\frac{\sum_{\substack{s,t \\ s \neq t}} N^{1:s,2:t} \left(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t}\right)}{N}$
Variance $\text{Var}_{1-2}^{\text{Neyman}}$	$\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}$	$\sum_{\substack{s,t \\ s \neq t}} \left(\frac{N^{1:s,2:t}}{N}\right)^2 \left(\frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} + \frac{\left(s_2^{1:s,2:t}\right)^2}{N_2^{1:s,2:t}}\right)$

The table summarizes the estimators and variances for the performance of a single policy and comparison of two policies under the standard and proposed approaches.

We proceed with a formal comparison of the efficiency of the standard and the proposed approaches and several comments, including a brief discussion of limitations.

Comparison of the Efficiency of the Proposed vs. the Standard Approach

The expressions in Table 2 allow us to analytically compare the efficiency of the proposed and standard approaches. We have argued that our proposed method has two efficiency advantages: it reduces variance introduced by between-segment differences; and it ensures the true and observed difference between two policies in segments in which they recommend the same action is zero, instead of introducing random noise. We now show analytically that the standard errors of the estimators strictly improve under the proposed method.

Our argument requires two assumptions. The first assumption is that the experimental design is balanced, i.e., for a policy we are evaluating, the proportion of customers assigned to that policy within a segment is the same as the proportion of customers assigned to that policy in any other segment:

$$\frac{N_1^{1:s,2:t}}{N^{1:s,2:t}} = \frac{N_1}{N}, \text{ for all } s, t,$$

and the same is true for all policies we are evaluating. For example, if we are also evaluating policy 2, then

$$\frac{N_2^{1:s,2:t}}{N^{1:s,2:t}} = \frac{N_2}{N}, \text{ for all } s, t.$$

The second assumption is that for all policies we are evaluating, the observed variance within any segment that received the policy is not larger than the observed variance across all observations in that condition:

$$\left(s_1^{1:s,2:t}\right)^2 \leq s_1^2, \quad \left(s^{1:s,2:s}\right)^2 \leq s_1^2, \quad \text{for all } s, t.$$

This assumption is aligned with the well-known intended benefit of stratification (e.g., Imbens and Rubin, 2015): we segment in order to achieve balance in the covariates. This means that the units within each block would be similar with respect to the covariates or some functions of the covariates. It therefore makes sense to expect that within any segment, the observed variance is not larger than in the aggregate.

The following theorem formalizes our result that the proposed method strictly improves efficiency. The proof is in Appendix B.5.

Theorem 1. *Assume random sampling and assignment without replacement as per Assumption 1. Further, assume that* $\frac{N_1^{1:s,2:t}}{N_1} = \frac{N_2^{1:s,2:t}}{N_2} = \frac{N^{1:s,2:t}}{N}$ *and that* $\left(s_1^{1:s,2:t}\right)^2 \leq s_1^2$, $\left(s_2^{1:s,2:t}\right)^2 \leq s_2^2$ *for all actions* s, t *with* $s \neq t$, *and* $\left(s^{1:s,2:s}\right)^2 \leq s_1^2$ *for all actions* s. *Then for the evaluation of policy* 1, *the estimated variance of the estimator under the proposed approach is strictly less than the estimated variance of the estimator under the standard approach. Furthermore, for the comparison of policies* 1 *and* 2, *the estimated variance of the estimator of the difference under the proposed approach is strictly less than the estimated variance of the estimator of the difference under the standard approach.*

We conclude this section showing how to evaluate a policy and how to compare two policies using the proposed approach on randomized-by-action data.

Randomizing by Action — Evaluating a Single Policy Using the Proposed Approach
We use $W_i = s$ to denote assignment to the policy that assigns action s, regardless of the customers' covariates. We want to evaluate targeting policy 1 without implementing it, using the outcomes from the experimental conditions pertaining to the actions.

Our goal (the estimand) is the population-level measure

$$y_1 = \frac{1}{N} \sum_{i=1}^{N} Y_i(1),$$

which can also be written as

$$y_1 = \frac{\sum_{s,t} \sum_{i \in g(1:s,2:t)} Y_i(s)}{\sum_{s,t} N^{1:s,2:t}} = \frac{\sum_{s,t} N^{1:s,2:t} \cdot y_1^{1:s,2:t}}{N},$$

where $y_1^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} Y_i(s)}{N^{1:s,2:t}}$.

We estimate y_1 by calculating the outcome in each segment and then aggregating across segments:

$$\hat{y}_1 = \frac{\sum_{s,t} N^{1:s,2:t} \cdot \hat{y}_1^{1:s,2:t}}{N},$$

with

$$\hat{y}_1^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} \mathbb{1}_{W_i = s} Y_i^{obs}}{N_s^{1:s,2:t}},$$

where $N_s^{1:s,2:t}$ is the number of customers in segment $g(1:s,2:t)$ that are randomly assigned to receive policy $W_i = s$. This estimator is at the heart of the proposed alternative experimental design: to evaluate a policy on a segment of customers to which the policy recommends a specific action, we look at customers within the segment that were hit with that specific action.

