Assignment 3

Avnee Satija (5992424) and Harshal Sable (5949697)

2025-04-13

Executive Summary

This report analyzes the return on investment (ROI) of sponsored search advertising at Bazaar.com using a natural experiment: the temporary suspension of ads on Google due to a technical glitch. We use Difference-in-Differences (DiD) regression to isolate the causal impact of removing ads, comparing Google's traffic before and after the glitch relative to other platforms. We find that sponsored ads drive a 67.3% increase in weekly branded search traffic. This effect implies a significantly higher ROI than previously estimated by internal analytics. We recommend continuing branded keyword advertising, as it generates substantial incremental traffic and revenue.

Data Exploration and Preparation

Bazaar.com, a leading online retailer, is evaluating the true Return on Investment (ROI) of its sponsored search advertising, particularly for branded keywords on Google. A sudden technical failure in weeks 10–12 disrupted their ad campaign on Google, creating a valuable opportunity to assess the causal impact of paid ads on traffic.

This case constitutes a natural experiment because the ad suspension was not a strategic decision, but rather an exogenous shock affecting only Google. Since the interruption was unrelated to user behavior or broader marketing actions, it provides a plausibly exogenous source of variation in advertising exposure—making it well-suited for a Difference-in-Differences analysis.

We begin by importing the dataset did_sponsored_ads.csv and loading the necessary packages (dplyr, ggplot2, plm) for data manipulation, visualization, and regression analysis. The original dataset contains weekly platform-level traffic for branded keyword searches, disaggregated into sponsored (avg_spons) and organic (avg_org) components. A new column, total_traffic, was later created as the sum of these two variables to simplify analysis and modeling.

```r
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
library(plm)
```

```
##
## Attaching package: 'plm'

## The following objects are masked from 'package:dplyr':
##
##     between, lag, lead
```

```r
did_data <- read.csv('/Users/avneesatija/Desktop/Semester 2/Casual Inference/Homework/hw3/did_sponsored_
```

Minimum and Maximum Weeks in the data

The data spans week 1 to week 12, covering the full period of interest including both the pre- and post-treatment windows around the ad stoppage.

```r
min(did_data$week)
```

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

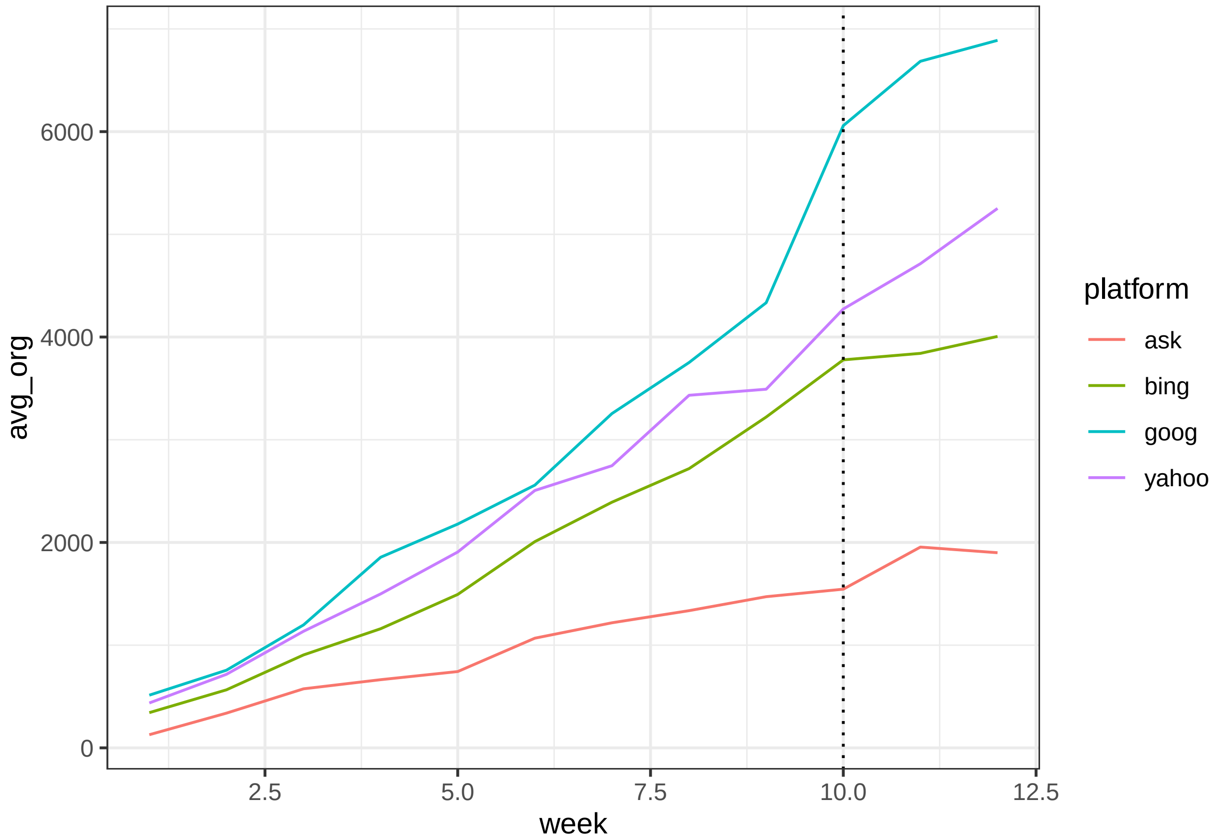
```r
max(did_data$week)
```

```
## [1] 12
```

