Causal - Measuring ROI on Sponsored Search Ads

Yu Chun Peng, Chien-Chu Hsu, Chia-Yen Ho, Carol Ng

4/16/2022

Business Overview

Bazaar.com is a leading online retailer in the US with extensive advertising presence in both display advertising and search engine advertising running paid search ads on Google and Bing. Bazaar's paid ads are broadly classified into two main categories by keywords: branded and nonbranded. Brand keywords contain the 'Bazaar' brand name such as 'Bazaar', 'Bazaar shoes', 'Bazaar clothes' and so on, while nonbranded are keywords such as 'shoes', 'dress' that do not contain 'Bazaar'.

With regards to traffic data by Google and Bing, Bob, who is from Bazaar's marketing analytics team, computed a 320% ROI on their sponsored ad spending. His result is skeptical because people who search with the word 'Bazaar' already had the intent to visit Bazaar.com, hence the effectiveness of branded keyword ads is implausible. Our goal is to understand the causal effect of the search ads and their effectiveness, the following questions will be addressed for an in-depth analysis overall:

1. What's wrong with Bob's ROI analysis?
2. What is the treatment and control of the experiment?
3. Is the First Difference reliable to estimate the causal effect of the treatment?
4. How should we compare with Difference-in-Difference estimation or Pre-Post estimation?

5. Based on our new treatment effect, what should be the corrected ROI? How is it compared with Bob's ROI?

Executive Summary

Diving deep into Bob's ROI calculation, we discovered two flaws in the calculation that make the figure inaccurate and not presentative - 1) Overvaluation of revenue and 2) Unaccounted opportunity cost

Overvaluation of revenue

The current ROI calculation assumes that everyone who clicked on the sponsored ads is casually driven by the sponsored ads, the error suggests the portion of people who have searched by the branded keywords already has the intent to visit Barzaar.com, this portion of people logically is not driven by the sponsored ads. Instead, they would still visit the website via organic link no matter if the sponsored ads existed or not, so they should not be included in the ROI calculation.

Therefore, if Bob included 'branded keyword search' clicks from these users visiting through the sponsored ads to be causal, his ROI calculation would have overestimated the revenue brought by the sponsored ads. We found that actually only 81% of total traffic via sponsored ads was a true result of displaying a sponsored ad.

Unaccounted opportunity cost

From the above, we knew that 81% is the true traffic driven by the sponsored ads. Currently, Bazaar.com is still paying a cost per click of $0.6 for the other 19% of branded keyword search visitors, who still visit the website via organic links in the absence of sponsored ads. This is an opportunity cost for the company since this advertising expense could have been spent on alternative marketing strategies.

To estimate the correct causal impact of sponsored ads on Bazaar.com's traffic, we used Difference-in-Difference through the following steps:

1) Calculate the first difference for weekly average traffic (Ads + Organic) over the course of before and after the technical hitch. This gives us the raw effect of sponsored ads of Google search instead of the overall treatment effect

2) Compare the first level pre-post difference in Google and the first level pre-post difference in other search engines. This determines the true incremental effect as this step handles possible confounders like seasonal variations across weeks as well as market factors.

As presented from the new treatment effects of the experiments, an average of 9.9K clicks per week are lost in the absence of sponsored ads, which is around 81% of the weekly traffic from the sponsored ads; thus we drew a conclusion of the rest of 19% traffic are visiting Bazaar.com through organic links instead of sponsored ads and this 19% should be ruled out from the original ROI calculation.

Based on the above two factors with new treatment effects, we re-calculated a corrected ROI from sponsored branded ads. Weeding out the false overvaluation of sponsored ads revenue and adjusted opportunity cost, the earlier calculation of ROI 320% is down to current adjusted ROI 241%, which is still a decent return for investment in sponsored ads advertising.

Experiment

We used 12 weeks of data including different platforms(Google, Bing, Yahoo, Ask), average sponsored branded keywords, and average organic search results. The technical glitch occurred on week 10-12, leading to no sponsored ads on Google platform. On the other hand, the other platforms run the keyword ads uninterruptedly for 12 weeks. Given the situation, we want to understand the causal effect of the ads. Treatment is the interrupted sponsored ads in Google during week 10-12. Google acts as the treatment group and other search engines such as Bing, Yahoo, and Ask act as the control group.

Threats to Causal Inference

Selection Bias: Ideally, the subjects in an experiment should be very similar to one another and to the larger population from which they are drawn to avoid selection bias. In our case, given that the ad strategies, keyword bids, and the mix of potential customers visiting a website are the same across Google and other search engines, it's safe to assume this experimental setup wouldn't cause selection bias.