It is straightforward to establish that $\hat{y}_1$ is an unbiased estimator of y_1. We can also write expressions for the variance of the estimator $\hat{y}_1$ within each segment:

$$\text{Var}_W\left(\hat{y}_1^{1:s,2:t}\right) = \frac{\left(S_1^{1:s,2:t}\right)^2}{N_s^{1:s,2:t}} \left(1 - \frac{N_s^{1:s,2:t}}{N^{1:s,2:t}}\right),$$

where $\left(S_1^{1:s,2:t}\right)^2$ is given by Equation (4). We estimate $\text{Var}_W\left(\hat{y}_1^{1:s,2:t}\right)$ by

$$\widehat{\text{Var}}_W\left(\hat{y}_1^{1:s,2:t}\right) = \frac{\left(s_1^{1:s,2:t}\right)^2}{N_s^{1:s,2:t}} \left(1 - \frac{N_s^{1:s,2:t}}{N^{1:s,2:t}}\right),$$

where

$$\left(s_1^{1:s,2:t}\right)^2 = \frac{1}{N_s^{1:s,2:t} - 1} \sum_{i \in g(1:s,2:t), W_i = s} \left(Y_i^{obs} - \hat{y}_1^{1:s,2:t}\right)^2.$$

Overall, the variance of estimator $\hat{y}_1$ across all segments is

$$\mathrm{Var}_W(\hat{y}_1) = \sum_{s,t} \left(\frac{N^{1:s,2:t}}{N} \right)^2 \mathrm{Var}_W \left(\hat{y}_1^{1:s,2:t} \right),$$

which we estimate by

$$\widehat{\mathrm{Var}}_W(\hat{y}_1) = \sum_{s,t} \left(\frac{N^{1:s,2:t}}{N} \right)^2 \widehat{\mathrm{Var}}_W \left(\hat{y}_1^{1:s,2:t} \right).$$

Next we describe an estimator for estimating the difference between two policies.

Randomizing by Action — Comparing Two Policies Using the Proposed Approach
Our target estimand is

$$y_{1-2} = y_1 - y_2 = \frac{1}{N} \sum_{i=1}^{N} \Big(Y_i(1) - Y_i(2) \Big),$$

which, accounting for the segmentation of the customers according to the recommended actions, can also be written as

$$\frac{\sum_{s,t} \sum_{i \in g(1:s,2:t)} \Big(Y_i(s) - Y_i(t) \Big)}{\sum_{s,t} N^{1:s,2:t}} = \frac{\sum_{\substack{s,t \\ s \neq t}} \sum_{i \in g(1:s,2:t)} \Big(Y_i(s) - Y_i(t) \Big)}{\sum_{s,t} N^{1:s,2:t}} = \frac{\sum_{\substack{s,t \\ s \neq t}} N^{1:s,2:t} \left(y_1^{1:s,2:t} - y_2^{1:s,2:t} \right)}{N},$$

where, for $s \neq t$,

$$y_1^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} Y_i(s)}{N^{1:s,2:t}}, \qquad y_2^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} Y_i(t)}{N^{1:s,2:t}}.$$

We estimate y_{1-2} with the unbiased estimator

$$\hat{y}_{1-2} = \frac{\sum_{\substack{s,t \\ s \neq t}} N^{1:s,2:t} \left(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t} \right)}{N},$$

where, for $s \neq t$,

$$\hat{y}_1^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} \mathbb{1}_{W_i = s} Y_i^{obs}}{N_s^{1:s,2:t}}, \qquad \hat{y}_2^{1:s,2:t} = \frac{\sum_{i \in g(1:s,2:t)} \mathbb{1}_{W_i = t} Y_i^{obs}}{N_t^{1:s,2:t}}.$$

In segments for which the two policies recommend the same action, the true difference in the outcome between the two policies is zero. In segments for which the two policies recommend different actions ($s \neq t$), the variance of the difference between the two policies is given by

$$\mathrm{Var}_W \left(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t} \right) = \frac{\left(S_1^{1:s,2:t} \right)^2}{N_s^{1:s,2:t}} + \frac{\left(S_2^{1:s,2:t} \right)^2}{N_t^{1:s,2:t}} - \frac{\left(S_{1,2}^{1:s,2:t} \right)^2}{N^{1:s,2:t}},$$

where variances $\left(S_1^{1:s,2:t}\right)^2, \left(S_2^{1:s,2:t}\right)^2, \left(S_{1,2}^{1:s,2:t}\right)^2$ are given by Equations (4) and (7).

We estimate $\left(S_1^{1:s,2:t}\right)^2$ and $\left(S_2^{1:s,2:t}\right)^2$ with

$$\left(s_1^{1:s,2:t}\right)^2 = \frac{1}{N_s^{1:s,2:t} - 1} \sum_{i \in g(1:s,2:t), W_i = s} \left(Y_i^{obs} - \hat{y}_1^{1:s,2:t}\right)^2$$

and

$$\left(s_2^{1:s,2:t}\right)^2 = \frac{1}{N_t^{1:s,2:t} - 1} \sum_{i \in g(1:s,2:t), W_i = t} \left(Y_i^{obs} - \hat{y}_2^{1:s,2:t}\right)^2,$$

respectively. The term $\left(S_{1,2}^{1:s,2:t}\right)^2$ is in general impossible to estimate empirically because we never observe the outcome of both actions s, t for the same customer. We use the Neyman variance estimator

$$\mathrm{Var}_{g(1:s,2:t)}^{\mathrm{Neyman}} = \frac{\left(s_1^{1:s,2:t}\right)^2}{N_s^{1:s,2:t}} + \frac{\left(s_2^{1:s,2:t}\right)^2}{N_t^{1:s,2:t}},$$

which is generally upwardly biased.