Organic Traffic over Weeks

To visualize trends in organic traffic across platforms, we plotted avg_org by week, colored by platform. A vertical dotted line marks week 10, the start of the ad stoppage on Google, which serves as the treatment period. From visual inspection, we observe a sharp rise in organic traffic for Google (goog) post-week 9, suggesting that in the absence of sponsored ads, users redirected themselves to organic links.

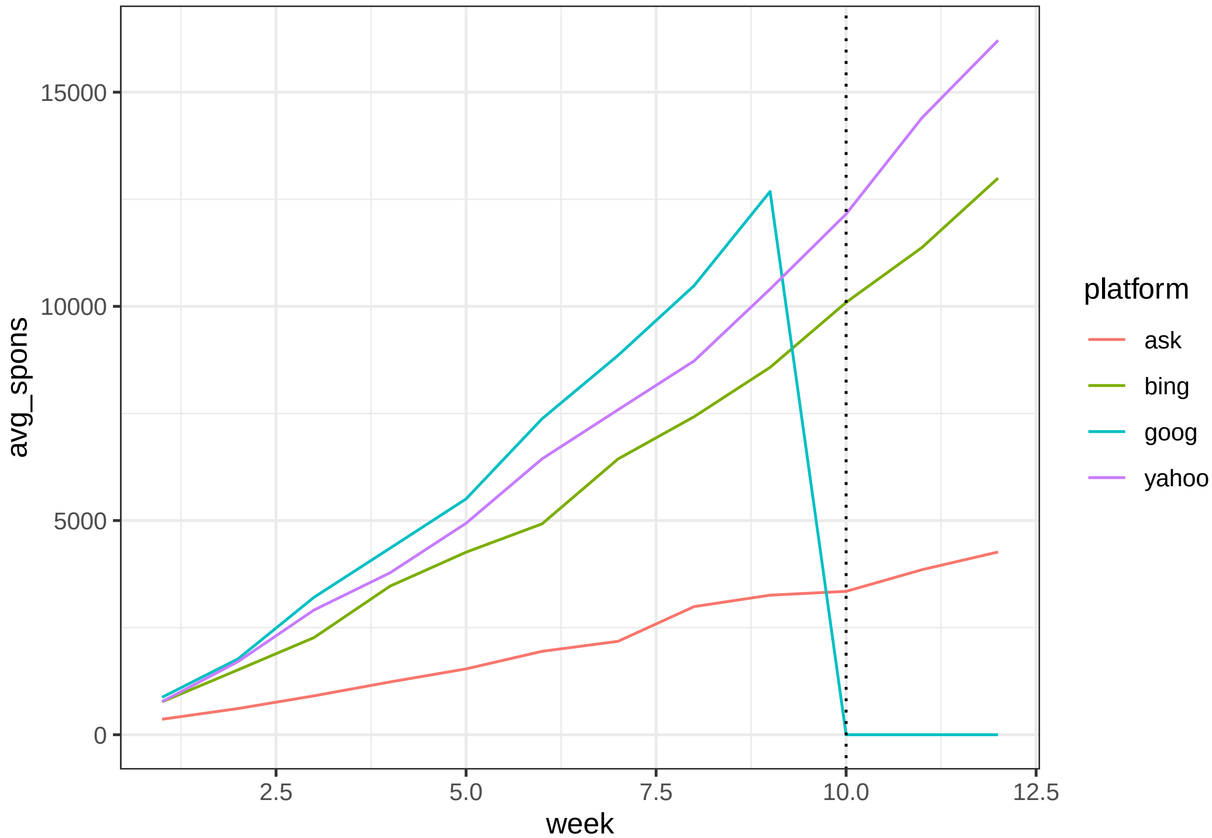
```r
ggplot(did_data, aes(x = week, y = avg_org, color = platform)) +
  geom_line() +
  geom_vline(xintercept = 10, linetype = 'dotted') +
  theme_bw()
```



