

# ROI on Sponsored Search for Bazaar.com - Aakash Patil and Mayank Singh

## Overview

Bazaar.com, an e-commerce company, is evaluating the effectiveness of its sponsored search advertising strategy. A natural experiment occurred when the company temporarily paused its sponsored search ads on Google while continuing them on other platforms such as Yahoo, Bing, and Ask. This provides a unique opportunity to apply Difference-in-Differences (DiD) techniques to estimate the causal impact of paid ads on website traffic and ultimately, return on investment (ROI).

## Problem Statement

The core question is: **What is the true ROI of Bazaar.com's sponsored search campaigns?**

To answer this, we must isolate the **incremental benefit** of paid search — that is, traffic and conversions that are attributable to the ads, beyond what would have occurred through organic search or other channels.

**Loading and Exploring the Dataset** We begin our analysis by loading the panel dataset, which contains weekly traffic data (sponsored and organic) from multiple search platforms.

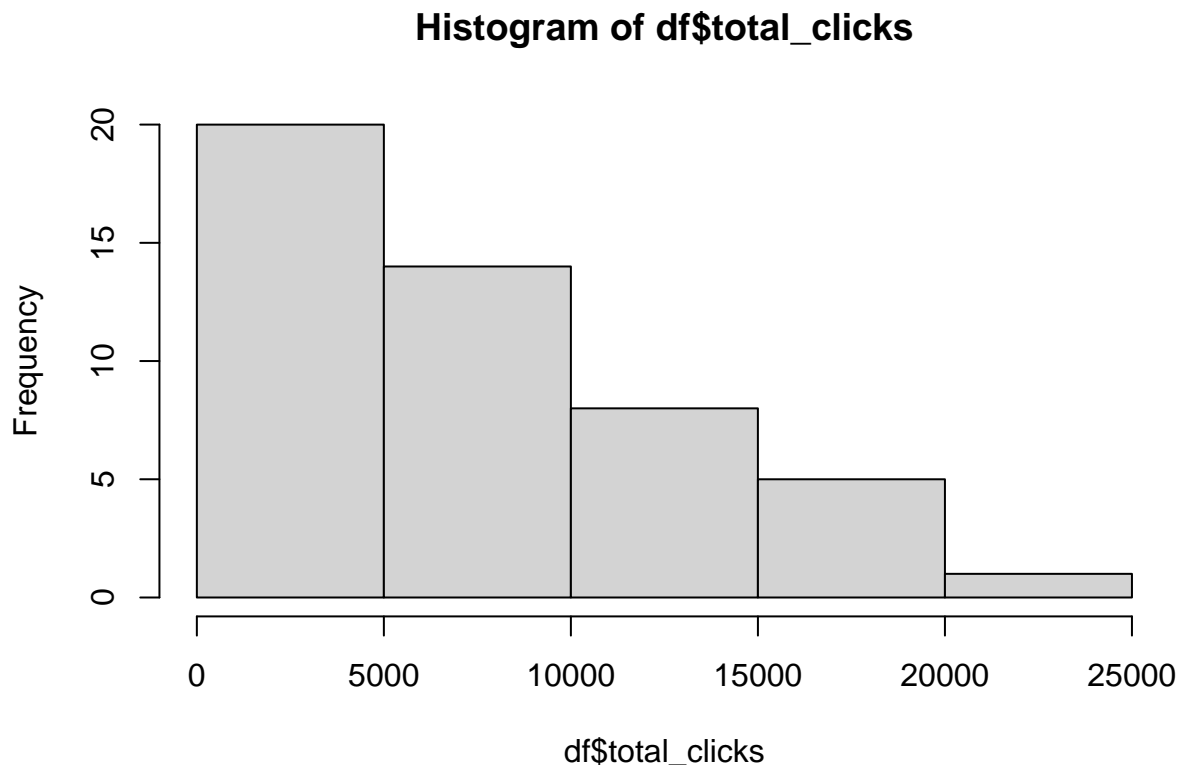
```
df = read.csv("~/Carlson/Sem2/Causal/HW3/did_sponsored_ads.csv")
summary(df)
```

```
##          id          platform          week          avg_spons
##  Min.      :1.00   Length:48      Min.      : 1.00   Min.      :    0
##  1st Qu.:1.75   Class :character  1st Qu.:  3.75   1st Qu.: 1665
##  Median :2.50   Mode  :character  Median :  6.50   Median : 3816
##  Mean    :2.50          Mean    :  6.50   Mean    : 5120
##  3rd Qu.:3.25          3rd Qu.:  9.25   3rd Qu.: 7834
##  Max.    :4.00          Max.    :12.00   Max.    :16211
##      avg_org
##  Min.      : 128
```

```
## 1st Qu.:1027
## Median :1904
## Mean   :2323
## 3rd Qu.:3448
## Max.   :6889
```

