Measuring ROI on Sponsored Search Ads

2022-11-18

Background Introduction

Bazaar.com is the leading online retailer in the United States using both display advertising and search engine advertising, running paid search ads on Google and Bing. It releases its ads in response to keywords from onlne customers and classifies into two categories: branded and nonbranded. Brand keywords contain the brand name such as 'Bazaar shoes' and 'Bazaar guitar'. Nonbranded keywords include items without brand name such as 'shoes' and 'guitar'.

Considering traffic data from Google and Bing, Bob who is from Bazaar's marketing analytics team computed that ROI is 320% associated with sponsored search ads. His result is problematic because people who search with the word 'Bazaar' already had the intent to visit Bazaar.com, so we doubt the effectiveness of branded keyword ads. Our goal is to understand the causal inference of the search ads and their effectiveness, the following analyzing will be proceeded:

(a) What's wrong with Bob's ROI analysis?
(b) Define the Treatment and Control.
(c) Consider a First Difference Estimate.
(d) Calculate the Difference-in-Differences.
(e) Given the Treatment Effect Estimate, Fix Bob's RoI Calculation.

(a) What's wrong with Bob's ROI analysis?

Bob states that the average probability of making a purchase from a website is 12%. He considers CTR (click-through rate) as a conversion on the landing page. The CTR in this case is Ad Clicks/(Sponsored Clicks + Organic Clicks). However, he doesn't consider people who use sponsored ads to navigate the website and make a purchase. That is, within all customers who click on sponsored ads, only some of them really visit Bazaar's website because of the sponsored ads. Even if they don't see the advertised products in search engine, they will still enter the website and purchase it from organic links by specifically entering 'Bazaar' in the search bar. Therefore, we cannot count the CTR as the contribution of sponsored ads. Bob's method will result in the overestimation of revenue and ROI.

(b) Define the Treatment and Control.

The unit of observation is the weekly clicks from each searching engine. We set the treatment group as Google, and others searching engine such as ask, bing, yahoo are control groups.

(c) Consider a First Difference Estimate

We built a linear regression model on the treated group(Google searching engine), which dependent variable is avg_org and independent variable is after(a dummy variable which represents the intervention).

```
# define the 'treatment' variable: 1=treat group, 0=control group
# define the 'after' variable: 1=treatment weeks, 0=pre-treatment weeks
data <- data %>% mutate(treatment = ifelse(platform == 'goog',1,0),
                        after = ifelse(week > 9,1,0))

#treat_week = wk10~12
data$treatment <- as.factor(data$treatment)
data$week <- as.factor(data$week)

goog <- data %>% filter(treatment==1)
summary(lm(avg_org ~ after, data = goog))

##
## Call:
## lm(formula = avg_org ~ after, data = goog)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1753.67  -631.92    26.67   505.83  2067.33
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2266.7      402.7   5.629 0.000219 ***
## after         4277.3      805.4   5.311 0.000342 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1208 on 10 degrees of freedom
## Multiple R-squared:  0.7383, Adjusted R-squared:  0.7121
## F-statistic: 28.21 on 1 and 10 DF,  p-value: 0.0003422
```

From the above result, p-value is 0.0003422 < alpha. The coefficient of intercept is 2266.7, which means the average organic click is 2266.7 before stopping sponsored search ads. The coefficient of after is 4277.3, which means it will averagely increase 4277.3 organic clicks after the intervention.

```
data %>% group_by(treatment,after) %>% summarise(mean_avg_org = mean(av
g_org))
```

```
## `summarise()` has grouped output by 'treatment'. You can override us
ing the
## `.groups` argument.
```

```
## # A tibble: 4 × 3
## # Groups:   treatment [2]
##   treatment after mean_avg_org
##   <fct>     <dbl>        <dbl>
## 1 0             0        1490.
## 2 0             1        3474.
## 3 1             0        2267.
## 4 1             1        6544
```

We assumed the market remains constant and the organic search will not change due to outside effects. However, as the avg_org change by time, we cannot rely on pre-post estimation. Instead, we need a control group to capture what would have happened without the intervention. Therefore, we need to conduct Difference-in-Differences (DiD).