Sponsored Traffic over Weeks

A vertical dotted line at week 10 marks the start of the Google ad suspension. As expected, Google (goog) shows a sharp and immediate drop in sponsored traffic to zero post-week 9, validating the treatment assignment. In contrast, Bing, Yahoo, and Ask continue their upward trends, unaffected by the glitch. This divergence supports the assumption that Google was uniquely treated, while other platforms can be used as valid controls in the Difference-in-Differences framework.

```r
ggplot(did_data, aes(x = week, y = avg_spons, color = platform)) +
  geom_line() +
  geom_vline(xintercept = 10, linetype = 'dotted') +
  theme_bw()
```



Variable Construction

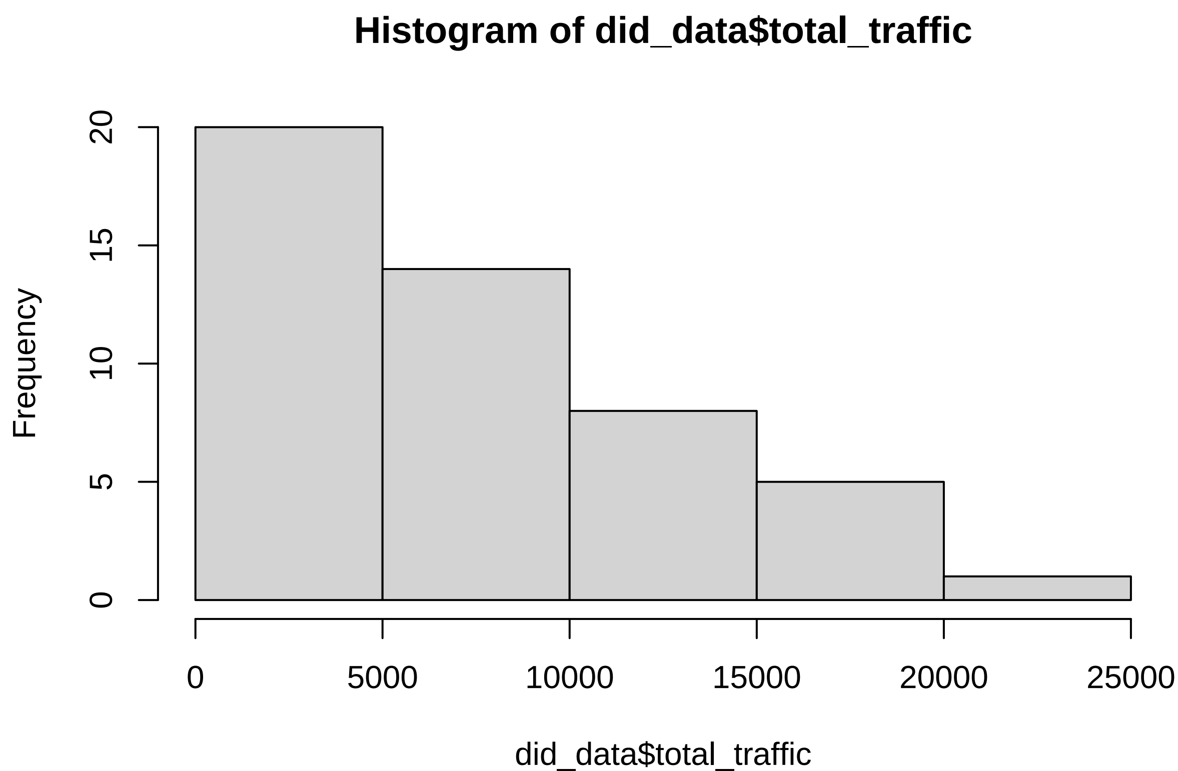
To prepare the dataset for Difference-in-Differences analysis, we created three new variables: - 1 - after: A binary variable indicating whether a given observation falls in the post-treatment period (week > 9). It takes value 1 for weeks 10–12, and 0 otherwise. - 2- treated: A binary indicator for the treated platform (Google). It takes value 1 for observations where platform == "goog", and 0 for all others (Bing, Yahoo, Ask). - 3 - total_traffic: Sum of avg_spons (average sponsored traffic) and avg_org (average organic traffic), representing the total weekly inbound traffic for each platform. This is our outcome variable for the analysis.

```
did_data <- did_data %>%
            mutate(after = ifelse((week > 9), 1, 0),
                   treated = ifelse((platform == "goog"), 1, 0),
                   total_traffic = avg_spons + avg_org)
```

