HW3_Causal

Olivia Boe & Aashi Patni

2025-04-16

Homework 3

Summary of events:

Bazaar uses two different types of ads, display and search engine ads. The search engine ads are broken down into two categories: branded and non branded key words. Bazaar wants to know if the search engine ads are necessary or if their organic webpage is effective enough. Essentially, we are trying to examine if the search engine ads are worth it given an ROI analysis.

Areas for Concern

Potential SUTVA Violations:

During the pre-period, when ads were displayed on Google, users might have seen the sponsored web link and chosen not to click on it. This could lead to spillover effects, as the exposure to the sponsored ad in the previous week may influence users to click on the organic link instead. Additionally, users may come across the sponsored ad on other browsers and then later search on Google, resulting in a click on the organic link. Conversely, users who didn't see the ad on Google may later see it on other search engines like Bing, Yahoo, or Ask. For those who exclusively use Google as their browser, they will automatically be included in the control group.

Other Concerns

People who exclusively use Google will be placed in the control group, and since the Google campaign was much larger than the one on Bing, this discrepancy could raise concerns about the potential impact on the results.

Omitted Variable Bias

There may be external factors we haven't considered, such as changes in Google's algorithms, a sudden competitor ad campaign, or seasonality, that could influence visits and sales. If these factors changed during the test period, it could create a misleading impression that the ads had a significant effect (or no effect at all).

Additionally, while Bazaar mentioned having display ads, the study doesn't clarify whether they plan to alter these ads. It may be important to consider including them in the analysis, as they could be influencing visits to either sponsored or organic links.

Simultaneity Bias

Users who are browsing back and forth between websites to compare prices could introduce simultaneity bias. In this case, the user's visit and exposure to the ad may be influencing each other at the same time, making it hard to determine which action (the ad click or the visit) is actually driving the other.

Measurement Error

Clicking an ad doesn't always indicate genuine engagement. Some users may click by accident or quickly leave the page, which inflates click numbers without reflecting real interest or conversions.

If the analysis is being used to assess ROI, focusing on conversion rates would be more meaningful. Visits alone may not accurately measure success for Bazaar, as they don't capture actual user intent or the desired outcomes like purchases or sign-ups.

Import Data and Packages

```r
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

setwd("C:\\Users\\olivi\\OneDrive\\Documents\\Causal Econometrics 6441\\HW 3"
)
data = read.csv("did_sponsored_ads.csv")
```

Question A

What is Wrong with Bob's ROI Calculation?

Suppose a potential customer searches for "buy hiking gear Bazaar" on Google. Even if Bazaar.com doesn't have a sponsored ad, the customer might still:

- Click on the organic (non-sponsored) search result
- Go directly to the website (e.g., via bookmarks or typing the URL)
- Discover it through other channels like email, referrals, or SEO

In these cases, Bazaar wouldn't incur the 0.60 ad cost but could still make the 21 margin.

Additionally, clicks are not independent. Users might click multiple times before making a purchase or click and leave without converting. Bazaar.com assumes each click equals a potential conversion, but non-converting clicks still cost money.

Another factor to consider is that Bazaar is assuming a fixed 21 margin. In reality, margins can vary widely—some customers may buy a 5 item, while others might spend 200.

Question B

Define the Treatment and Control. What is the unit of observation here?

The control group will consist of Yahoo, Ask, and Bing, as they will be running ads throughout the entire 12-week period. Google will be the treatment group, where ads are only run during weeks 1-9. The treatment involves removing ads.

The post-treatment period will be weeks 10-12, while the pre-treatment period will be weeks 1-9.

The unit of observation is the number of visits for each search engine, tracked on a weekly basis.

First lets create a column for treatment. The treatment group will receive a value of 1.

```r
# New column for treated and control groups
data <- data %>%
  mutate(treatment = ifelse(platform == "goog", 1, 0))
```

Lets create a new column for post and pre. Weeks 1-9 will be pre and weeks 10 - 12 will be post.

