

# Unit Selection Based on Counterfactual Logic

Ang Li and Judea Pearl

Cognitive Systems Laboratory, University of California, Los Angeles

{angli, judea}@cs.ucla.edu

## Abstract

The unit selection problem aims to identify a set of individuals who are most likely to exhibit a desired mode of behavior, which is defined in counterfactual terms. A typical example is that of selecting individuals who would respond one way if encouraged and a different way if not encouraged. Unlike previous works on this problem, which rely on ad-hoc heuristics, we approach this problem formally, using counterfactual logic, to properly capture the nature of the desired behavior. This formalism enables us to derive an informative selection criterion which integrates experimental and observational data. We demonstrate the superiority of this criterion over A/B-test-based approaches.

## 1 Introduction

The problem of selecting individuals with a desired response pattern is encountered in many areas of industry, marketing, and health science. For example, in customer relationship management (CRM) [Berson *et al.*, 1999; Lejeune, 2001; Hung *et al.*, 2006; Tsai and Lu, 2009], it is of interest to predict which customers are about to churn but are likely to change their minds if enticed toward retention. The cost associated with such programs compels management to limit enticement to customers who are most likely to exhibit the behavior of interest. In online advertising [Yan *et al.*, 2009; Bottou *et al.*, 2013; Li *et al.*, 2014; Sun *et al.*, 2015], as another example, companies are interested in identifying users who would click on an advertisement if and only if the said advertisement is highlighted. The difficulty in identifying these users stems from the fact that the desired response pattern is not observed directly but rather is defined counterfactually in terms of what the individual would do under hypothetical unrealized conditions. For example, when we observe that a user has clicked on a highlighted advertisement, we do not know whether they would click on that same advertisement if it were not highlighted.

It is useful to classify individual behavior into four response types, labeled complier, always-taker, never-taker, and defier [Angrist *et al.*, 1996; Balke and Pearl, 1997]. Compliers are individuals who would respond positively if encouraged and negatively if not encouraged. Always-takers are in-

dividuals who always respond positively whether or not they are encouraged. Never-takers are individuals who always respond negatively whether or not they are encouraged. Defiers are individuals who would respond negatively if encouraged and positively if not encouraged.

A typical objective of the unit selection problem is to select individuals with those characteristics that maximize the percentage of compliers since compliers represent the effectiveness of the encouragement.

A common solution that is explored in the literature is an A/B-test-based approach, where a controlled experiment is performed and the result is used as a criterion for selection. Specifically, users are randomly split into two groups called control and treatment. Users in the control group are served un-highlighted advertisements, whereas those in the treatment group are served highlighted advertisements. Then, those characteristics that resulted in a higher difference between the two groups are used as predictors for the benefit of selection.

Departing from the prevailing literature, we will treat the unit selection problem using the structural causal model (SCM) [Pearl, 2009], which accounts for the counterfactual nature of the desired behavior, and in which a large body of theoretical work has been established [Galles and Pearl, 1998; Halpern, 2000].

The unit selection problem entails two sub-problems, evaluation and search. The evaluation problem is to devise an estimable objective function that, if optimized over the set of observed characteristics  $\mathcal{C}$  (available for each individual), would ensure an optimal counterfactual behavior for the selected group. The search task is to devise a search algorithm to select individuals based both on their observed characteristics and the objective function devised above. This task is nontrivial due to the large number of characteristics available for each individual and the sparsity of data available in each cell (of characteristics).

In this study, we focus on the evaluation sub-problem. In section 4, we define the counterfactual expression that should serve as the objective function for selection. This expression consists of the probabilities of causation, such as the probability of necessity-and-sufficiency (PNS), which was studied in [Pearl, 1999; Tian and Pearl, 2000; Kuroki and Cai, 2011].

Next, we provide two conditions under which the prevailing heuristic used in the literature can become optimal. Our

analysis shows that a selection criterion based on an A/B test can be made optimal (by fine tuning) under conditions of *monotonicity* or *gain equality*, to be defined formally in section 5.1.

In the general case, however, the counterfactual criterion is not identifiable. In the Main Results, we derive a tight bound for this criterion, based on experimental and observational data and use the midpoint of this bound as a selection criterion. Finally, by simulation, we demonstrate that sets of individuals selected by the derived criterion yield greater overall benefit than those selected by standard methods.