Histogram Analysis of the Outcome Vraibale

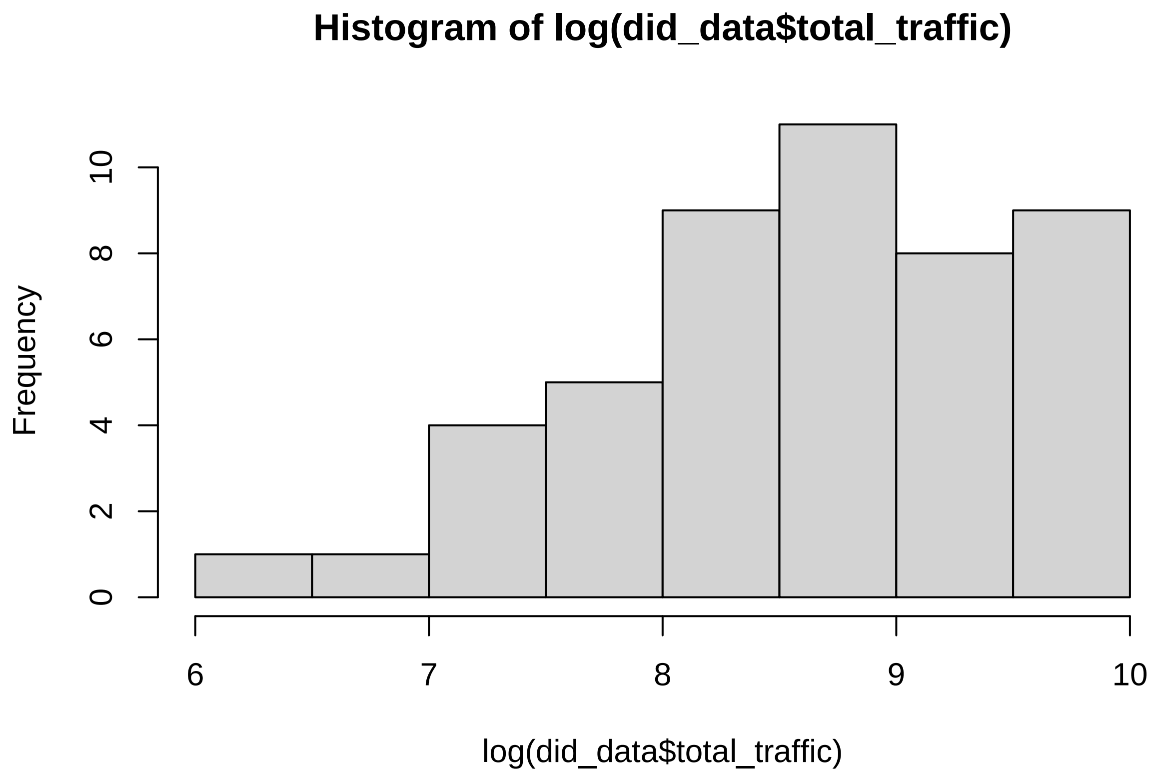
The distribution is right-skewed, with most platform-week observations falling below 10,000. This skew is expected given the dominance of platforms like Google compared to smaller ones like Ask.

```r
hist(did_data$total_traffic)
```