Overall, the variance of estimator $\hat{y}_{1-2}$ across all segments is

$$\mathrm{Var}_W\left(\hat{y}_{1-2}\right) = \sum_{\substack{s,t \\ s \neq t}} \left(\frac{N^{1:s,2:t}}{N}\right)^2 \mathrm{Var}_W\left(\hat{y}_1^{1:s,2:t} - \hat{y}_2^{1:s,2:t}\right),$$

which we estimate with

$$\mathrm{Var}_{1-2}^{\mathrm{Neyman}} = \sum_{\substack{s,t \\ s \neq t}} \left(\frac{N^{1:s,2:t}}{N}\right)^2 \mathrm{Var}_{g(1:s,2:t)}^{\mathrm{Neyman}}.$$

This is generally upwardly biased and, as a result, confidence intervals will be conservative.

B.1 Evaluating a Single Policy Using the Standard Approach - the Variance

We have $\mathrm{Var}_W(\hat{y}_1) = \mathbb{E}_W[\hat{y}_1^2] - (\mathbb{E}_W[\hat{y}_1])^2$. For the first term, we write

$$\mathbb{E}_W[\hat{y}_1^2] = \mathbb{E}_W\left[\left(\frac{1}{N_1}\sum_{i=1}^{N} \mathbb{1}_{W_i=1} Y_i^{obs}\right)^2\right].$$

The squared terms of the sum are of the form

$$\mathbb{E}_W\left[\left(\mathbb{1}_{W_i=1} Y_i^{obs}\right)^2\right] = \frac{N_1}{N} \cdot Y_i^2(1),$$

while the cross terms are of the form

$$\begin{aligned}
\mathbb{E}_W\left[\left(\mathbb{1}_{W_i=1}Y_i^{obs}\right)\left(\mathbb{1}_{W_j=1}Y_j^{obs}\right)\right] &= \Pr\left(W_i = W_j = 1\right)\cdot Y_i(1)\cdot Y_j(1) \\
&= \frac{\binom{N-2}{N_1-2}}{\binom{N}{N_1}}\cdot Y_i(1)\cdot Y_j(1) \\
&= \frac{N_1(N_1-1)}{N(N-1)}\cdot Y_i(1)\cdot Y_j(1),
\end{aligned}$$

for $i \neq j$. Overall, we can write

$$\mathbb{E}_W[\hat{y}_1^2] = \frac{1}{N_1^2}\left(\frac{N_1}{N}\sum_{i=1}^{N} Y_i^2(1) + \frac{N_1(N_1-1)}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(1)\right)$$

and

$$\left(\mathbb{E}_W[\hat{y}_1]\right)^2 = \left(\frac{1}{N}\sum_{i=1}^{N} Y_i(1)\right)^2 = \frac{1}{N^2}\left(\sum_{i=1}^{N} Y_i^2(1) + \sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(1)\right).$$

Combining the two, we have that

$$\mathrm{Var}_W(\hat{y}_1) = \left(\frac{1}{N_1\cdot N} - \frac{1}{N^2}\right)\sum_{i=1}^{N} Y_i^2(1) - \left(\frac{1}{N^2} - \frac{N_1-1}{N_1\cdot N\cdot(N-1)}\right)\sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(1) \quad (8)$$

Starting from the definition of S_1^2, we write

$$\begin{aligned}
S_1^2 &= \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i(1) - \frac{1}{N}\sum_{j=1}^{N} Y_j(1)\right)^2 \\
&= \frac{1}{N-1}\left(\left(\sum_{i=1}^{N} Y_i^2(1)\right) - N\left(\frac{\sum_{i=1}^{N} Y_i(1)}{N}\right)^2\right) \\
&= \left(\frac{1}{N-1} - \frac{1}{(N-1)N}\right)\sum_{i=1}^{N} Y_i^2(1) - \frac{1}{(N-1)N}\sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(1). \quad (9)
\end{aligned}$$

We compare expressions (8) and (9). We show that the coefficients for the terms $\sum_{i=1}^{N} Y_i^2(1)$, $\sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(1)$ in Equation (8) equal the coefficients in Equation (9), scaled by $\frac{1}{N_1}\left(1 - \frac{N_1}{N}\right)$.

Indeed, we have

$$\begin{aligned}
\frac{1}{N_1}\left(1 - \frac{N_1}{N}\right)\left(\frac{1}{N-1} - \frac{1}{(N-1)N}\right) &= \left(\frac{1}{N_1} - \frac{1}{N}\right)\left(\frac{1}{N-1} - \frac{1}{(N-1)N}\right) \\
&= \frac{N-N_1}{N_1\cdot N^2} \\
&= \frac{1}{N_1\cdot N} - \frac{1}{N^2}.
\end{aligned}$$

We also have that
$$\frac{1}{N_1}\left(1-\frac{N_1}{N}\right)\frac{1}{(N-1)N}=\frac{N-N_1}{N_1\cdot N^2\cdot(N-1)},$$
which is equal to the respective coefficient of Equation (8), because
$$\frac{1}{N^2}-\frac{N_1-1}{N_1\cdot N\cdot(N-1)}=\frac{N_1(N-1)-(N_1-1)N}{N_1\cdot N^2\cdot(N-1)}=\frac{N-N_1}{N_1\cdot N^2\cdot(N-1)}.$$

We have thus shown that $\mathrm{Var}_W(\hat{y}_1)=\frac{S_1^2}{N_1}\left(1-\frac{N_1}{N}\right)$.