## 2 Motivating Example

Consider a mobile carrier that wants to identify customers likely to discontinue their services within the next quarter based on customer characteristics (the company management has access to user data, such as income, age, usage, and monthly payments). The carrier will then offer these customers a special renewal deal to dissuade them from discontinuing their services and to increase their service renewal rate. These offers provide considerable discounts to the customers, and the management prefers that these offers be made only to those customers who would continue their service if and only if they receive the offer. Note that some customers may discontinue service if and only if offered the renewal discount. Reasons for this could include being reminded that they are paying for a service they no longer want, feeling that discounts cheapen the service, reflecting on how much they are paying, being turned off by the promotional wording, or being annoyed by the process to claim the discount.

A typical aim is to select a subset of individuals with the characteristics  $c$  (a concrete instantiation of all characteristics) that maximizes the percentage of compliers and minimizes the percentages of defiers, always-takers, and never-takers among the selected customers (compliers are the customers who would continue the service if they received special offers and would not otherwise; defiers are the customers who would continue the service if they received no special offers and would not otherwise; always-takers are the customers who would continue the service whether or not they received special offers; never-takers are the customers who would not continue the service whether or not they received special offers).

### 2.1 Related Work and Our Contributions

There are two main approaches for handling the unit selection problem, as described extensively in books, articles, and software packages.

The first approach relies on A/B testing and statistical analysis [Sundar *et al.*, 1998; Blumenthal *et al.*, 2001; Winer, 2001; Resnick *et al.*, 2006; Lewis and Reiley, 2014]. Specifically, an experiment is conducted on a randomized controlled group of individuals. Then, the desired individuals (with concrete instantiation  $c$  of all characteristics) are identified by maximizing the difference in the probability  $P(\text{positive response}|c, \text{encouraged}) - P(\text{positive response}|c, \text{not encouraged})$ . However, the counterfactual nature of the desired behavior is not handled properly. A linear combination of  $P(\text{positive response}|c, \text{encouraged})$  and  $P(\text{positive response}|c, \text{not encouraged})$  does not maximize the percentage of compliers and minimize the percentages of defiers, always-takers, and never-takers among the selected individuals, because the first term comprises compliers and always-takers and the second term comprises always-takers and defiers.

The second approach is machine-learning based. Hung [Hung *et al.*, 2006] summarized and compared the most popular methods for churn prediction, including the regression, decision tree, and neural network methods. Using these approaches, a model is constructed using historical data to identify which customers are likely to discontinue their services. Then, the carrier offers a special renewal deal to the customers identified by the model as most likely to churn. However, an analysis of the set of customers who have accepted the special deal (hence, not churned) does not immediately reveal the customers who would have continued their services anyway and the customers who renewed their services only because of the special deal. Of course, we can run another A/B test, however, this leads to the same scenario as that encountered when employing the above statistical approach.

Our treatment differs fundamentally from those of previous studies by appealing to SCM, which is more robust and less prone to model misspecifications. First, the SCM model makes no assumptions about the data-generating process. Second, in most cases, the experimental data can be evaluated in terms of the observational data when a causal graph is available. In such a case, observational data alone is sufficient for this approach. Third, and most importantly, the SCM properly accounts for the counterfactual nature of the desired behavior.

Our treatment differs fundamentally from those of previous studies by appealing to SCM, which is more robust and less prone to model misspecifications. First, the SCM model makes no assumptions about the data-generating process. Second, in most cases, the experimental data can be evaluated in terms of the observational data when a causal graph is available. In such a case, observational data alone is sufficient for this approach. Third, and most importantly, the SCM properly accounts for the counterfactual nature of the desired behavior.

## 3 Preliminaries

In this section, we review the counterfactual logic [Galles and Pearl, 1998; Halpern, 2000; Pearl, 2009] associated with Pearl’s SCM, which is used in the remainder of this paper. Readers who are familiar with SCM may want to skip this section.

### 3.1 Counterfactual Logic

The basic counterfactual statement associated with model  $M$  is denoted by  $Y_x(u) = y$ , and stands for: “ $Y$  would be  $y$  had  $X$  been  $x$  in unit  $U = u$ .”. Let  $M_x$  denote a modified version of  $M$ , with the equation(s) of set  $X$  replaced by  $X = x$  (i.e., all edges that go into  $X$  have been removed). Then, the formal definition of the counterfactual  $Y_x(u)$  is as follows:

$$Y_x(u) \triangleq Y_{M_x}(u) \tag{1}$$

In words, the counterfactual  $Y_x(u)$  in model  $M$  is defined as the solution of  $Y$  in the “modified” submodel  $M_x$ . In [Galles and Pearl, 1998; Halpern, 2000], a complete axiomatization of structural counterfactuals, embracing both recursive and nonrecursive models, is given.