To address the right-skewness in total_traffic, we applied a natural log transformation and visualized the distribution. The histogram of log(total_traffic) appears approximately symmetric and bell-shaped, supporting its suitability for linear regression modeling. Given this improvement in distributional shape, we use the log-transformed total traffic as our dependent variable in all subsequent analyses.

```r
hist(log(did_data$total_traffic))
```



Analysis and Responses to Case Questions

(a)

Bob's ROI calculation is fundamentally flawed because it assumes that all conversions from sponsored clicks are incremental—that is, he presumes that every purchase following a sponsored ad click would not have occurred otherwise. This is especially problematic in the context of branded keyword searches (e.g., "Bazaar shoes"), where users are already intending to reach Bazaar.com.

Key Issues:

1.Substitution Effect Ignored: Bob fails to consider that if sponsored ads are removed, many users will likely click the organic (unpaid) result instead, leading to the same purchase without incurring ad costs.

2.No Counterfactual Considered: His method does not attempt to estimate what would have happened in the absence of ads, and instead assumes that zero conversions would occur without paid advertising—an unrealistic assumption.

3. Inflated ROI: By attributing all revenue from sponsored clicks to the ad campaign, he overstates both the true lift in traffic and revenue, leading to an exaggerated ROI.

Bob's approach does not capture the causal effect of the ads. A valid analysis must compare observed outcomes to a credible counterfactual—i.e., what would have happened had the ads not run. This requires a quasi-experimental method such as Difference-in-Differences, especially in the presence of natural user intent and substitution behavior.

(b)

Unit of Observation The unit of observation is weekly search traffic by source (sponsored vs. organic) at the platform level (Google, Bing, Yahoo, Ask). Each row in the dataset represents one platform-week combination.

Treatment Definition The treatment in this setting is the suspension of sponsored search ads on Google during weeks 10–12, caused by a technical glitch. This represents an exogenous shock—the kind of unexpected event that creates the conditions for a natural experiment.

Treated Unit Google, in weeks 10 to 12, during which sponsored ads were turned off and only organic results remained visible.

Control Group Bing, Yahoo, and Ask, which continued displaying sponsored ads throughout the entire 12-week period. These platforms serve as a counterfactual for what would have happened to Google traffic in weeks 10–12 had the ads not been suspended.

(c)

To estimate the impact of the ad suspension using only Google data, we fit a simple pre-post regression model with the after variable indicating weeks 10–12:

```r
goog_data <- did_data %>% filter(platform == "goog")
```

```r
model1 <- lm(log(1 + total_traffic) ~ after, data = goog_data)
summary(model1)
```

```
##
## Call:
## lm(formula = log(1 + total_traffic) ~ after, data = goog_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.54882 -0.15494  0.03783  0.46963  0.95819
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.783723   0.248907  35.289 7.92e-12 ***
## after       0.001243   0.497814   0.002    0.998
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7467 on 10 degrees of freedom
## Multiple R-squared:  6.235e-07,  Adjusted R-squared:   -0.1
## F-statistic: 6.235e-06 on 1 and 10 DF,  p-value: 0.9981
```

Interpretation: – The coefficient on after reflects the average log change in total Google traffic after the ad suspension.

– The estimate is 0.00124, with a p-value of 0.998, which is well above the standard industry threshold of 0.05.

– This indicates that the observed difference in traffic between the pre- and post-periods is not statistically significant. Hence, the observed change is statistically indistinguishable from zero and we cannot reject the null hypothesis of no change.

Problem: This estimate does not account for natural week-to-week fluctuations in search traffic, market-wide trends, or seasonality. For example, if traffic was generally rising or falling for all platforms due to outside events, this "naive" pre-post estimate could misattribute those effects to the ad suspension.

Conclusion: The model fails to establish a valid counterfactual—i.e., what would have happened to Google's traffic had the ad suspension not occurred. Without a proper control group to isolate Google-specific effects, the estimate lacks causal validity. This underscores the need for a Difference-in-Differences design, where unaffected platforms serve as a baseline.