(d) Calculate the Difference-in-Difference

Before the Did, we check the parallel assumption to see whether the treated subjects would have continued in parallel with the control.

```
summary(lm(avg_org ~ treatment*week, data=data))
```

```
## 
## Call:
## lm(formula = avg_org ~ treatment * week, data = data)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -1819.33   -43.67    12.67   299.42  1532.67 
## 
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)    
## (Intercept)          302.7      556.5   0.544 0.591544    
## treatment1           210.3     1113.0   0.189 0.851698    
## week2                237.0      787.0   0.301 0.765901    
## week3                569.0      787.0   0.723 0.476673    
## week4                805.0      787.0   1.023 0.316574    
## week5               1078.7      787.0   1.371 0.183186    
## week6               1558.0      787.0   1.980 0.059323 .  
## week7               1816.3      787.0   2.308 0.029937 *  
## week8               2193.3      787.0   2.787 0.010235 *  
## week9               2425.7      787.0   3.082 0.005101 ** 
```

```
## week10                2896.0      787.0   3.680 0.001178 **
## week11                3200.7      787.0   4.067 0.000445 ***
## week12                3416.7      787.0   4.341 0.000222 ***
## treatment1:week2         6.0     1574.0   0.004 0.996990
## treatment1:week3       115.0     1574.0   0.073 0.942363
## treatment1:week4       538.0     1574.0   0.342 0.735476
## treatment1:week5       587.3     1574.0   0.373 0.712318
## treatment1:week6       487.0     1574.0   0.309 0.759688
## treatment1:week7       925.7     1574.0   0.588 0.561964
## treatment1:week8      1045.7     1574.0   0.664 0.512812
## treatment1:week9      1395.3     1574.0   0.886 0.384156
## treatment1:week10     2649.0     1574.0   1.683 0.105350
## treatment1:week11     2971.3     1574.0   1.888 0.071211 .
## treatment1:week12     2959.3     1574.0   1.880 0.072284 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 963.9 on 24 degrees of freedom
## Multiple R-squared:  0.8399, Adjusted R-squared:  0.6864
## F-statistic: 5.473 on 23 and 24 DF,  p-value: 5.034e-05
```

We found that the p-value of interaction from week1 to week9 are larger than alpha, which are not significant. Therefore, the parallel assumption is effective. Also, the p-value of interaction from week10 to week12 are smaller. We can conclude that the treatment effect is effective.

```
summary(lm(avg_org ~ treatment + after + treatment * after, data=data))

##
## Call:
## lm(formula = avg_org ~ treatment + after + treatment * after,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1928.78  -847.92   -52.67   825.00  2067.33
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1489.7      215.4   6.917 1.51e-08 ***
## treatment1          777.0      430.7   1.804   0.0781 .
## after              1984.1      430.7   4.607 3.49e-05 ***
## treatment1:after   2293.2      861.4   2.662   0.0108 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1119 on 44 degrees of freedom
## Multiple R-squared:  0.6043, Adjusted R-squared:  0.5773
## F-statistic:  22.4 on 3 and 44 DF,  p-value: 5.881e-09
```

We built a linear regression model, the dependent variable is avg_org and the independent variables are treatment, after and interaction between treatment and after. The p-value is 5.881e-09. We can conclude that there is a 2293.2 change in average organic over time between the treatment and control group. The causal effect of stopping sponsored serving is 2293.2.

(e) Given the Treatment Effect Estimate, Fix Bob's RoI Calculation

```
summary(lm(avg_spons~treatment,data=filter(data,after==0)))

## 
## Call:
## lm(formula = avg_spons ~ treatment, data = filter(data, after == 
##     0))
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -5250.2 -2331.0  -701.3  2663.2  6627.7 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   3775.3      612.7   6.161 5.32e-07 ***
## treatment1    2347.9     1225.5   1.916   0.0638 .  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 3184 on 34 degrees of freedom
## Multiple R-squared:  0.09744,    Adjusted R-squared:  0.0709 
## F-statistic: 3.671 on 1 and 34 DF,  p-value: 0.06382
```

The p-value is 0.06382 which is larger than alpha. We can conclude that the average sponsors search for treatment and control group are similar before week 10.That is, Google's average sponsored clicks after week 10 doesn't have significant difference from other platforms. We then try to use the average number of sponsored clicks in the control group after week 10 to estimate the treatment group.

```
data2 <- data %>% filter(after==1 & treatment==0)
mean(data2$avg_spons)

## [1] 9855.889
```

The average sponsored clicks on Google is 9855.889 if it hasn't stopped the sponsored ads.

```
google = 9855.9
organic = 2293.2
realclick = google-organic
percent = 0.12 # Average probability of purchasing from Bazaar.com
```

```r
margin = 21 # margin per conversion
rev = percent*margin # average revenue per click
cost = 0.6 # Average cost per click
roi = (rev-cost)*realclick/(google * cost)*100
roi

## [1] 245.5447
```

The new ROI should be 245.5447%.