Equation (1) implies that the distribution  $P(u)$  induces a well-defined probability for the counterfactual event  $Y_x = y$ , written as  $P(Y_x = y)$ , which is equal to the probability that a random unit  $u$  would satisfy the equation  $Y_x(u) = y$ . Therefore, the probability of the event “ $Y$  would be  $y$  had  $X$  been

$x$ ”,  $P(Y_x = y)$ , is well-defined and  $P(Y_x = y) = P(Y = y|do(X = x))$ .  $P(Y = y|do(X = x))$  can be interpreted as experimental data [Pearl, 1995]. With the same reasoning, the SCM model assigns a probability to every counterfactual or combination of counterfactuals that are defined using the variables in SCM.

Using the above formal language for the counterfactual expression, all events involving a counterfactual scenario can be well defined, because the event represented by the subscript does not actually occur. For example,  $P(Y_x = y|X = x')$  defines the probability of the event “ $Y$  would be  $y$  had  $X$  been  $x$  if we observed  $X = x'$ ” (note that  $x$  and  $x'$  are counterfactual scenarios),  $P(Y_x = y, Y_{x'} = y')$  defines the probability of the event “ $Y$  would be  $y$  had  $X$  been  $x$  and  $Y$  would be  $y'$  had  $X$  been  $x'$ ” (note that  $x$  and  $x'$  is a counterfactual scenario;  $y$  and  $y'$  is a counterfactual scenario), and  $P(Y_x = y|X = x', Y = y')$  defines the probability of the event “ $Y$  would be  $y$  had  $X$  been  $x$ , if we observed  $X = x'$  and  $Y = y'$ ”.

For simplicity purposes, in the rest of the paper, we use  $y_x$  to denote the event  $Y_x = y$ ,  $y_{x'}$  to denote the event  $Y_{x'} = y$ ,  $y'_x$  to denote the event  $Y_x = y'$ , and  $y'_{x'}$  to denote the event  $Y_{x'} = y'$ .

## 4 Counterfactual Formulation of the Unit Selection Problem

Our objective is to find a set of characteristics  $c$  that maximizes the benefit associated with the resulting mixture of compliers, defiers, always-takers, and never-takers. Suppose the benefit of selecting a complier is  $\beta$ , the benefit of selecting an always-taker is  $\gamma$ , the benefit of selecting a never-taker is  $\theta$ , and the benefit of selecting a defier is  $\delta$ . Our objective, then, should be to find  $c$  that maximizes the following expression:

$$\begin{aligned} \operatorname{argmax}_c & \beta P(\text{complier}|c) + \gamma P(\text{always-taker}|c) + \\ & + \theta P(\text{never-taker}|c) + \delta P(\text{defier}|c) \end{aligned}$$

Suppose  $A = a$  denotes encouragement is received and  $A = a'$  otherwise;  $R = r$  denotes a positive response and  $R = r'$  otherwise. The objective function that maximizes the benefit on average of the selected individuals can be formulated as follows:

$$\begin{aligned} \operatorname{argmax}_c & \beta P(r_a, r'_a|c) + \gamma P(r_a, r'_a|c) + \\ & + \theta P(r'_a, r'_a|c) + \delta P(r'_a, r'_a|c) \end{aligned} \quad (2)$$

Most importantly, this objective function can be bounded using observational and experimental data, as will be demonstrated in the following section.

## 5 Main Results

In this section, we demonstrate how an explicit solution of the unit selection problem can be derived using the benefit function with observational and experimental data via the following theorem.

**Theorem 1.** *The benefit function  $f(\beta, \gamma, \theta, \delta) = \beta P(y_x, y'_{x'}|z) + \gamma P(y_x, y_{x'}|z) + \theta P(y'_{x'}, y'_{x'}|z) + \delta P(y_{x'}, y'_{x'}|z)$  is bounded as follows:*

$$\max\{p_1, p_2, p_3, p_4\} \leq f \leq \min\{p_5, p_6, p_7, p_8\} \text{ if } \sigma < 0 \quad (3)$$

$$\max\{p_5, p_6, p_7, p_8\} \leq f \leq \min\{p_1, p_2, p_3, p_4\} \text{ if } \sigma > 0 \quad (4)$$