B.2 Evaluating a Single Policy Using the Proposed Approach - Unbiasedness

For $s\neq t$, we have

$$\begin{aligned}
\mathbb{E}_W\left[\hat{y}_1^{1:s,2:t}\right] &= \frac{1}{N_1^{1:s,2:t}}\sum_{i\in g(1:s,2:t)}\mathbb{E}_W[\mathbb{1}_{W_i=1}]\cdot Y_i^{obs}\\
&= \frac{1}{N_1^{1:s,2:t}}\frac{N_1^{1:s,2:t}}{N^{1:s,2:t}}\sum_{i\in g(1:s,2:t)}Y_i(s)\\
&= \frac{1}{N^{1:s,2:t}}\sum_{i\in g(1:s,2:t)}Y_i(s)\\
&= y_1^{1:s,2:t}.
\end{aligned}$$

We also have

$$\begin{aligned}
\mathbb{E}_W\left[\hat{y}_1^{1:s,2:s}\right] &= \frac{1}{N_1^{1:s,2:s}+N_2^{1:s,2:s}}\sum_{i\in g(1:s,2:s)}\mathbb{E}_W\left[\mathbb{1}_{W_i=1}+\mathbb{1}_{W_i=2}\right]\cdot Y_i^{obs}\\
&= \frac{1}{N_1^{1:s,2:s}+N_2^{1:s,2:s}}\left(\frac{N_1^{1:s,2:s}}{N^{1:s,2:s}}+\frac{N_2^{1:s,2:s}}{N^{1:s,2:s}}\right)\sum_{i\in g(1:s,2:s)}Y_i(s)\\
&= \frac{1}{N^{1:s,2:s}}\sum_{i\in g(1:s,2:s)}Y_i(s)\\
&= y_1^{1:s,2:s}.
\end{aligned}$$

We can therefore write

$$\mathbb{E}_W[\hat{y}_1]=\frac{\sum_{s,t}N^{1:s,2:t}\cdot\mathbb{E}_W\left[\hat{y}_1^{1:s,2:t}\right]}{N}=\frac{\sum_{s,t}N^{1:s,2:t}\cdot y_1^{1:s,2:t}}{N}=y_1.$$

B.3 Evaluating a Single Policy Using the Proposed Approach - the Variance

Showing that, for $s \neq t$,

$$\mathrm{Var}_W\left(\hat{y}_1^{1:s,2:t}\right) = \frac{\left(S_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}}\left(1 - \frac{N_1^{1:s,2:t}}{N^{1:s,2:t}}\right),$$

follows the exact same steps as the derivation in Appendix B.1, restricted to the segment $g(1:s,2:t)$.

The derivation for the case of a segment in which the two policies recommend the same action is similar: a calculation shows that

$$\begin{aligned}
\mathrm{Var}_W\left(\hat{y}_1^{1:s,2:s}\right) &= \left(\frac{1}{\left(N_1^{1:s,2:s}+N_2^{1:s,2:s}\right)N^{1:s,2:s}} - \frac{1}{\left(N^{1:s,2:s}\right)^2}\right)\sum_{i\in g(1:s,2:s)} Y_i^2(s) \\
&\quad - \left(\frac{1}{\left(N^{1:s,2:s}\right)^2} - \frac{N_1^{1:s,2:s}+N_2^{1:s,2:s}-1}{\left(N_1^{1:s,2:s}+N_2^{1:s,2:s}\right)N^{1:s,2:s}\left(N^{1:s,2:s}-1\right)}\right) \\
&\quad \cdot \sum_{i\in g(1:s,2:s)}\sum_{j\neq i} Y_i(s)\cdot Y_j(s) \\
&= \frac{\left(S^{1:s,2:s}\right)^2}{N_1^{1:s,2:s}+N_2^{1:s,2:s}}\left(1 - \frac{N_1^{1:s,2:s}+N_2^{1:s,2:s}}{N^{1:s,2:s}}\right).
\end{aligned}$$

B.4 Comparing Two Policies Using the Standard Approach - the Variance

We start by writing

$$\mathrm{Var}_W\left(\hat{y}_{1-2}\right) = \mathrm{Var}_W\left(\hat{y}_1 - \hat{y}_2\right) = \mathrm{Var}_W\left(\hat{y}_1\right) + \mathrm{Var}_W\left(\hat{y}_2\right) - 2\cdot \mathrm{Cov}_W\left(\hat{y}_1,\hat{y}_2\right).$$

We calculate the covariance. We have

$$\mathbb{E}_W\left[\hat{y}_1\cdot\hat{y}_2\right] = \mathbb{E}_W\left[\frac{1}{N_1}\sum_{i=1}^{N}\mathbb{1}_{W_i=1}Y_i^{obs}\cdot\frac{1}{N_2}\sum_{i=1}^{N}\mathbb{1}_{W_i=2}Y_i^{obs}\right].$$

We first calculate, for $i \neq j$,

$$\begin{aligned}
\mathbb{E}_W\left[\left(\mathbb{1}_{W_i=1}Y_i^{obs}\right)\cdot\left(\mathbb{1}_{W_j=2}Y_i^{obs}\right)\right] &= \Pr\left(W_i=1, W_j=2\right)\cdot Y_i(1)\cdot Y_j(1) \\
&= \frac{\binom{N-2}{N_1-1,N_2-1,N-N_1-N_2}}{\binom{N}{N_1,N_2,N-N_1-N_2}}\cdot Y_i(1)\cdot Y_j(1) \\
&= \frac{N_1N_2}{N(N-1)}\cdot Y_i(1)\cdot Y_j(1),
\end{aligned}$$

therefore

$$\mathbb{E}_W\left[\hat{y}_1 \cdot \hat{y}_2\right] = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq 1} Y_i(1) \cdot Y_j(1).$$