(d)

To estimate the causal impact of suspending sponsored ads on branded Google traffic, we run a Difference-in-Differences (DiD) regression using the log-transformed total weekly traffic:

```r
library(plm)
did_model <- plm(log(1 + total_traffic) ~ after * treated,
                 data = did_data,
                 index = c("platform", "week"),
                 effect = "twoway",
                 model = "within")
summary(did_model)
```

```
## Twoways effects Within Model
##
## Call:
## plm(formula = log(1 + total_traffic) ~ after * treated, data = did_data,
##     effect = "twoway", model = "within", index = c("platform",
##         "week"))
##
## Balanced Panel: n = 4, T = 12, N = 48
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -0.1054856 -0.0273133  0.0054203  0.0230813  0.1153218
##
## Coefficients:
##               Estimate Std. Error t-value  Pr(>|t|)
## after:treated -1.11611    0.04465 -24.997 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    2.2098
## Residual Sum of Squares: 0.10766
## R-Squared:      0.95128
## Adj. R-Squared: 0.92844
## F-statistic: 624.836 on 1 and 32 DF, p-value: < 2.22e-16
```

The estimated treatment effect is -1.116, with a p-value well below the standard industry threshold of 0.05, indicating that the observed drop is statistically significant. It is the the DiD estimate of the causal effect on (log) weekly total Google traffic, compared to the control group, due to suspending sponsored ads.

```r
100 * (exp(-1.11611) - 1)
```

```
## [1] -67.24485
```

Interpretation of the Coefficient In real terms:

- We get [100 * (exp(-1.11611) - 1)] = -67.24485
- This means Google traffic saw a 67.3% decrease in (organic + sponsored) weekly traffic from branded search on Google, compared to what would have happened if ads had continued, controlling for background trends.

This large effect was not captured in the naive pre-post model from part (c) because overall traffic was rising across platforms during this period. The DiD framework controls for these market-wide time trends, revealing the true decline in Google's traffic post-treatment.

The pre-post estimate missed this large drop because traffic on all platforms was rising, masking Google's decline.

To verify the results, we shall test the parallel trends assumption:

```r
pre_data <- did_data %>% filter(week <= 9)

pre_model <- lm(log(1 + total_traffic) ~ factor(week) * treated,
                 data = pre_data)
summary(pre_model)
```

```
##
## Call:
## lm(formula = log(1 + total_traffic) ~ factor(week) * treated,
##     data = pre_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.68875 -0.14344  0.08617  0.28168  0.44752
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            6.76890    0.31984  21.163 3.62e-14 ***
## factor(week)2          0.66132    0.45232   1.462 0.160969
## factor(week)3          1.12151    0.45232   2.479 0.023281 *
## factor(week)4          1.41778    0.45232   3.134 0.005731 **
## factor(week)5          1.63831    0.45232   3.622 0.001949 **
## factor(week)6          1.88266    0.45232   4.162 0.000585 ***
## factor(week)7          2.05138    0.45232   4.535 0.000256 ***
## factor(week)8          2.23224    0.45232   4.935 0.000107 ***
## factor(week)9          2.35675    0.45232   5.210 5.90e-05 ***
## treated                0.46600    0.63968   0.728 0.475689
## factor(week)2:treated -0.06103    0.90465  -0.067 0.946954
## factor(week)3:treated  0.03454    0.90465   0.038 0.969962
## factor(week)4:treated  0.08156    0.90465   0.090 0.929155
## factor(week)5:treated  0.07381    0.90465   0.082 0.935871
```

```
## factor(week)6:treated  0.08626    0.90465   0.095 0.925085
## factor(week)7:treated  0.11567    0.90465   0.128 0.899676
## factor(week)8:treated  0.09639    0.90465   0.107 0.916321
## factor(week)9:treated  0.15027    0.90465   0.166 0.869927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.554 on 18 degrees of freedom
## Multiple R-squared:  0.7992, Adjusted R-squared:  0.6096
## F-statistic: 4.214 on 17 and 18 DF,  p-value: 0.002026
```