where  $\sigma, p_1, \dots, p_8$  are given by,

$$\begin{aligned} \sigma &= \beta - \gamma - \theta + \delta \\ p_1 &= (\beta - \theta)P(y_x|z) + \delta P(y_{x'}|z) + \theta P(y'_{x'}|z) \\ p_2 &= \gamma P(y_x|z) + \delta P(y_{x'}|z) + (\beta - \gamma)P(y'_{x'}|z) \\ p_3 &= (\gamma - \delta)P(y_x|z) + \delta P(y_{x'}|z) + \theta P(y'_{x'}|z) \\ &+ (\beta - \gamma - \theta + \delta)[P(y, x|z) + P(y', x'|z)] \\ p_4 &= (\beta - \theta)P(y_x|z) - (\beta - \gamma - \theta)P(y_{x'}|z) + \theta P(y'_{x'}|z) \\ &+ (\beta - \gamma - \theta + \delta)[P(y, x'|z) + P(y', x|z)] \\ p_5 &= (\gamma - \delta)P(y_x|z) + \delta P(y_{x'}|z) + \theta P(y'_{x'}|z) \\ p_6 &= (\beta - \theta)P(y_x|z) - (\beta - \gamma - \theta)P(y_{x'}|z) + \theta P(y'_{x'}|z) \\ p_7 &= (\gamma - \delta)P(y_x|z) - (\beta - \gamma - \theta)P(y_{x'}|z) \\ &+ \theta P(y'_{x'}|z) + (\beta - \gamma - \theta + \delta)P(y|z) \\ p_8 &= (\beta - \theta)P(y_x|z) + \delta P(y_{x'}|z) + \theta P(y'_{x'}|z) \\ &- (\beta - \gamma - \theta + \delta)P(y|z) \end{aligned}$$

*Proof.* See Appendix <sup>1</sup>. □

### 5.1 Identifiability under Additional Assumptions

We will now show that equation 2 can be evaluated precisely from pure experimental data under either of two conditions, monotonicity (definition 2) and gain equality (definition 3). Moreover, both conditions lead to the same result.

Monotonicity expresses the assumption that a change from  $X = \text{false}$  to  $X = \text{true}$  cannot, under any circumstance make  $Y$  change from true to false [Tian and Pearl, 2000]. In epidemiology, this assumption is often expressed as “no prevention,” that is, no individual in the population can be helped by exposure to the risk factor.

**Definition 2.** (Monotonicity) *A Variable  $Y$  is said to be monotonic relative to variable  $X$  in a causal model  $M$  iff*

$$y'_x \wedge y_{x'} = \text{false}$$

Gain equality states that the benefit of selecting a complier and a defier is the same as the benefit of selecting an always-taker and a never-taker (i.e.,  $\beta + \delta = \gamma + \theta$ ).

**Definition 3.** (Gain Equality) *The benefit of selecting a complier ( $\beta$ ), an always-taker ( $\gamma$ ), a never-taker ( $\theta$ ), and a defier ( $\delta$ ) is said to satisfy gain equality iff*

$$\beta + \delta = \gamma + \theta$$

**Theorem 4.** *Given that  $Y$  is monotonic relative to  $X$  or that  $(\beta, \gamma, \theta, \delta)$  satisfies gain equality, the benefit function  $f(\beta, \gamma, \theta, \delta)$  is given by*

$$\begin{aligned} & f(\beta, \gamma, \theta, \delta) \\ &= \beta P(y_x, y'_{x'}|z) + \gamma P(y_x, y_{x'}|z) + \\ & \quad \theta P(y'_{x'}, y'_{x'}|z) + \delta P(y_{x'}, y'_{x'}|z) \\ &= (\beta - \theta)P(y_x|z) + (\gamma - \beta)P(y_{x'}|z) + \theta \end{aligned}$$

*Proof.* See Appendix <sup>1</sup>. □

<sup>1</sup>The detail proof is at the appendix in [https://ftp.cs.ucla.edu/pub/stat\\_ser/r488.pdf](https://ftp.cs.ucla.edu/pub/stat_ser/r488.pdf)

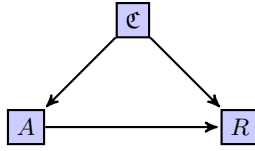


Figure 1: Causal graph for the customer selection model.

Note that for the special case where the perceived benefit is proportional to the final number of customers in the system, the A/B heuristic will maximize the expression  $(H - D) \times P(y_x|z) - H \times P(y_{x'}|z)$ , where  $H$  is the unit profit per remaining customer and  $D$  is the discount offered. This case corresponds to the following parameters in our notation  $\beta = H - D, \gamma = -D, \theta = 0, \delta = -H$ . And indeed, theorem 4 implies an identical benefit function  $f = (H - D) \times P(y_x|z) - H \times P(y_{x'}|z)$ . In other words, the A/B heuristic is optimal for this special case. For slightly more elaborate combinations of  $(\beta, \gamma, \theta, \delta)$ , however, theorem 4 dictates a benefit function that is not captured by A/B heuristics.