Omitted Variable Bias: We suspect that some external variables which are not included in the data set may also affect this experiment, for example, Google search engine may have other technical issues that we were not aware of, or Bazaar's competitors may happen to conduct a short-term but high-intensity advertising strategy on a certain search engine.

Simultaneity Bias: In this experiment, the bidirectional effect between the variable and dependent variable is not considered. In other words, only ad clicks are considered to influence the website visits in the experiment but sometimes website visits may also influence ad clicks. Because customers may search back and forth between Bazaar products and their competitors in the final stage of their purchase decision, resulting in an increase in ad clicks.

Measurement Error: There may be some level of issues with click-through measurement for sponsored ads because we're not sure if consumers actually viewed the ad. It's quite possible that consumers clicked on the ad by mistake, and immediately closed the landing page after clicking.

Data Overview

The data set is weekly average traffic data through four platforms for 12 weeks. After importing the data, we created a "treatment" variable for test vs. control groups, an "after" variable that indicates 1 for treatment weeks and 0 for pre-treatment weeks and a "total_traffic" variable to aggregate the sponsored and organic traffic.

```
data = read.csv('did_sponsored_ads.csv')
treatment_week = c(10,11,12)
data <- data %>% mutate(treatment = ifelse(platform == 'goog',1,0),
                        after = ifelse(week %in% treatment_week,1,0),
                        total_traffic = avg_spons + avg_org)
```

What is Wrong with Current RoI Calculation

Sponsored ads carry the role of promoting the survival and development of enterprises, and Sponsored ads expenditures also account for a large part of enterprise expenses. But assessing the effectiveness of ad spend has always been a challenge. We found that the main reason for Bob's ROI miscalculation is revenue inflation. That is to say, Bob didn't consider one of the behavioral factors of people who use sponsored ads to navigate to the website to make a purchase. Of all customers who click on sponsored ads, only some of them actually visited Bazaar's website because of the sponsored ads. Others, even if they don't see the product being advertised in a search engine, they will still enter the website by organic links to purchase products. As a result, the revenue these people bring in is not the contribution of sponsored ads. That is, if the marketing analytics team treats all user clicks from sponsored ads as true cause and effect, this can lead to an overestimation of revenue and ROI.
Below is a quick example to demonstrate use case:
If 100 users reach Bazaar through sponsored ads, we would have estimated the ROI based on all 100 users whose average probability of making a purchase is 12%, the average margin per customer is $21, and the average cost per click is $0.6.

ROI = (($21 * 0.12 * 100) - (100 * $0.6)) / (100 * $0.6) = 320%

```
ROI = ((21 * 0.12 * 100) - (100*0.6) )/ (100 * 0.6)
ROI
```

```
## [1] 3.2
```

Whereas in reality, perhaps only 40 were actually caused by the ads. Adj_ROI = (($21 * 0.12 * 40) - (100 * $0.6)) / (100 * $0.6) = 68%

```
Adj_ROI = ((21 * 0.12 * 40) - (100* 0.6)) / (100 * 0.6)
Adj_ROI
```

```
## [1] 0.68
```

Define the Treatment and Control

Treatment here is the stop of sponsored search ads on Google during 10th, 11th, 12th weeks. We used Google as the treatment group while Bing, Yahoo and Ask acted as the control groups.

Consider a First Difference Estimate

```r
# Try a simple pre-post estimator
data_google<- data %>%
  filter(platform == "goog")

l1 <- lm(log(total_traffic) ~ after, data = data_google)
summary(l1)
```

```
##
## Call:
## lm(formula = log(total_traffic) ~ after, data = data_google)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.54933 -0.15495  0.03784  0.46975  0.95834
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.783506   0.248968  35.280 7.94e-12 ***
## after       0.001306   0.497936   0.003    0.998
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7469 on 10 degrees of freedom
## Multiple R-squared:  6.88e-07,   Adjusted R-squared:   -0.1
## F-statistic: 6.88e-06 on 1 and 10 DF,  p-value: 0.998
```

```r
exp(coef(summary(l1))[2])-1
```

```
## [1] 0.001306972
```

Interpretation:

With no sponsored search ads on Google, we observed a 0.13% decline in weekly total traffic to the website. The p-value is 0.998, thus we could not conclude that there are differences with and without sponsored ads. This is not a proper way of estimating the effect of sponsored ads. Using the pre-post estimate means that we assumed the market remains constant and the traffic will not change due to outside effects. So we simply compared before and after treatment in the treatment group. To be more specific, the nature of search is the same in week 1-9 and week 10-12. There is no seasonal variations across weeks.