We will add `total_clicks`, `after` and `treatment` columns in the dataset and explore the distribution of `total_clicks`:

```
df <- df %>% mutate(treatment = ifelse(platform == "goog", 1, 0), after = ifelse(week %>= 10, 1, 0))
hist(df$total_clicks)
```



Since, it is right-skewed, we will consider taking log transform wherever required.

### (a) Bob's ROI Calculation

Bob's ROI calculation incorrectly assumes that all clicks from sponsored search ads are incremental — meaning they represent new customers who wouldn't have visited Bazaar.com otherwise. However, in reality, many users who click on a sponsored ad may have still visited the site via organic search results if the ad hadn't been shown. This leads to over-attribution of conversions and revenue to the ads.

Bob's calculation would give us an ROI of 320%, which we know is suspect because we are not considering clicks that would divert to organic search in the absence of paid ads.

Now let's simulate what should happen if we remove users who would have come via organic clicks anyway:

```
total_clicks <- 12203
revenue_per_conversion <- 21
conversion_rate <- 0.12
ad_cost_per_click <- 0.60
total_cost <- total_clicks * ad_cost_per_click
total_revenue <- total_clicks * conversion_rate * revenue_per_conversion
true_incremental_clicks <- 8000 # Let's say only 8000 clicks were truly incremental
organic_would_have_come <- 2000

# Proportion of truly incremental clicks
true_proportion <- true_incremental_clicks / (true_incremental_clicks + organic_would_have_come)

# Adjusted revenue using only incremental clicks
adjusted_revenue <- total_clicks * conversion_rate * revenue_per_conversion * true_proportion

roi_adjusted <- (adjusted_revenue - total_cost) / total_cost
roi_adjusted * 100 # New ROI
```

```
## [1] 236
```

The actual ROI you get is 236%. In reality, the ROI is lower than Bob thinks, unless we adjust for what users would do without the ad (i.e., click organically).

## (b) Treatment and Control Definitions

### Unit of Observation

Each row represents a **platform-week** (e.g., Google in week 3), forming a panel dataset observed over time across multiple platforms.

### Treatment Group

- **Platform:** Google
- **Treatment:** Removal of sponsored ads
- **Timing:** Weeks 10–12

## Control Group

- **Platforms:** Yahoo, Bing and Ask (ads continued throughout)
- **Timing:** All weeks

These control platforms provide the **counterfactual** — what would have happened to Google traffic if ads had not been turned off.

## (c) First Difference Estimate

We're estimating the pre-post change in sponsored traffic for Google only, using a simple regression:

```
google_df <- df %>% filter(platform == "goog") # Filter Google only

# Run regression of sponsored traffic on the post-treatment dummy (with log transform)
model_first_diff <- lm(log(total_clicks) ~ after, data = google_df)
summary(model_first_diff)
```

```
##
## Call:
## lm(formula = log(total_clicks) ~ after, data = google_df)
##
## 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
```

- **after coefficient (0.0013):** Indicates a **0.13% increase** in total clicks after ads were turned off. But...
- **p-value (0.998):** Not statistically significant, shows that we cannot attribute any meaningful change in traffic to the treatment using this model.