Without either monotonicity or gain equality, we can only obtain bounds for the objective function. However, the next section demonstrates (by simulation) that taking the midpoint of these bounds as a criterion results in a greatly improved selection of individuals.

## 6 Simulation Results

In this section, we present two simulated examples, one to demonstrate that the midpoints of the bounds of the objective function given by equations (3, 4) are adequate for selecting the desired individuals, and one to demonstrate the case that satisfies gain equality. In addition, we illustrate that, in the traditional A/B-test-based statistical approach, the selected individuals are different from those selected using the proposed approach and have a lower benefit on average.

### 6.1 Example in Churn Management

First, we consider the motivating example. Let  $A = a$  denote the event that a customer receives the special deal and  $A = a'$  denote the event that a customer receives no special deal. Let  $R = r$  denote the event that a customer continues the services and  $R = r'$  denote the event that a customer discontinues the services. Let  $\mathcal{C}$  (the set of variables) denote the characteristics of the customer (e.g., income, age, usage, and monthly payments). Figure 1 depicts the customer selection model.

The management estimates that the benefit of selecting a complier is \$100 as the profit is \$140 but the discount is \$40, the benefit of selecting an always-taker is -\$60 as the customer would continue the service anyway, so the company loses the value of the discount and an extra cost (because the always-taker may require additional discounts in the future), the benefit of selecting a never-taker is 0 as the cost of issuing the discount is negligible, and the benefit of selecting a defier is -\$140 as we lose customer due to the special offer.

Suppose we have two groups of customers, group 1 with characteristics  $c_1$  and group 2 with characteristics  $c_2$ . In addition, we have prior information that  $P(r|c_1) = 0.7$  and

		$do(a)$	$do(a')$
Group 1	$r$	262	175
	$r'$	88	175
Group 2	$r$	87	52
	$r'$	263	298

Table 1: Results of a simulated study for churn management.

	$f_1$	$f_2$	$f_3$
Group 1	\$25	\$4.86	-\$2.63
Group 2	\$10	\$4.06	\$3.09

Table 2: Results of the three objective functions based on the data from the simulated study.

$P(r|c_2) = 0.3$ . We randomly select 700 customers from each group and offer the special renewal deal to 350 customers in each group. Table 1 summarizes the results.

Let us compare three selection strategies, each using a different objective function. The first is based on the simple A/B test, that is:

$$\begin{aligned} Obj_1 &= \operatorname{argmax}_c f_1(c) \\ &= \operatorname{argmax}_c 100 \cdot P(r|c, do(a)) - 100 \cdot P(r|c, do(a')). \end{aligned}$$

The second is based on weighted A/B test approach, where

$$\begin{aligned} Obj_2 &= \operatorname{argmax}_c f_2(c) \\ &= \operatorname{argmax}_c 100 \cdot P(r|c, do(a)) - 140 \cdot P(r|c, do(a')). \end{aligned}$$

The third is based on the analysis of this paper, where equation (2) yields

$$\begin{aligned} Obj_3 &= \operatorname{argmax}_c f_3(c) \\ &= \operatorname{argmax}_c 100 \cdot P(r_a, r'_a|c) + (-60) \cdot P(r_a, r'_a|c) \\ &\quad + 0 \cdot P(r'_a, r'_a|c) + (-140) \cdot P(r'_a, r'_a|c). \end{aligned}$$

Then, we enter the data in table 1 into the objective functions of groups 1 and 2. Table 2 summarizes the results (note that we use the midpoint of the bound as the selection criterion for  $Obj_3$  and  $P(r_a|c) = P(r|c, do(a))$ ). The proposed approach selected group 2; however, the first and second objective functions selected group 1 as the desired individuals.

An informer with access to the fractions of compliers, always-takers, never-takers, and defiers in both groups (as summarized in table 3) would easily conclude that the A/B-test-based approach had reached a wrong conclusion. (Note that we will never know these numbers in reality) In detail, the expected benefit of selecting an individual in group 1 is  $100 \times 0.3 - 60 \times 0.45 + 0 \times 0.2 - 140 \times 0.05 = -\$4$ , which means offering the special deal to group 1 would reduce the profit; the expected benefit of selecting an individual in group 2 is  $100 \times 0.2 - 60 \times 0.05 + 0 \times 0.65 - 140 \times 0.1 = \$3$ . Thus, the management should only offer the special deal to group 2.

