An Evaluation of the Causal Effect of a Promotion Campaign on Software Usage

Febriany Lete
Maysen Pagan

April 2024

1 Introduction

2 Exploratory Data Analysis

Before performing any causal analysis, we first explore the data by looking at a correlation matrix of our quantitative variables.

```
##                 IT.Spend Employee.Count PC.Count Size
## IT.Spend            1.00           0.02     0.02 0.88
## Employee.Count      0.02           1.00     0.95 0.02
## PC.Count            0.02           0.95     1.00 0.03
## Size                0.88           0.02     0.03 1.00
```

From this matrix, we can see a strong linear association of 0.95 between the number of PCs a customer has and the number of employees a customer has. Similarly, there is a large correlation of 0.88 between the size of the customer, given by their yearly total revenue, and the dollars spent on IT purchases by the customer. As a result, we choose to omit the dollars spend on IT purchases and the number of PCs the customer has from further analysis.

3 Methods and Analysis

3.1 Assumptions

We then consider the three main identification assumptions to identify our average treatment effect: Positivity, Consistency, and Conditional Exchangeability.

3.1.1 Positivity

For the Positivity assumption, we observed the covariate balance plots for each of our covariates, comparing the distribution of the values between the treatment and control. These plots can be viewed in the Appendix. All plots show the pattern of an equal distribution of covariate values between those customers who received the discount and those who did not. For example, the first covariate balance plot for the employee count variable displays a skewed right distribution with the count ranging from 0 to about 500 employees both for customers who received a discount and customers who did not receive a discount. This demonstrates that for given values of each of the covariates, the probability of receiving the treatment and control are both non-zero. Therefore,

$$0 < p(A = a|X) < 1, \ a \in \{0, 1\}$$

3.1.2 Consistency

We will also perform our causal analysis under the Consistency assumption which states that a customer's potential outcome under the observed treatment assignment is equal to the observed outcome for that customer. Therefore,

If $A_i = a$, then $Y_i = Y_i^{(a)}$

3.1.3 Conditional Exchangeability

For the Conditional Exchangeability assumption, we thought of possible variables that could be a confounder of the treatment-outcome relation. We believe that the size of the customer could be this confounder. The startup company might tend to give discounts to larger companies either to keep their loyalty or prevent them from purchasing from their competitors. In this sense, the size of the customer directly affects the treatment assignment. Additionally, the size of the customer may also directly affect the outcome because, for example, larger customers will most likely purchase more software than smaller customers and therefore bring in more revenue for the company.