## Why It's Not a Good Idea to Solely Rely on This

Even if the **after** coefficient were significant, this method is still not reliable for causal inference.

### 1. No Control Group (No Counterfactual)

- You're only looking at Google.
- You don't know what would have happened to Google traffic if ads had not been turned off.
- Maybe traffic would have declined anyway (e.g., seasonal dip).

### 2. External Time Trends Are Ignored

- If overall web demand declined during weeks 10–12, the observed drop might have happened with or without the ad change.
- You can't separate the treatment effect from background noise.

## (d) Difference in Differences Approach

We turn to a more robust approach of dealing with potential confounding factors and selection bias: the difference-in-differences (DiD) method, which allows us to estimate causal effects by comparing changes over time between a treatment group and a control group.

### ### Checking Parallel Trends Assumptions

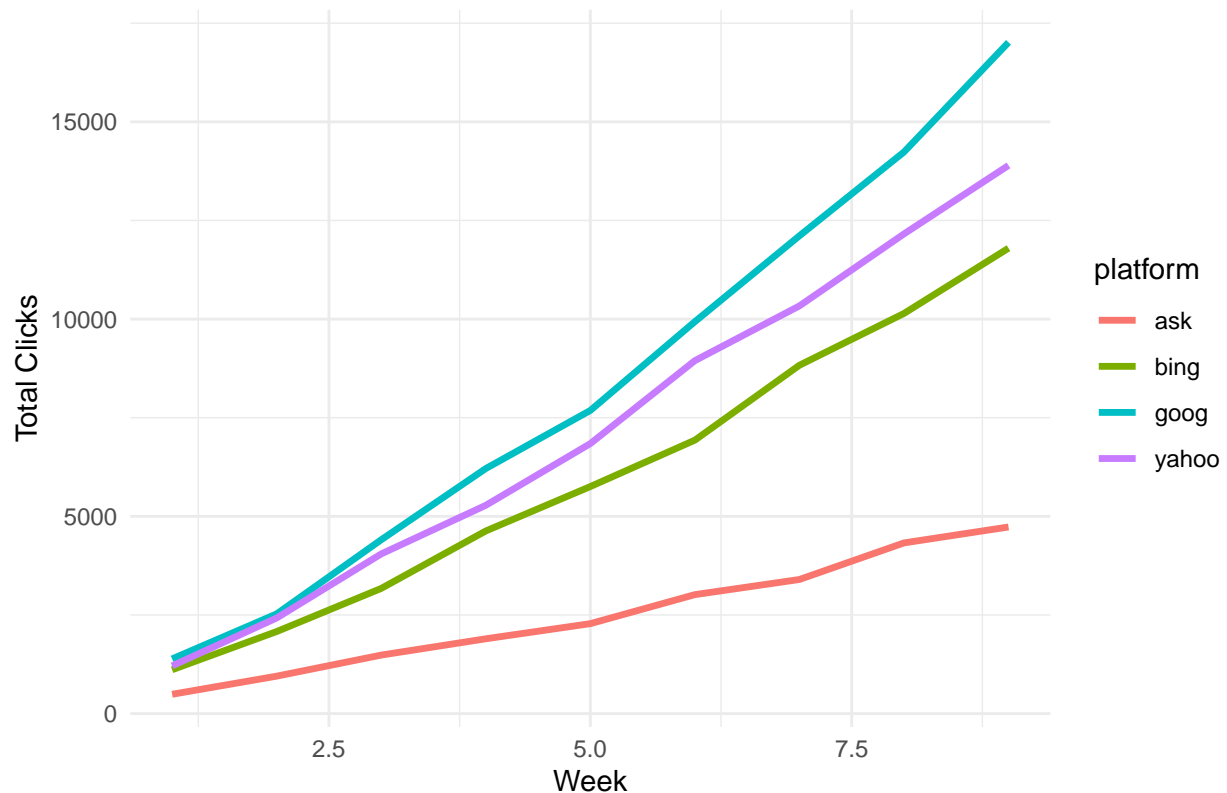
First, we check if control and treatment groups (Google vs others) moved in parallel in the pre-treatment period.

```
# Filter pre-treatment period
pre_df <- df %>% filter(after == 0)

# Visualize trends
ggplot(pre_df, aes(x = week, y = total_clicks, color = platform)) + geom_line(size = 1.2)
  labs(title = "Pre-Treatment Trends in Total Clicks (Parallel Trends Test)", y = "Total
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

Pre-Treatment Trends in Total Clicks (Parallel Trends Test)