With all these assumptions, it would be difficult to establish causality of sponsored ads. Hence, we resorted to Difference in Difference(DiD).

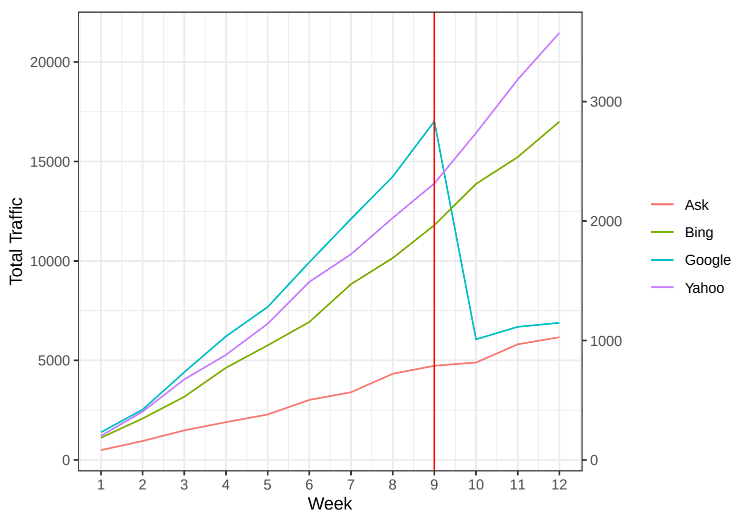
Calculate the Difference-in-Differences

Before calculating the difference in differences estimate of the treatment effect, we first checked the assumption of parallel trends. In pre-treatment weeks, treatment and control groups should have similar behavior. Also, the control group needs to be a good counterfactual for the treated group.

Visualization of parallel trend

```
temp1 = data %>%  filter(platform %in%  c('bing')) %>% select(week, total_traffic)
temp2 = data %>%  filter(platform %in%  c('yahoo')) %>% select(week, total_traffic)
temp3 = data %>%  filter(platform %in%  c('ask')) %>% select(week, total_traffic)

ggplot(data %>% filter(platform == 'goog'), aes(x=week, y= total_traffic, color = 'Google')) +
  geom_line() +
  geom_line(aes(x=week, y= total_traffic, color = 'Bing'), data = temp1) +
  geom_line(aes(x=week, y= total_traffic, color = 'Yahoo'), data = temp2) +
  geom_line(aes(x=week, y= total_traffic, color = 'Ask'), data = temp3) +
  geom_vline(xintercept = 9,color='red') +
  scale_y_continuous(sec.axis = sec_axis(~./6)) +
  scale_x_continuous(breaks = seq(1, 12, by = 1)) +
  labs(y = "Total Traffic", x = "Week") +
  theme_bw() +
  theme(legend.title = element_blank())
```



From the above visualization, we didn't see parallel trends pre-treatment. Still, we used the dynamic DiD to check if the pre-treatment weeks are not significant, while the post-treatment weeks are significant.

```
did_dyn <- lm(total_traffic ~ treatment + factor(week) + treatment * factor(week),data=data)
summary(did_dyn)
```

```
##
## Call:
## lm(formula = total_traffic ~ treatment + factor(week) + treatment *
##     factor(week), data = data)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8710.7  -111.8    87.3  1422.3  6586.3
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 936.3     2465.2   0.380 0.707414
## treatment                   449.7     4930.3   0.091 0.928087
## factor(week)2               881.3     3486.3   0.253 0.802574
## factor(week)3              1964.7     3486.3   0.564 0.578291
## factor(week)4              2998.3     3486.3   0.860 0.398274
## factor(week)5              4023.3     3486.3   1.154 0.259840
## factor(week)6              5361.0     3486.3   1.538 0.137190
## factor(week)7              6584.7     3486.3   1.889 0.071069 .
## factor(week)8              7940.0     3486.3   2.278 0.031955 *
## factor(week)9              9204.3     3486.3   2.640 0.014337 *
## factor(week)10            10794.3     3486.3   3.096 0.004932 **
## factor(week)11            12445.3     3486.3   3.570 0.001550 **
## factor(week)12            13940.3     3486.3   3.999 0.000529 ***
## treatment:factor(week)2     259.7     6972.5   0.037 0.970600
## treatment:factor(week)3    1055.3     6972.5   0.151 0.880960
## treatment:factor(week)4    1826.7     6972.5   0.262 0.795571
## treatment:factor(week)5    2274.7     6972.5   0.326 0.747075
## treatment:factor(week)6    3187.0     6972.5   0.457 0.651723
## treatment:factor(week)7    4140.3     6972.5   0.594 0.558196
## treatment:factor(week)8    4909.0     6972.5   0.704 0.488177
## treatment:factor(week)9    6424.7     6972.5   0.921 0.365997
## treatment:factor(week)10  -6122.3     6972.5  -0.878 0.388613
## treatment:factor(week)11  -7146.3     6972.5  -1.025 0.315616
## treatment:factor(week)12  -8437.3     6972.5  -1.210 0.238030
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4270 on 24 degrees of freedom
## Multiple R-squared:  0.6819, Adjusted R-squared:  0.3771
## F-statistic: 2.237 on 23 and 24 DF,  p-value: 0.0278
```