However, this size of the customer confounder can be observed in our data set through variables such as the number of employees the customer has or the binary variable that indicates if the customer has global offices. Larger businesses tend to have more employees as well as global offices. Since this confounder can be observed in our data set, we continue with the assumption of Conditional Exchangeability which states
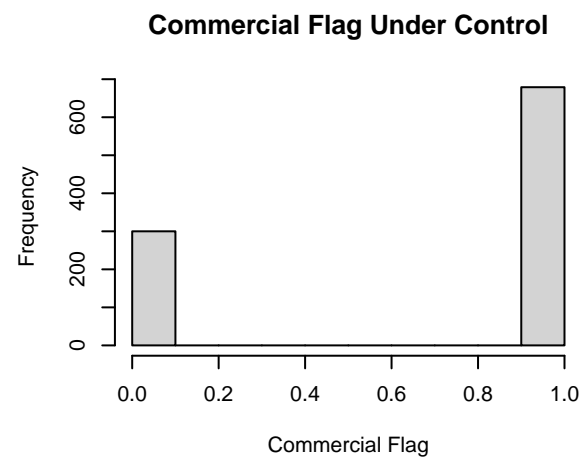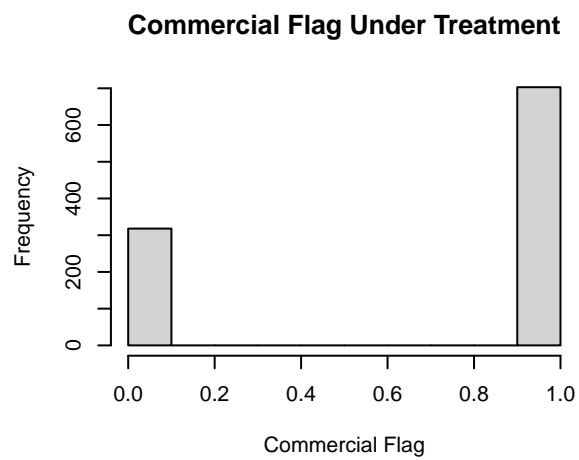
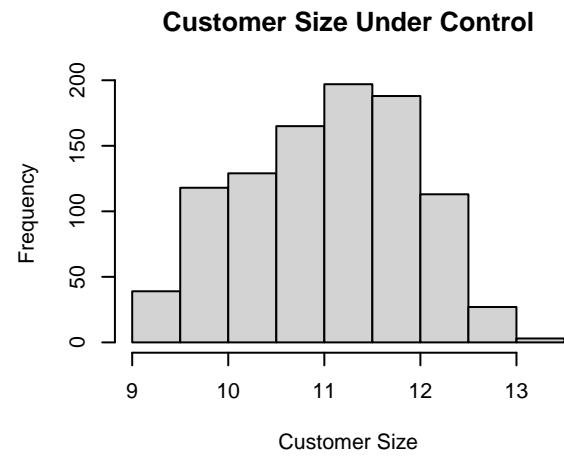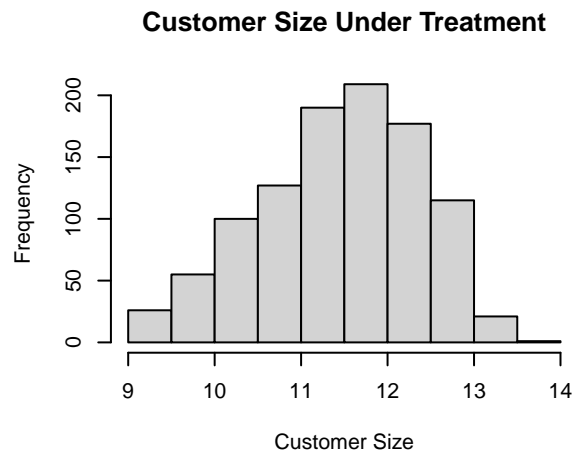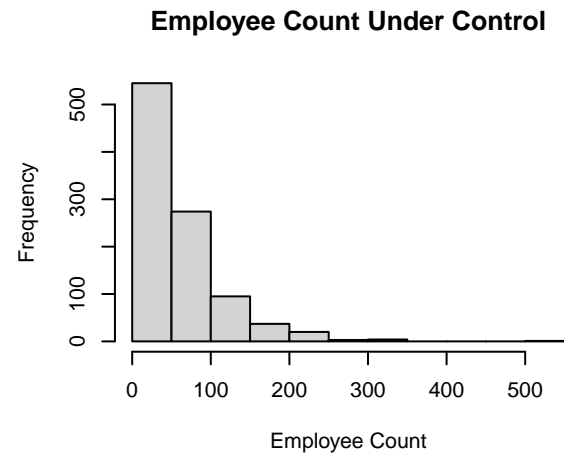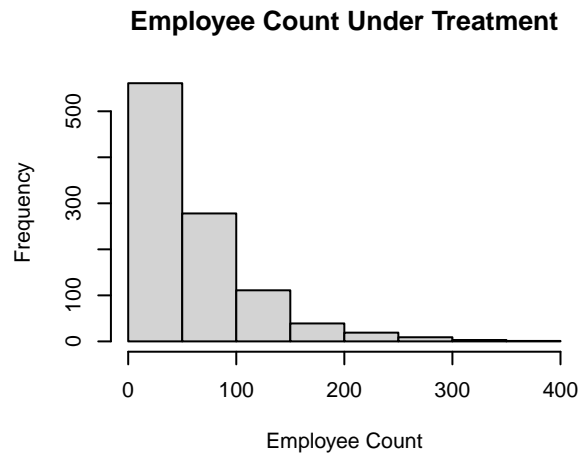$\{Y^{(0)}, Y^{(1)}\}A|X$