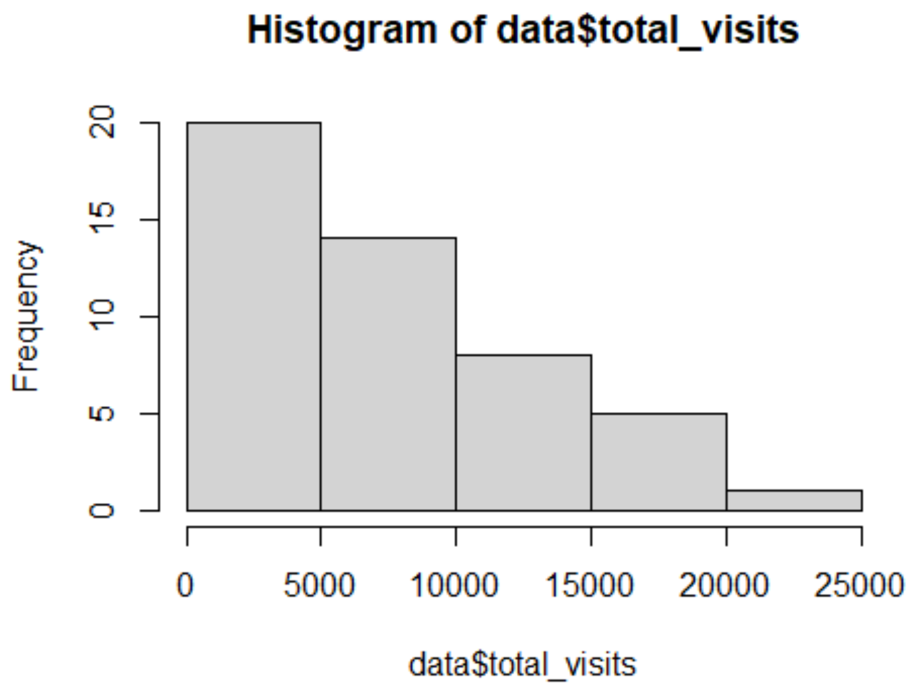
```r
# New column for post and pre treatments
data <- data %>%
  mutate(post = ifelse(week >= 10 & week <= 12, 1, 0))
```

Create a new column to get total visits

```r
# Add the average sponsored visits with average organic visits
data <- data %>%
  mutate(total_visits = avg_spons + avg_org)
```

We want to check the normality of the total number of visits to ensure that the data meets the assumptions required for statistical tests, such as parametric analysis or hypothesis testing.

```r
# Check if the data is skewed
hist(data$total_visits)
```



The total number of visits is right skewed, we will need to transform this in our analysis with a log transformation.

Question C

Consider a First Difference Estimate.

```
# We have to log transform the total visits
summary(lm(log(1+ total_visits) ~ post*treatment, data = data))

##
## Call:
## lm(formula = log(1 + total_visits) ~ post * treatment, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0591 -0.5441  0.1413  0.5809  1.2858
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      8.2536     0.1542  53.527  < 2e-16 ***
## post             1.1174     0.3084   3.623  0.00075 ***
## treatment        0.5302     0.3084   1.719  0.09262 .
## post:treatment  -1.1161     0.6168  -1.810  0.07720 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8012 on 44 degrees of freedom
```

```
## Multiple R-squared:  0.2415, Adjusted R-squared:  0.1898
## F-statistic: 4.671 on 3 and 44 DF,  p-value: 0.00643
```

This is not significant as the p-value is above our acceptable level of error (0.05). If the p-value would've been below our acceptable level of error, there would be a 111.61% decrease in the total number of visits for Google relative to what we would expect if Google followed the same trend as other search engines (Yahoo, Bing, Ask), after removing the ads.

```
# We would have to log transform total_visits
avg_visits <- data %>%
  group_by(treatment, post) %>%
  summarize(mean(log(1+total_visits)))

## `summarise()` has grouped output by 'treatment'. You can override using th
e
## `.groups` argument.

avg_visits

## # A tibble: 4 × 3
## # Groups:   treatment [2]
##   treatment  post `mean(log(1 + total_visits))`
##       <dbl> <dbl>                         <dbl>
## 1         0     0                          8.25
## 2         0     1                          9.37
## 3         1     0                          8.78
## 4         1     1                          8.78

Treatment_pre = 8.783
Treatment_post = 8.784
Control_pre = 8.253
Control_post = 9.370
diff_treat <- Treatment_post - Treatment_pre
diff_contr <- Control_post - Control_pre
diff_treat - diff_contr

## [1] -1.116

change = (Treatment_post - Treatment_pre)
pct_change = ((Treatment_post - Treatment_pre)/Treatment_pre)*100
pct_change

## [1] 0.01138563
```

As for calculating the difference in means and then looking at percent change, the overall visits across all platforms increased slightly before to after. The overall total visits for Google increased slightly from the pre-period to the post-period, as seen from the 1.13% increase in the sum of total visits (log transformed).