We also calculate

$$\begin{aligned}
\mathbb{E}_W\left[\hat{y}_1\right] \cdot \mathbb{E}_W\left[\hat{y}_2\right] &= y_1 \cdot y_2 \\
&= \frac{1}{N} \sum_{i=1}^{N} Y_i(1) \cdot \frac{1}{N} \sum_{i=1}^{N} Y_i(2) \\
&= \frac{1}{N^2} \left[\sum_{i=1}^{N} Y_i(1) \cdot Y_i(2) + \sum_{i=1}^{N} \sum_{j \neq i} Y_i(1) \cdot Y_j(2) \right].
\end{aligned}$$

Therefore the covariance is

$$\begin{aligned}
\mathrm{Cov}_W\left(\hat{y}_1, \hat{y}_2\right) &= \mathbb{E}_W\left[\hat{y}_1 \cdot \hat{y}_2\right] - \mathbb{E}_W\left[\hat{y}_1\right] \cdot \mathbb{E}_W\left[\hat{y}_2\right] \\
&= -\frac{1}{N^2} \sum_{i=1}^{N} Y_i(1) \cdot Y_i(2) + \frac{1}{N^2(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} Y_i(1) \cdot Y_j(2). \qquad (10)
\end{aligned}$$

Overall, we have

$$\begin{aligned}
\mathrm{Var}_W\left(\hat{y}_{1-2}\right) &= \mathrm{Var}_W\left(\hat{y}_1\right) + \mathrm{Var}_W\left(\hat{y}_2\right) - 2 \cdot \mathrm{Cov}_W\left(\hat{y}_1, \hat{y}_2\right) \\
&= \left(\frac{1}{N_1 N} - \frac{1}{N^2}\right) \sum_{i=1}^{N} Y_i^2(1) + \left(\frac{1}{N_2 N} - \frac{1}{N^2}\right) \sum_{i=1}^{N} Y_i^2(2) \\
&\quad - \left(\frac{1}{N^2} - \frac{N_1 - 1}{N_1 N(N-1)}\right) \sum_{i=1}^{N} \sum_{j \neq i} Y_i(1) \cdot Y_j(1) \\
&\quad - \left(\frac{1}{N^2} - \frac{N_2 - 1}{N_2 N(N-1)}\right) \sum_{i=1}^{N} \sum_{j \neq i} Y_i(2) \cdot Y_j(2) \\
&\quad + \frac{2}{N^2} \sum_{i=1}^{N} Y_i(1) \cdot Y_i(2) - \frac{2}{N^2(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} Y_i(1) \cdot Y_j(2), \qquad (11)
\end{aligned}$$

where the second equality follows from Equations (8) and (10).

We write

$$\begin{aligned}
\frac{S_1^2}{N_1} &= \frac{1}{N_1(N-1)} \sum_{i=1}^{N} \left(Y_i(1) - \frac{1}{N} \sum_{j=1}^{N} Y_j(1) \right)^2 \\
&= \frac{1}{N_1(N-1)} \left[\left(1 - \frac{1}{N}\right) \sum_{i=1}^{N} Y_i^2(1) - \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i} Y_i(1) \cdot Y_j(1) \right], \qquad (12)
\end{aligned}$$

$$
\begin{aligned}
\frac{S_2^2}{N_2} &= \frac{1}{N_2(N-1)}\sum_{i=1}^{N}\left(Y_i(2)-\frac{1}{N}\sum_{j=1}^{N}Y_j(2)\right)^2 \\
&= \frac{1}{N_2(N-1)}\left[\left(1-\frac{1}{N}\right)\sum_{i=1}^{N}Y_i^2(2)-\frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i}Y_i(2)\cdot Y_j(2)\right],
\end{aligned} \tag{13}
$$

and

$$
\begin{aligned}
\frac{S_{1,2}^2}{N} &= \frac{1}{N(N-1)}\sum_{i=1}^{N}\left(Y_i(1)-Y_i(2)-\frac{1}{N}\sum_{j=1}^{N}\Big(Y_j(1)-Y_j(2)\Big)\right)^2 \\
&= \frac{1}{N(N-1)}\left[\sum_{i=1}^{N}Y_i^2(1)+\sum_{i=1}^{N}Y_i^2(2)+\frac{1}{N}\left(\sum_{i=1}^{N}Y_i(1)\right)^2+\frac{1}{N}\left(\sum_{i=1}^{N}Y_i(2)\right)^2\right. \\
&\qquad -2\cdot\sum_{i=1}^{N}Y_i(1)\cdot Y_i(2)-2\cdot\sum_{i=1}^{N}Y_i(1)\frac{\sum_{j=1}^{N}Y_j(1)}{N} \\
&\qquad +2\cdot\sum_{i=1}^{N}Y_i(1)\frac{\sum_{j=1}^{N}Y_j(2)}{N}+2\cdot\sum_{i=1}^{N}Y_i(2)\frac{\sum_{j=1}^{N}Y_j(1)}{N} \\
&\qquad \left.-2\cdot\sum_{i=1}^{N}Y_i(2)\frac{\sum_{j=1}^{N}Y_j(2)}{N}-\frac{2}{N}\left(\sum_{i=1}^{N}Y_i(1)\right)\left(\sum_{i=1}^{N}Y_i(2)\right)\right] \\
&= \frac{1}{N(N-1)}\left[\left(1-\frac{1}{N}\right)\left(\sum_{i=1}^{N}Y_i^2(1)+\sum_{i=1}^{N}Y_i^2(2)\right)\right. \\
&\qquad -\frac{1}{N}\left(\sum_{i=1}^{N}\sum_{j\neq i}Y_i(1)\cdot Y_j(1)+\sum_{i=1}^{N}\sum_{j\neq i}Y_i(2)\cdot Y_j(2)\right) \\
&\qquad \left.-\left(2-\frac{2}{N}\right)\sum_{i=1}^{N}Y_i(1)\cdot Y_i(2)+\frac{2}{N}\sum_{i=1}^{N}\sum_{j\neq i}Y_i(1)\cdot Y_j(2)\right].
\end{aligned} \tag{14}
$$

We compare the terms of Equation (11) with the terms of

$$
\frac{S_1^2}{N_1}+\frac{S_2^2}{N_2}-\frac{S_{1,2}^2}{N}, \tag{15}
$$

using Equations (12), (13), and (14).