From the result above, we noticed that the assumption didn't pass. Despite the assumption didn't pass, we still performed a Difference in Difference regression between treatment and control groups to estimate the true causality of the sponsored ads. Independent variables for the DiD regression would be Treatment, After and the Interaction between Treatment and After.

```r
did <- lm(total_traffic ~ treatment + after + treatment * after, data=data)
summary(did)
```

```
##
## Call:
## lm(formula = total_traffic ~ treatment + after + treatment *
##     after, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -8437.7 -3231.0  -510.5  3591.6  8630.0
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        5265.0      882.5   5.966 3.79e-07 ***
## treatment          3124.9     1765.0   1.770  0.08357 .
## after              8064.7     1765.0   4.569 3.94e-05 ***
## treatment:after   -9910.6     3530.0  -2.808  0.00741 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4586 on 44 degrees of freedom
## Multiple R-squared:  0.3274, Adjusted R-squared:  0.2816
## F-statistic: 7.141 on 3 and 44 DF,  p-value: 0.0005211
```

Interpretation:

Without sponsored ads on the Google platform, Bazaar.com on average loses 9910 clicks per week. The new treatment effect estimated over and above the control group behavior captures the true causality for the sponsored ads. This approach is more accurate in comparison to the pre-post estimate as we can understand the control and treatment groups' behavior.

Fix RoI Calculation, given New Treatment Effect Estimate

Average probability of purchasing from Bazaar.com is 12% given they clicked onto the website and average margin per conversion is \$21, hence average revenue per click is 0.12 x \$21 = \$2.52. Average cost per click for a sponsored ad is \$0.6, so Bob's ROI calculation is (Margin per conversion * probability of click - cost per click / cost per click) - (\$2.52 - \$0.60) / (\$0.60) = 320%.

To compute the correct ROI, we needed a more accurate estimate of the proportion who were causally driven by the sponsored ads. In the previous part, we already calculated the total traffic causally driven by the ads, this step we needed to determine the traffic that would have used organic search results, granted that the sponsored ads were not running anymore, to arrive at the proportion of traffic that was causally driven by ads.

```r
did_org <- lm(avg_org ~ treatment + after + treatment * after, data=data)
summary(did_org)
```

```
##
## 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 ***
## treatment           777.0      430.7   1.804   0.0781 .
## after              1984.1      430.7   4.607 3.49e-05 ***
```

```
## treatment: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
```

In the absence of running sponsored ads on Google, 2293 would have used organic search results to visit Bazaar.com.
Total clicks from sponsored ads (C) = clicks truly motivated by sponsored ads (A) + clicks by customers who would still visit Bazaar.com in the absence of sponsored ads (B)

A = 9,910 (new treatment effect in did) B = 2,293 (new treatment effect in did_org)

Proportion of true traffic = A / (A+B) = 9,910 / (9,910+2,293) = 0.8120954 (81%)

The new ROI should be (Margin per conversion * probability of click * proportion - cost per click) / cost per click

New_ROI = ($21 * 0.12 * 0.8120954- 0.6)/0.6 New_ROI = 241.080%

```r
New_ROI = (21 * 0.12 * 0.8120954 - 0.6)/0.6
New_ROI
```

```
## [1] 2.410801
```

Our corrected ROI from sponsored branded ads is 241% based on the new estimated treatment effect. Without the false augmentation of sponsored ads revenue, the earlier calculation of ROI 320% is down to 241%, which is still a decent return for investment in sponsored ads advertising.