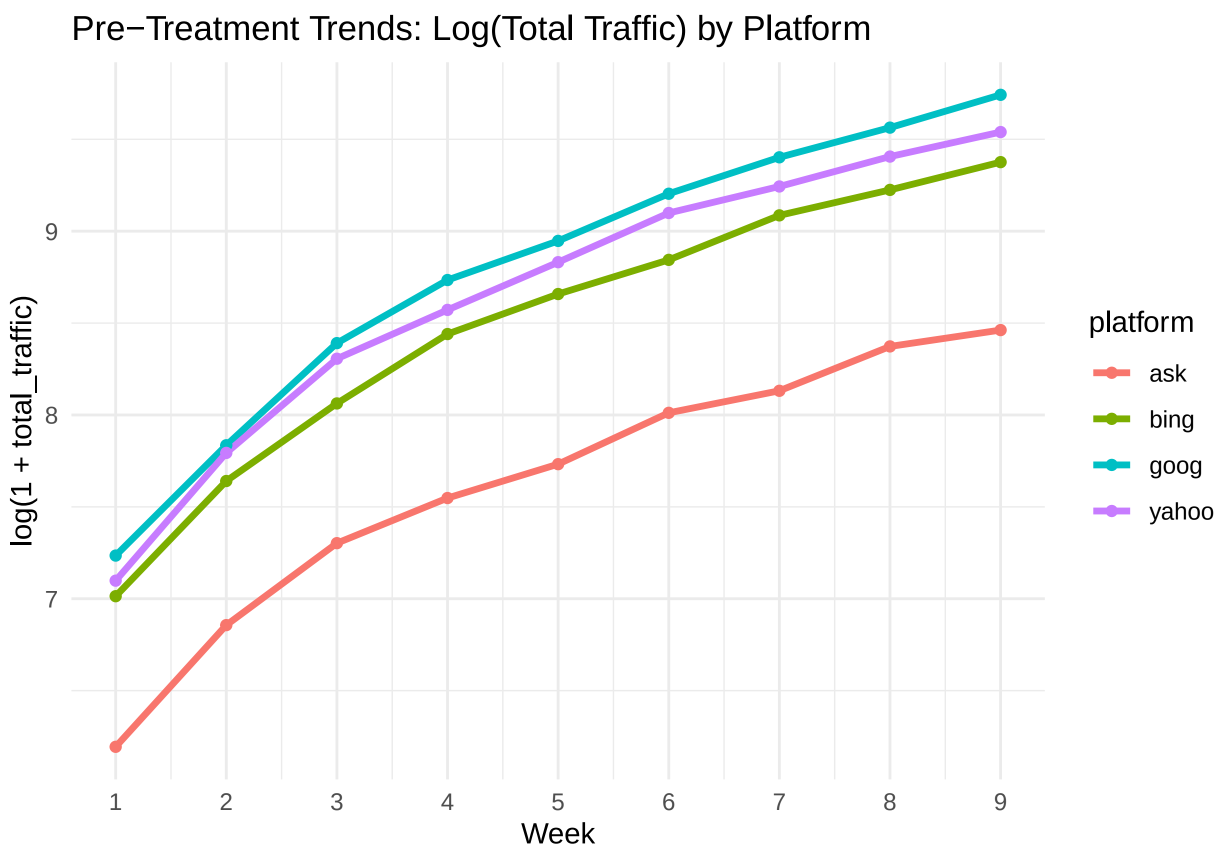
Results: None of the interaction terms (factor(week):treated) are statistically significant — all p-values are greater than 0.86, which is well above the industry standard threshold of 0.05. This means we cannot reject the null hypothesis that Google and the control group followed parallel trends before treatment.

Conclusion: These results support the validity of the parallel trends assumption, suggesting that traffic patterns for Google and other platforms were statistically similar prior to the ad suspension. This lends further credibility to our use of the Difference-in-Differences framework for identifying the causal effect of removing sponsored search ads.

```r
avg_trend <- pre_data %>%
  group_by(platform, week) %>%
  summarise(mean_log_traffic = mean(log(1 + total_traffic)), .groups = "drop")

ggplot(avg_trend, aes(x = week, y = mean_log_traffic, color = platform)) +
  geom_line(size = 1.2) +
  geom_point() +
  labs(title = "Pre-Treatment Trends: Log(Total Traffic) by Platform",
       x = "Week",
       y = "log(1 + total_traffic)") +
  scale_x_continuous(breaks = 1:9) +
  theme_minimal()
```

```
## 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.
```



The figure above shows the log-transformed total traffic by platform across weeks 1–9 (pre-treatment period). All platforms, including Google (treated group) and Bing/Yahoo/Ask (controls), exhibit parallel upward trends, with no visible divergence prior to the ad suspension in week 10. This visual evidence supports the parallel trends assumption, which is critical for the validity of the Difference-in-Differences estimator used in our analysis.