```
# Regress total_clicks on factor(week) * treatment interaction (only pre-period)
parallel_test <- lm(total_clicks ~ factor(week) * treatment, data = pre_df)
summary(parallel_test)
```

```
##
## Call:
## lm(formula = total_clicks ~ factor(week) * treatment, data = pre_df)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-5410.7	-111.8	87.3	1176.2	3754.3

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	936.3	1648.2	0.568	0.576999
factor(week)2	881.3	2331.0	0.378	0.709780
factor(week)3	1964.7	2331.0	0.843	0.410370
factor(week)4	2998.3	2331.0	1.286	0.214637
factor(week)5	4023.3	2331.0	1.726	0.101465
factor(week)6	5361.0	2331.0	2.300	0.033635 *
factor(week)7	6584.7	2331.0	2.825	0.011222 *

```
## factor(week)8          7940.0      2331.0    3.406 0.003147 **
## factor(week)9          9204.3      2331.0    3.949 0.000941 ***
## treatment              449.7       3296.5    0.136 0.893013
## factor(week)2:treatment  259.7      4661.9    0.056 0.956195
## factor(week)3:treatment 1055.3      4661.9    0.226 0.823462
## factor(week)4:treatment 1826.7      4661.9    0.392 0.699787
## factor(week)5:treatment 2274.7      4661.9    0.488 0.631494
## factor(week)6:treatment 3187.0      4661.9    0.684 0.502920
## factor(week)7:treatment 4140.3      4661.9    0.888 0.386193
## factor(week)8:treatment 4909.0      4661.9    1.053 0.306274
## factor(week)9:treatment 6424.7      4661.9    1.378 0.185060
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2855 on 18 degrees of freedom
## Multiple R-squared:  0.7859, Adjusted R-squared:  0.5836
## F-statistic: 3.886 on 17 and 18 DF,  p-value: 0.003244
```

### Trend Plot Interpretation:

From the plot of total clicks by platform over weeks 1–9: Control platforms (Yahoo, Ask, Bing) have very similar growth trajectories. Google starts lower and grows more slowly — it visibly diverges from the control group. The lines are not parallel, especially as weeks progress.

### Regression Test Interpretation:

Interaction terms (treatment:factor(week)): These capture the difference in Google’s trend vs. control for each week, relative to week 1. All interaction p-values are  $> 0.18$ , and none are statistically significant. The largest interaction (week 9: 6424.7) is still not significant ( $p = 0.185$ ).

**Conclusion:** While the plot in your earlier test hinted at visual divergence, this more flexible regression using factor(week) shows no statistically significant differences in trends between Google and the control group during the pre-period.

**Now, we proceed with DiD to solve the problem:**

treatment:after (interaction) will be the true treatment effect — the causal impact of removing ads on Google traffic:

```
# DiD regression on full dataset (no log transform since we are taking differences)
did_model <- lm(total_clicks ~ treatment + after + treatment:after, data = df)
summary(did_model)
```

```
##
## Call:
## lm(formula = total_clicks ~ treatment + after + treatment:after,
##     data = df)
##
## 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
```

## New Treatment Effect Estimate

### −9,910.6 total clicks

This is the **Difference-in-Differences (DiD)** estimate of the causal effect of stopping ads on total clicks to Bazaar.com from Google. It implies that removing the ads resulted in a loss of approximately **9,911 clicks** that would have otherwise occurred.