3.2 Estimation

Under the above assumptions, which are known as the three main identification assumptions, our parameter of interest can be identified and then estimated using the Outcome Regression estimator, the Inverse Probability Weighting estimator, the Hajek estimator as well as the Doubly Robust estimator.
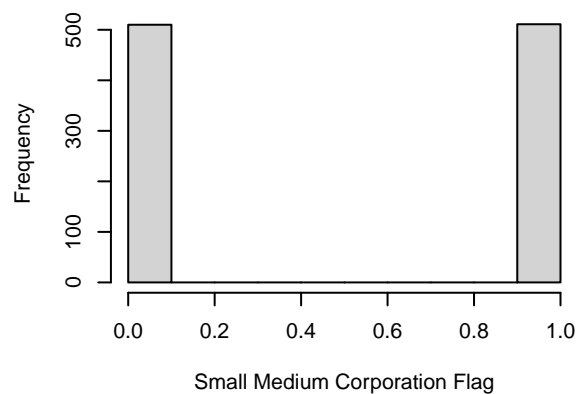
## 3.3 Sensitivity Analysis

4 Conclusion

5 Appendix

Employee Count Under Treatment

Employee Count Under Control

Customer Size Under Treatment

Customer Size Under Control

Commercial Flag Under Treatment

Commercial Flag Under Control

Small Medium Corporation Flag Under Treatment

Frequency

Small Medium Corporation Flag

Small Medium Corporation Flag Under Control

Frequency

Small Medium Corporation Flag
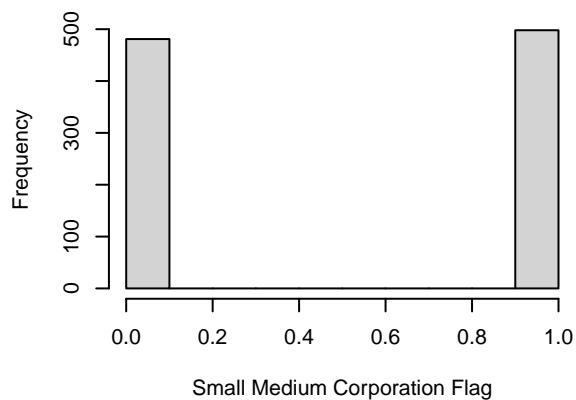
Large Consumer Flag Under Treatment

Frequency

Large Consumer Flag
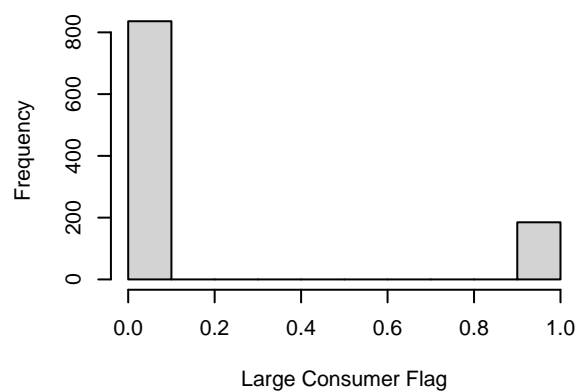
Large Consumer Flag Under Control

Frequency

Large Consumer Flag
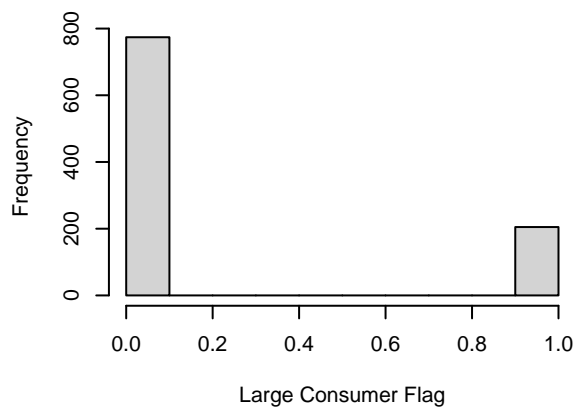
Global Offices Flag Under Treatment

Frequency

Global Offices Flag
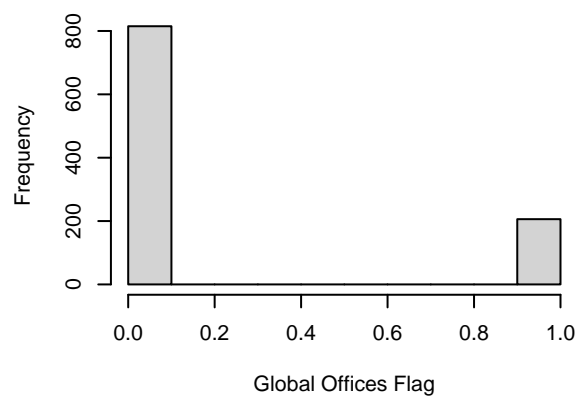
Global Offices Flag Under Control

Frequency

Global Offices Flag