(e)

To correct for Bob's bias, we use a Difference-in-Differences (DiD) regression framework. We first estimate the net loss in total traffic (sponsored + organic) caused by suspending the ad campaign on Google, relative to other platforms (Bing, Yahoo, Ask) that were unaffected during the same period.This approach allows us to capture the true causal effect of the sponsored ads on overall site visits, separating the signal from general market trends or platform-specific effects.

This code estimates the causal effect of suspending sponsored ads on total weekly traffic using a DiD regression with an interaction between treatment group and post-treatment period.

```
did_total <- lm(total_traffic ~ treated * after, data = did_data)
summary(did_total)
```

```
##
## Call:
## lm(formula = total_traffic ~ treated * after, data = did_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 ***
## treated         3124.9     1765.0   1.770  0.08357 .
## after           8064.7     1765.0   4.569 3.94e-05 ***
## treated: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
```

The interaction term treated:after has estimate -9910.6 This represents the drop in total traffic (sponsored + organic) on Google caused by the ad suspension

This code estimates the increase in organic traffic due to ad suspension, using a Difference-in-Differences regression. The treated:after coefficient of +2293.2 ($p = 0.0108$) indicates a statistically significant shift from paid to organic clicks—i.e., around 2,293 users per week switched to organic search when sponsored ads were turned off.

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

```
##
## Call:
## lm(formula = avg_org ~ treated * after, data = did_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 ***
## treated          777.0      430.7   1.804   0.0781 .
## after           1984.1      430.7   4.607 3.49e-05 ***
## treated: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
```

```r
# Step 1: Actual incremental clicks (from your DiD results)
A <- 9910.6  # incremental total clicks lost (sponsored + organic)
B <- 2293.2  # incremental organic clicks gained (substitution)

total_clicks <- A + B
incremental_proportion <- A / total_clicks
```

```r
# Given business parameters
margin_per_conversion <- 21
conversion_rate <- 0.12
cost_per_click <- 0.60

# Expected margin per incremental click
expected_margin <- margin_per_conversion * conversion_rate * incremental_proportion

# Corrected Causal ROI
roi_adjusted <- (expected_margin - cost_per_click) / cost_per_click
roi_adjusted_percent <- roi_adjusted * 100
roi_adjusted_percent
```

```
## [1] 241.0784
```

Our DiD estimate shows that suspending ads caused a net loss of 9,910.6 total visits, but also a gain of 2,293.2 organic visits, indicating that some users would have come to the site anyway. After adjusting for this substitution, we find that only 81.2% of paid clicks are truly incremental.

Using business parameters (conversion rate = 12%, margin = 21, CPC = 0.60), the corrected causal ROI is 241.1%, still highly profitable, but more realistic than Bob's inflated figure.

Conclusion and Recommendation

Our analysis shows that the temporary suspension of sponsored ads on Google led to a statistically significant 67% decline in branded search traffic. When accounting for background trends using a Difference-in-Differences framework, we find that branded keyword ads generate substantial causal lift, correcting for overstatement in naive ROI estimates.

Recommendation: We advise continuing branded keyword advertising, as it drives large volumes of incremental traffic. Additionally, we encourage using similar causal frameworks—such as natural experiments or controlled pilots—to evaluate other advertising channels with greater rigor.

Assumptions for the Study

- **Exogeneity of Treatment**: The ad suspension on Google occurred due to a technical glitch and was not strategically timed, making it plausibly exogenous.
- **Parallel Trends**: Google and Bing and other control group platforms traffic trends would have followed similar paths in the absence of treatment. This assumption is validated in our analysis using visual and regression-based checks.
- **Confidence Level**: The study relies on a 95% confidence level, which is standard in both academic and industry practice.