Difference in mean without controlling for treatment group can be misleading as you are ignoring the treatment vs control effect. Additionally, there may be confounding factors where other search engines have different trends than Google even before ad removal, meaning the average change across all platforms might not reflect the true impact of Google's ad removal.

Question D

Calculate the Difference-in-Differences.

We used plm() because we have *panel data* and want to control for both platform-specific and time-specific factors that could bias your estimate of the ad removal effect.

```r
library(plm)

## Warning: package 'plm' was built under R version 4.4.3

##
## Attaching package: 'plm'

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

summary(plm(log(1 + total_visits) ~ treatment*post,
            data = data,
            index = c("platform","week"),
            effect = 'twoway',
            model = 'within'))

## Twoways effects Within Model
##
## Call:
## plm(formula = log(1 + total_visits) ~ treatment * post, data = 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|)
## treatment:post -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 p-value is below our acceptable level of error (0.05), therefore there is indeed a change in the total number of visits for Google. There is a roughly 111.61% decrease in the total number of visits for Google relative to what we would expect if Google followed the same trend as other search engines (Yahoo, Bing, Ask), after removing the ads. A more precise estimate would be 67.24% drop in visits attributable to ad removal on Google.

In contrast, a simple pre-post comparison for Google alone shows almost no change (~0.01% increase). This large discrepancy shows why pre-post estimates can be misleading — they don't consider that visits might have naturally increased or decreased over time, regardless of the ads. By using other platforms as a comparison group, Difference-in-Differences isolates the true effect of removing the ads on Google's visits.

We ran a regression to allow for week-specific effects and to estimate a separate treatment effect for each week, rather than assuming a constant treatment effect across all weeks.

```
#Let's run a dynamic DiD model where we include dummies for all weeks
did_dyn <- lm(log(1 + total_visits) ~ treatment * factor(week),
    data = data)
summary(did_dyn)