We first compare the terms with $\sum_{i=1}^{N}Y_i^2(1)$. Expression (15) has this term multiplied by $\frac{1}{N_1(N-1)}\left(1-\frac{1}{N}\right)-\frac{1}{N(N-1)}\left(1-\frac{1}{N}\right)$. This is equal to $\frac{1}{N_1N}-\frac{1}{N^2}$, which is the multiplier of term $\sum_{i=1}^{N}Y_i^2(1)$ in expression (11).

The comparison for the terms with $\sum_{i=1}^{N}Y_i^2(2)$ is similar.

We now compare the terms with $\sum_{i=1}^{N}\sum_{j\neq i}Y_i(1)\cdot Y_j(1)$. Expression (11) has this term multiplied by $\frac{N_1-1}{N_1N(N-1)}-\frac{1}{N^2}$. This is equal to $-\frac{1}{N_1(N-1)}\frac{1}{N}+\frac{1}{N(N-1)}\frac{1}{N}$, which is the multiplier of term $\sum_{i=1}^{N}\sum_{j\neq i}Y_i(1)\cdot Y_j(1)$ in expression (15).

The comparison for the terms with $\sum_{i=1}^{N}\sum_{j\neq i} Y_i(2)\cdot Y_j(2)$ is similar.

We now compare the terms with $\sum_{i=1}^{N} Y_i(1)\cdot Y_i(2)$. Expression (15) has this term multiplied by $\frac{1}{N(N-1)}\left(2-\frac{2}{N}\right)$. This is equal to $\frac{2}{N^2}$, which is the multiplier of term $\sum_{i=1}^{N} Y_i(1)\cdot Y_i(2)$ in expression (11).

We finally compare the terms with $\sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(2)$. In both expressions (11) and (15), term $\sum_{i=1}^{N}\sum_{j\neq i} Y_i(1)\cdot Y_j(2)$ is multiplied with $-\frac{2}{N^2(N-1)}$.

We have thus shown that

$$\mathrm{Var}_W\left(\hat{y}_{1-2}\right) = \frac{S_1^2}{N_1} + \frac{S_2^2}{N_2} - \frac{S_{1,2}^2}{N}.$$

B.5 Comparing the Efficiency of the Standard and the Proposed Approaches

We first compare the estimated variances under each approach for single policy evaluation.

We start with the estimated variance under the proposed approach:

$$\begin{aligned}&\sum_{\substack{s,t\\ s\neq t}}\left(\frac{N^{1:s,2:t}}{N}\right)^2\frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}}\left(1-\frac{N_1^{1:s,2:t}}{N^{1:s,2:t}}\right)\\&\quad+\sum_{s}\left(\frac{N^{1:s,2:s}}{N}\right)^2\frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s}+N_2^{1:s,2:s}}\cdot\left(1-\frac{N_1^{1:s,2:s}+N_2^{1:s,2:s}}{N^{1:s,2:s}}\right),\end{aligned}\tag{16}$$

which we want to compare to the estimated variance under the standard approach,

$$\frac{s_1^2}{N_1}\left(1-\frac{N_1}{N}\right).\tag{17}$$

We assume a balanced experimental design, such that the proportion of customers assigned to a policy within a segment is the same as the proportion of customers assigned to that policy in any other segment, i.e.,

$$\frac{N_1^{1:s,2:t}}{N^{1:s,2:t}} = \frac{N_1}{N}$$

for all s, t.

For $s\neq t$, we can write

$$\begin{aligned}\left(\frac{N^{1:s,2:t}}{N}\right)^2\frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} &= \left(\frac{N_1^{1:s,2:t}}{N_1}\right)^2\frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}}\\ &= \frac{1}{N_1}\frac{N_1^{1:s,2:t}}{N_1}\left(s_1^{1:s,2:t}\right)^2,\end{aligned}$$

and for the segments where the two policies recommend the same action, we can write

$$\begin{aligned}
\left(\frac{N^{1:s2:s}}{N}\right)^2 \frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s}+N_2^{1:s,2:s}} &= \left(\frac{N_1^{1:s,2:s}}{N_1}\right)^2 \frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s}+N_2^{1:s,2:s}} \\
&< \left(\frac{N_1^{1:s,2:s}}{N_1}\right)^2 \frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s}} \\
&= \frac{1}{N_1}\frac{N_1^{1:s,2:s}}{N_1}\left(s^{1:s,2:s}\right)^2.
\end{aligned}$$

We have $1-\frac{N_1^{1:s,2:t}}{N_{1:s,2:t}} = 1-\frac{N_1}{N}$, and $1-\frac{N_1^{1:s,2:s}+N_2^{1:s,2:s}}{N^{1:s,2:s}} = 1-\frac{N_1+N_2}{N} < 1-\frac{N_1}{N}$. We can now compare the expressions (16) and (17), after eliminating the adjustment for the size of the super-population.