- **Pre-post estimate (Google only):**
  - Estimate: **−0.13%** (−1,846)
  - Not statistically significant
- **DiD estimate (with controls):**
  - Estimate: **−9,911**
  - Statistically significant ( $p = 0.007$ )

## Why This Matters

- The **pre-post model underestimates** the true effect by ignoring rising traffic trends across all platforms.
- Google should have seen a **gain of ~8,065 clicks**, like Yahoo, Bing, and Ask, but instead saw a loss.



- This reinforces why we **cannot rely on pre-post estimates**:
  - Ignores external trends
  - Leads to biased, misleading estimates
  - Risks incorrect conclusions about ad effectiveness

## (e) Fixing Bob's ROI Calculation Using Causal Estimates

Bob's original ROI calculation mistakenly assumed that **all paid sponsored clicks were incremental**, which led to an inflated ROI. To properly adjust our ROI calculation, we also need to estimate how many users would have visited Bazaar.com organically in the absence of sponsored ads. These are users who may have clicked on the ad but would have come through organic search anyway — representing non-incremental traffic. We estimate this using a Difference-in-Differences (DiD) regression on organic clicks, comparing changes in organic traffic for Google versus the control platforms before and after the ad suspension.

```
# DiD regression for organic clicks
did_org_model <- lm(avg_org ~ treatment + after + treatment:after, data = df)
summary(did_org_model)
```

```
##
## Call:
## lm(formula = avg_org ~ treatment + after + treatment:after, data = df)
##
## 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
```

- The statistically significant treatment effect ( $p = 0.0108$ ) shows that turning off ads caused a shift of ~2,293 clicks from paid to organic search.

Our Difference-in-Differences (DiD) analysis gives us a more accurate breakdown of user behavior: - **A = 9,910**: Incremental clicks driven by sponsored ads (DiD estimate on total clicks) - **B = 2,293**: Clicks from users who would have visited Bazaar.com via organic search if ads weren't present (DiD estimate on organic clicks)

To ensure our ROI calculation reflects only the causal impact of sponsored ads, we adjust the revenue per click using the proportion of truly incremental clicks - First, we compute the share of sponsored clicks that were caused by the ads, excluding users who would have clicked organically anyway. Then, we calculate the expected revenue per click, adjusting for this incrementality.

This gives a more accurate picture of the actual value generated per paid click, which is essential for a valid ROI estimate.

```
A <- 9910      # Incremental clicks from DiD on total clicks
B <- 2293      # Organic clicks that would have come anyway
total_clicks <- A + B
margin_per_conversion <- 21
conversion_rate <- 0.12
cost_per_click <- 0.60

# Proportion of clicks that are truly incremental
incremental_proportion <- A / total_clicks
# Revenue per paid click (adjusted for only truly incremental ones)
expected_margin <- margin_per_conversion * conversion_rate * incremental_proportion
# Adjusted ROI
roi_adjusted <- (expected_margin - cost_per_click) / cost_per_click
roi_adjusted_percent <- roi_adjusted * 100
roi_adjusted_percent
```

```
## [1] 241.0801
```

**Why we subtract organic?** Because not all paid clicks are incremental — some users would've come via free organic clicks anyway.

**Why we scale margin by proportion?** Only a portion of the cost actually led to incremental profit.

**Why this matters?** Bob assumed 100% of clicks were valuable, which overstates ROI and misguides ad budgeting.

The adjusted ROI of **241.08%** reflects the true return on sponsored search ads, accounting only for incremental clicks that wouldn't have occurred without advertising. This more accurate measure excludes users who would have visited via organic search, preventing overstatement of ad effectiveness.