Furthermore, the plot in figure 2 depicts the benefit of group 1 from objective functions as a function of  $\delta$  ( $\beta, \gamma$ , and  $\theta$  are fixed), with each curve representing an objective function. The first two objective functions are the most common

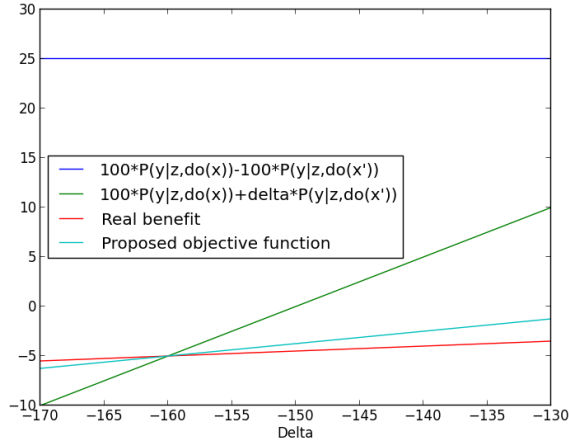


Figure 2: Benefit calculated by objective functions versus  $\delta$  of group 1 in the churn management model.

	Complier	Always-taker	Never-taker	Defier
Group 1	30%	45%	20%	5%
Group 2	20%	5%	65%	10%

Table 3: Percentages of four response types in each group for churn management.

heuristics in the A/B-test-based approach. The third objective function is the real expected benefit. The last objective function is the midpoint of the bounds for the proposed objective function. We see that the midpoint of the bounds for the proposed objective function is the closest one to the real benefit.

## 6.2 Example in Online Advertisement

### Task 1

A search engine company management wants to decide whether it is worth recommending an advertisement to a group of users, so as to increase user satisfaction. The management estimates that the satisfaction of recommending an advertisement to a complier is 2 degrees as users would gain new information that they needed, the satisfaction of recommending the advertisement to an always-taker is 1 degree as users got a shortcut to the advertisement, the satisfaction of recommending the advertisement to a never-taker is -1 degrees as users got unnecessary information, and the satisfaction of recommending the advertisement to a defier is -2 degrees as the recommendation would prevent users to get needed information (compliers are the users who would click on the advertisement if the advertisement is recommended and would not otherwise; always-takers are the users who would click on the advertisement whether or not the advertisement is recommended; never-takers are the users who would not click on the advertisement whether or not the advertisement is recommended; defiers are the users who would click on the advertisement if the advertisement is not recommended and would not otherwise).

	$do(a)$	$do(a')$
$r$	140	175
$r'$	210	175

Table 4: Results of simulated study for advertisement recommendation.

Complier	Always-taker	Never-taker	Defier
30%	10%	20%	40%

Table 5: Percentages of four response types for advertisement recommendation.

Let  $A = a$  denote the event that the given advertisement is recommended and  $A = a'$  denote the event that the given advertisement is not recommended. Let  $R = r$  denote the event that the user clicks on the advertisement and  $R = r'$  denote the event that the user does not click on the advertisement.

Since no other data is available about the users, the management decides to conduct a randomized experiment and measure the degree to which the recommendation increases users' click rate. The study involved 700 randomly selected users of whom 350 were recommended the advertisement. Table 4 summarizes the results.

A simple A/B test approach concluded that recommending the advertisement to this group of users would increase the user satisfaction because (satisfaction with recommendation)  $\times P(r|do(a)) -$  (satisfaction without recommendation)  $\times P(r|do(a')) = 2 \times 0.4 - 1 \times 0.5 = 0.3$ .

However, an informer with access to the fractions of compliers, always-takers, never-takers, and defiers in the group (as summarized in table 5, note that we will never know these numbers in reality since there is no monotonicity) claimed that the simple A/B test approach had reached a wrong conclusion. According to the company's assessment, the expected satisfaction per customer of recommending the advertisement to this group is  $2 \times 0.3 + 1 \times 0.1 - 1 \times 0.2 - 2 \times 0.4 = -0.3$ . This analysis shows that recommending the advertisement to this group of users would reduce the satisfaction. This is because only 30% of users are compliers and 10% of users are always-takers; therefore, a lot of advertisements are recommended on never-takers and defiers, which makes the recommendation reduce satisfaction.

In contrast, looking at the benefit parameter  $(2, 1, -1, -2)$ , we see that it satisfies gain equality, which means that we can obtain the true average satisfaction, despite the fact that we cannot determine the fraction of individuals in each response type. Accordingly, applying the benefit function of theorem 4, we obtain that the expected satisfaction per user of recommending the advertisement to the group is  $3 \times P(r|do(a)) - 1 \times P(r|do(a')) - 1 = 3 \times 0.4 - 1 \times 0.5 - 1 = -0.3$ , which is precisely the satisfaction computed knowing the type distribution. This implies that the company should NOT recommend the advertisement to the group.