$$\begin{aligned}
&\sum_{\substack{s,t\\ s\neq t}}\left(\frac{N^{1:s,2:t}}{N}\right)^2 \frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} + \sum_s \left(\frac{N^{1:s,2:s}}{N}\right)^2 \frac{\left(s^{1:s,2:s}\right)^2}{N_1^{1:s,2:s}+N_2^{1:s,2:s}} \\
&\quad < \frac{1}{N_1}\left(\sum_{\substack{s,t\\ s\neq t}}\frac{N_1^{1:s,2:t}}{N_1}\left(s_1^{1:s,2:t}\right)^2 + \sum_s \frac{N_1^{1:s,2:s}}{N_1}\left(s^{1:s,2:s}\right)^2\right).
\end{aligned}$$

If it holds that $\left(s_1^{1:s,2:t}\right)^2 \leq s_1^2$ for all actions s,t with $s\neq t$, and $\left(s^{1:s,2:s}\right)^2 \leq s_1^2$ for all actions s, then the previous expression is bounded above by

$$\frac{1}{N_1}\sum_{s,t}\frac{N_1^{1:s,2:t}}{N_1}s_1^2 = \frac{s_1^2}{N_1},$$

showing that the proposed method strictly improves the variance over the standard method.

We now compare the estimated variances under the standard and the proposed approach for comparing two policies. We start with the estimated variance under the proposed approach

$$\sum_{\substack{s,t\\ s\neq t}}\left(\frac{N^{1:s,2:t}}{N}\right)^2\left(\frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} + \frac{\left(s_2^{1:s,2:t}\right)^2}{N_2^{1:s,2:t}}\right),$$

which can be written as

$$\begin{aligned}
&\textstyle\sum_{\substack{s,t\\ s\neq t}}\left(\frac{N_1^{1:s,2:t}}{N_1}\right)^2\frac{\left(s_1^{1:s,2:t}\right)^2}{N_1^{1:s,2:t}} + \sum_{\substack{s,t\\ s\neq t}}\left(\frac{N_2^{1:s,2:t}}{N_2}\right)^2\frac{\left(s_2^{1:s,2:t}\right)^2}{N_2^{1:s,:t}} \\
&= \textstyle\frac{1}{N_1}\sum_{\substack{s,t\\ s\neq t}}\frac{N_1^{1:s,2:t}}{N_1}\left(s_1^{1:s,2:t}\right)^2 + \frac{1}{N_2}\sum_{\substack{s,t\\ s\neq t}}\frac{N_2^{1:s,2:t}}{N_2}\left(s_2^{1:s,2:t}\right)^2,
\end{aligned}$$

where we have used the assumption of a balanced experimental design:

$$\frac{N_1^{1:s,2:t}}{N_1} = \frac{N_2^{1:s,2:t}}{N_2} = \frac{N^{1:s,2:t}}{N}.$$

If it holds that $\left(s_1^{1:s,2:t}\right)^2 \leq s_1^2$ and $\left(s_2^{1:s,2:t}\right)^2 \leq s_2^2$ for all actions s, t with $s \neq t$, then the expression for the variance is strictly less than

$$\frac{1}{N_1} \sum_{s,t} \frac{N_1^{1:s,2:t}}{N_1} s_1^2 + \frac{1}{N_2} \sum_{s,t} \frac{N_2^{1:s,2:t}}{N_2} s_2^2 = \frac{s_1^2}{N_1} + \frac{s_2^2}{N_2},$$

which is the estimated variance under the standard approach.

C Targeting Variables

Table 3: Definition of Targeting Variables

Variable	**Definition**
Age	Age of head of household
Home Value	Estimated home value
Income	Estimated household income
Single Family	A binary flag indicating whether the home is a single family home
Multi-Family	A binary flag indicating whether the home is a multi-family home
Distance	Distance to nearest store for this retailer
Comp. Distance	Distance to nearest competitors' store
Penetration Rate	% of households in zip code that are members
3yr Response	Average response rate to mailings to this zip code over the last 3 years
F Flag	Binary flag indicating whether the retailer considers the zip code "far" from its closest store
M Flag	Binary flag indicating whether the retailer considers the zip code a "medium" distance from its closest store
Past Paids	The proportion of households in the zip code that were previously paid members
Trialists	The proportion of households in the zip code that have been identified as households who repeatedly sign up for trial memberships

D Double Machine Learning Approach for Comparing Two Targeting Policies Using Counterfactual Policy Logging

We present the algorithm for the double machine learning approach we propose in Section 5 and implement in Section 7.

Data:
We have N observations. X_i are the covariates for observation i. P_i is a binary indicator such that

- $P_i = 0$ for all observations i that received an action recommended by policy P_1;
- $P_i = 1$ for all observations i that received an action recommended by policy P_2.

$P_1(i)$ and $P_2(i)$ are actions recommended for observation i by policies P_1 and P_2, respectively. S and K are the number of cross-validation iterations, and folds at each iteration, respectively.

Result:
Estimate the mean and standard error of the difference in the performance of targeting policies P_1 and P_2.

begin

Define segments S_1, ..., S_J by grouping observations $1, \ldots, N$ using policies P_1 and P_2:

1. Observations within a segment S_j share the action assignments $P_1(S_j)$ and $P_2(S_j)$.
2. Calculate the segment sizes $\delta_{S_j} = \frac{|S_j|}{N}$

foreach S_j **do**

if $P_1(S_j) = P_2(S_j)$ **then**

$\hat{\theta}^{S_j} = 0, \hat{\sigma}^{S_j} = 0$

else

for $s = 1, \ldots, S$ **do**

Calculate $w_{P_1} = \frac{|\{i \in S_j : P_i = 0\}|}{|S_j|}$ and $w_{P_2} = \frac{|\{i \in S_j : P_i = 1\}|}{|S_j|}$

Randomly split a sample S_j into K equal folds

for $k = 1, \ldots, K$ **do**

Use (K-1) folds to estimate $\hat{g}(P_1, X)$ and $\hat{g}(P_2, X)$:

1. Fit a model $y = \hat{g}(P_1, X)$ using {obs in (K-1) folds for which $P_i = 0$}
2. Fit a model $y = \hat{g}(P_2, X)$ using {obs in (K-1) folds for which $P_i = 1$}

foreach obs. i in fold k **do**

$$\hat{\phi}_i = \left(\hat{g}(P_1, X_i) + \frac{(y_i - \hat{g}(P_1, X_i))(1 - P_i)}{w_{P_1}}\right) - \left(\hat{g}(P_2, X_i) + \frac{(y_i - \hat{g}(P_2, X_i))P_i}{w_{P_2}}\right)$$

end

end

$$\hat{\theta}_s^{S_j} = \frac{1}{|S_j|} \sum_{i=1}^{|S_j|} \hat{\phi}_i, \quad \left(\hat{\sigma}_s^{S_j}\right)^2 = \frac{1}{|S_j|} \sum_{i=1}^{|S_j|} \left(\hat{\phi}_i - \hat{\theta}_s^{S_j}\right)^2$$

end

$$\hat{\theta}^{S_j} = \frac{1}{S} \sum_{s=1}^{S} \hat{\theta}_s^{S_j}, \quad \left(\hat{\sigma}^{S_j}\right)^2 = \frac{1}{S} \sum_{s=1}^{S} \left(\left(\hat{\sigma}_s^{S_j}\right)^2 + \left(\hat{\theta}_s^{S_j} - \hat{\theta}^{S_j}\right)^2\right)$$

end

end

$$\hat{\theta} = \sum_{j=1}^{J} \delta_{S_j} \hat{\theta}^{S_j}, \quad s.e.(\hat{\theta}) = \sqrt{\frac{\sum_{j=1}^{J} \delta_{S_j} \left(\hat{\sigma}^{S_j}\right)^2}{N}}$$

end

E STZ Study: Performance Differences and Standard Errors

Table 4: Carrier Route Level Results

		Simple OLS	OLS with Covariates	DML OLS	DML Lasso
Randomized-by-Action	No Segmentation	$0.049 ($0.101) [$0.101]	$0.064 ($0.086) [$0.086]	$0.065 ($0.086)	$0.046 ($0.087)
	Partial Segmentation	$0.023 ($0.044) [$0.045]	$0.037 ($0.039) [$0.039]	$0.045 ($0.040)	$0.044 ($0.039)
Randomized-by-Policy	No Segmentation	$0.100 ($0.107) [$0.107]	$0.042 ($0.090) [$0.090]	$0.042 ($0.090)	$0.016 ($0.091)
	Partial Segmentation	$0.091 ($0.047) [$0.047]	$0.109 ($0.041) [$0.040]	$0.112 ($0.041)	$0.108 ($0.041)

Note: The table reports the estimated difference in the performance of Policy 1 and Policy 2 in the STZ study under different estimation approaches. The data is aggregated to the carrier route level (the unit of analysis is a carrier route). Standard errors are in round parentheses. Robust standard errors using the Eicker-Huber-White adjustment are in square brackets. To preserve confidentiality, the profits are multiplied by a common random number.

Table 5: Individual Customer Level Results

		Simple OLS	OLS with Covariates	DML OLS	DML Lasso
Randomized-by-Action	No Segmentation	$0.038 ($0.052) [$0.052]	$0.028 ($0.052) [$0.052]	$0.028 ($0.052)	$0.022 ($0.052)
	Partial Segmentation	$0.011 ($0.024) [$0.024]	$0.016 ($0.024) [$0.024]	$0.018 ($0.024)	$0.020 ($0.024)
Randomized-by-Policy	No Segmentation	$0.089 ($0.055) [$0.055]	$0.126 ($0.055) [$0.055]	$0.127 ($0.055)	$0.100 ($0.055)
	Partial Segmentation	$0.064 ($0.027) [$0.027]	$0.106 ($0.027) [$0.028]	$0.108 ($0.027)	$0.113 ($0.027)

Note: The table reports the estimated difference in the performance of Policy 1 and Policy 2 in the STZ study under different estimation approaches. The unit of analysis is an individual customer. Standard errors are in round parentheses. Robust standard errors using the Eicker-Huber-White adjustment are in square brackets. To preserve confidentiality, the profits are multiplied by a common random number.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge (2014), "Finite population causal standard errors." Working Paper 20325, National Bureau of Economic Research, URL `http://www.nber.org/papers/w20325`.

Athey, S. and G.W. Imbens (2017), "The econometrics of randomized experiments." *Handbook of Economic Field Experiments*, 1, 73 – 140, URL `http://www.sciencedirect.com/science/article/pii/S2214658X16300174`. Handbook of Field Experiments.

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

Neyman, Jerzy (1934), "On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection." *Journal of the Royal Statistical Society*, 97, 558–625, URL `http://www.jstor.org/stable/2342192`.

Simester, D., A. Timoshenko, and S. I. Zoumpoulis (2018), "Targeting prospective customers: Robustness of machine learning methods to typical data challenges." *Working Paper*.