### Task 2

Let us look at two groups, c1 and c2. A study by the same company was conducted with 1400 randomly selected users

		$do(a)$	$do(a')$
Group 1	$r$	140	88
	$r'$	210	262
Group 2	$r$	192	210
	$r'$	158	140

Table 6: Results of a simulated study for advertisement recommendation.

	Complier	Always-taker	Never-taker	Defier
Group 1	20%	20%	55%	5%
Group 2	30%	25%	10%	35%

Table 7: Percentages of four response types in each group for advertisement recommendation.

(700 in each group) where the advertisement was recommended to 700 of those users (350 in each group). Table 6 summarizes the results.

A simple A/B test approach concluded that recommending the advertisement to both group of customers would increase the satisfaction because (satisfaction with recommendation)  $\times P(r|do(a), c_1)$  - (satisfaction without recommendation)  $\times P(r|do(a'), c_1) = 2 \times 0.4 - 1 \times 0.25 = 0.55$ , and (satisfaction with recommendation)  $\times P(r|do(a), c_2)$  - (satisfaction without recommendation)  $\times P(r|do(a'), c_2) = 2 \times 0.55 - 1 \times 0.6 = 0.5$ .

However, an informer with access to the fractions of compliers, always-takers, never-takers, and defiers in both groups (as summarized in table 7, note that we will never know these numbers in reality) claimed that the simple A/B test approach had reached a wrong conclusion. In detail, the expected satisfaction per user of recommending the advertisement to group 1 is  $2 \times 0.2 + 1 \times 0.2 - 1 \times 0.55 - 2 \times 0.05 = -0.05$ , which means offering coupons to this group of users would reduce the satisfaction; the expected satisfaction per user of recommending the advertisement to group 2 is  $2 \times 0.3 + 1 \times 0.25 - 1 \times 0.1 - 2 \times 0.35 = 0.05$ , which means recommending the advertisement to this group of users would increase the satisfaction. Thus, the company should only recommend the advertisement to group 2.

In contrast, looking at the benefit parameter  $(2, 1, -1, -2)$ , we see that it satisfies gain equality, which means that we can obtain the true average satisfaction despite the fact that we cannot determine the fraction of individuals in each response type. Accordingly, applying the benefit function of theorem 4, we obtain that the expected satisfaction per user of recommending the advertisement to the group 1 is  $3 \times P(r|do(a)) - 1 \times P(r|do(a')) - 1 = 3 \times 0.4 - 1 \times 0.25 - 1 = -0.05$ , and the expected satisfaction per user of recommending the advertisement to the group 2 is  $3 \times P(y|do(x)) - 1 \times P(y|do(x')) - 1 = 3 \times 0.55 - 1 \times 0.6 - 1 = 0.05$ , which is precisely the satisfaction computed knowing the type distribution. This implies that the company should NOT recommend the advertisement to group 1.

### 6.3 Discussion

In this section, we discuss additional features of the counterfactual-logic-based approach. First, as discussed in section 4, the objective function properly accounts for the counterfactual nature of the desired behavior. Theorem 4 provides theoretical assurance that the A/B-test-based approach can be made optimal under certain conditions. However, the first simulated experimental example demonstrates that this approach selected individuals with a lower expected benefit when the cost-benefit structure is ignored. Even though the proposed objective function is, in general, not identifiable and cannot be used in selection, we demonstrated in the previous section that the mid-point of the tight bounds in theorem 1 is adequate for selecting the desired individuals.

Second, given a causal graph and a set of observed variables that satisfies the backdoor criterion [Pearl, 1993], theorem 1 can be applied using purely the observational data via adjustment formula [Pearl, 1995].

Third, the proposed approach could be used to evaluate machine learning models as well as to generate labels for machine learning models. The accuracy of such a machine learning model would be higher because it would consider the counterfactual scenarios.

Fourth, theorem 4 provides a way for identifying the weight coefficients in the extensively used statistical approach when the additional assumption is satisfied.

Finally, the proposed approach is applicable universally to any application in which the manager can assess the benefits associated with selecting a unit in each of the four types of units. Theorem 1 ensures that for any benefits input, we obtain the desired output. The input is not determined by the model, but by the manager who can use the algorithm for any combination of inputs.

## 7 Conclusions

We demonstrated the advantages of the SCM framework in addressing the unit selection problem. We defined an objective function for selection that properly accounts for the counterfactual nature of the desired behavior. We derived tight bounds (theorem 1) to ensure that the objective function can be evaluated using experimental and observational data. We further identified via theorem 4 the conditions under which the standard A/B test heuristic used in the literature can become optimal. In summary, we have analyzed and demonstrated what can be gained by exploiting causal knowledge, when solving the unit selection problem.

## Acknowledgments

This research was supported in parts by grants from International Business Machines Corporation (IBM) [#A1771928], National Science Foundation [#IIS-1527490 and #IIS-1704932], and Office of Naval Research [#N00014-17-S-B001]. The authors thank Manabu Kuroki and Scott Mueller for all valuable comments.

## References

- [Angrist *et al.*, 1996] Joshua D Angrist, Guido W Imbens, and Donald B Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996.
- [Balke and Pearl, 1997] Alexander Balke and Judea Pearl. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92(439):1171–1176, 1997.
- [Berson *et al.*, 1999] Alex Berson, Stephen Smith, and Kurt Thearling. *Building data mining applications for CRM*. McGraw-Hill Professional, 1999.
- [Blumenthal *et al.*, 2001] Marsha Blumenthal, Charles Christian, Joel Slemrod, and Matthew G Smith. Do normative appeals affect tax compliance? evidence from a controlled experiment in minnesota. *National Tax Journal*, pages 125–138, 2001.
- [Bottou *et al.*, 2013] Léon Bottou, Jonas Peters, Joaquin Quiñonero-Candela, Denis X Charles, D Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. Counterfactual reasoning and learning systems: The example of computational advertising. *The Journal of Machine Learning Research*, 14(1):3207–3260, 2013.
- [Galles and Pearl, 1998] David Galles and Judea Pearl. An axiomatic characterization of causal counterfactuals. *Foundations of Science*, 3(1):151–182, 1998.
- [Halpern, 2000] Joseph Y Halpern. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research*, 12:317–337, 2000.
- [Hung *et al.*, 2006] Shin-Yuan Hung, David C Yen, and Hsiu-Yu Wang. Applying data mining to telecom churn management. *Expert Systems with Applications*, 31(3):515–524, 2006.
- [Kuroki and Cai, 2011] Manabu Kuroki and Zhihong Cai. Statistical analysis of probabilities of causation using covariate information. *Scandinavian Journal of Statistics*, 38(3):564–577, 2011.
- [Lejeune, 2001] Miguel APM Lejeune. Measuring the impact of data mining on churn management. *Internet Research*, 11(5):375–387, 2001.
- [Lewis and Reiley, 2014] Randall A Lewis and David H Reiley. Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on yahoo! *Quantitative Marketing and Economics*, 12(3):235–266, 2014.
- [Li *et al.*, 2014] Lihong Li, Shunbao Chen, Jim Kleban, and Ankur Gupta. Counterfactual estimation and optimization of click metrics for search engines. *arXiv preprint arXiv:1403.1891*, 2014.
- [Pearl, 1993] J Pearl. Aspects of graphical models connected with causality. *Proceedings of the 49th Session of the International Statistical Institute, Italy*, pages 399–401, 1993.
- [Pearl, 1995] Judea Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 1995.
- [Pearl, 1999] Judea Pearl. Probabilities of causation: three counterfactual interpretations and their identification. *Synthese*, 121(1-2):93–149, 1999.
- [Pearl, 2009] Judea Pearl. *Causality*. Cambridge university press, 2009.
- [Resnick *et al.*, 2006] Paul Resnick, Richard Zeckhauser, John Swanson, and Kate Lockwood. The value of reputation on ebay: A controlled experiment. *Experimental economics*, 9(2):79–101, 2006.
- [Sun *et al.*, 2015] Wei Sun, Pengyuan Wang, Dawei Yin, Jian Yang, and Yi Chang. Causal inference via sparse additive models with application to online advertising. In *AAAI*, pages 297–303, 2015.
- [Sundar *et al.*, 1998] S Shyam Sundar, Sunetra Narayan, Rafael Obregon, and Charu Uppal. Does web advertising work? memory for print vs. online media. *Journalism & Mass Communication Quarterly*, 75(4):822–835, 1998.
- [Tian and Pearl, 2000] Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. *Annals of Mathematics and Artificial Intelligence*, 28(1-4):287–313, 2000.
- [Tsai and Lu, 2009] Chih-Fong Tsai and Yu-Hsin Lu. Customer churn prediction by hybrid neural networks. *Expert Systems with Applications*, 36(10):12547–12553, 2009.
- [Winer, 2001] Russell S Winer. A framework for customer relationship management. *California management review*, 43(4):89–105, 2001.
- [Yan *et al.*, 2009] Jun Yan, Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen. How much can behavioral targeting help online advertising? In *Proceedings of the 18th international conference on World wide web*, pages 261–270. ACM, 2009.