## 
## Call:
## lm(formula = log(1 + total_visits) ~ treatment * factor(week),
##     data = data)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -0.75376 -0.14344  0.08617  0.30058  0.49340 
## 
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                  6.76890    0.33466  20.226  < 2e-16 ***
## treatment                    0.46600    0.66932   0.696 0.492973    
## factor(week)2                0.66132    0.47328   1.397 0.175106    
## factor(week)3                1.12151    0.47328   2.370 0.026189 *  
## factor(week)4                1.41778    0.47328   2.996 0.006270 ** 
## factor(week)5                1.63831    0.47328   3.462 0.002026 ** 
## factor(week)6                1.88266    0.47328   3.978 0.000557 ***
## factor(week)7                2.05138    0.47328   4.334 0.000226 ***
## factor(week)8                2.23224    0.47328   4.717 8.54e-05 ***
## factor(week)9                2.35675    0.47328   4.980 4.38e-05 ***
## factor(week)10               2.47779    0.47328   5.235 2.29e-05 ***
## factor(week)11               2.61642    0.47328   5.528 1.10e-05 ***
## factor(week)12               2.71183    0.47328   5.730 6.64e-06 ***
## treatment:factor(week)2     -0.06103    0.94656  -0.064 0.949123    
## treatment:factor(week)3      0.03454    0.94656   0.036 0.971191    
## treatment:factor(week)4      0.08156    0.94656   0.086 0.932047    
```

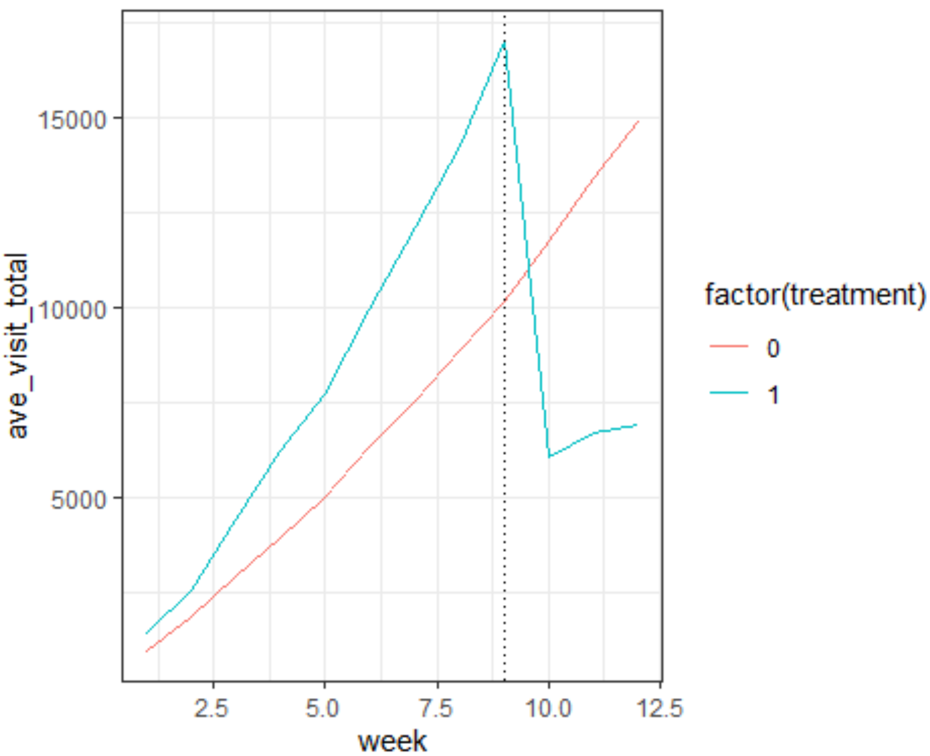
```
## treatment:factor(week)5   0.07381    0.94656   0.078 0.938491
## treatment:factor(week)6   0.08626    0.94656   0.091 0.928142
## treatment:factor(week)7   0.11567    0.94656   0.122 0.903757
## treatment:factor(week)8   0.09639    0.94656   0.102 0.919732
## treatment:factor(week)9   0.15027    0.94656   0.159 0.875195
## treatment:factor(week)10 -1.00338    0.94656  -1.060 0.299682
## treatment:factor(week)11 -1.04355    0.94656  -1.102 0.281194
## treatment:factor(week)12 -1.10890    0.94656  -1.172 0.252894
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5796 on 24 degrees of freedom
## Multiple R-squared:  0.7835, Adjusted R-squared:  0.576
## F-statistic: 3.776 on 23 and 24 DF,  p-value: 0.0009808
```

With this model we can see there is 100.33% deviation in week 10 when treatment begins. For Google compared to Bing/Yahoo/Ask, if they would've followed a similar trend, the deviation would be 100.33%. Additionally, week 11 and week 12 follow similar deviations, suggesting week 10 is not a one-off occasion.

```
library(ggplot2)
week_ave = data %>% group_by(week,treatment) %>% summarise(ave_visit_total =
mean(total_visits))
```

```
## `summarise()` has grouped output by 'week'. You can override using the
## `.groups` argument.
```

```
# plot for total visit
ggplot(week_ave, aes(x = week, y = ave_visit_total, color = factor(treatment)
)) +
  geom_line() +
  geom_vline(xintercept = 9, linetype='dotted') +
  xlim(1,12) +
  theme_bw()
```



The parallel trends assumption holds, meaning our difference-in-differences methodology is valid for estimating the causal effect of the treatment.

Question E

Given Your Treatment Effect Estimate, Fix Bob's ROI Calculation.

Bob's ROI really needs to consider ONLY the sponsored ads.As the ads follow a parallel trend, we can use week 9 to impute for proportion of sponsored ads. We use the below formula to calculate ROI.

ROI = Income - Cost / Cost

A = Additional clicks from sponsored ads

B = The number of sponsored clicks

C = Conversion Rate

D = Margin

E = Cost per click

Formula to calculate ROI = (A * B * C * D) - (B * E) / (B * E)

```
# Calculate the mean of sponsored ad visits for just google
data %>%
```

```r
  filter(treatment == 1, week ==9) %>%
  summarize(total_spons = sum(avg_spons))

##   total_spons
## 1       12681

# Set constants for ROI calculations
conversion_rate <- 0.12
margin <- 21
cost_per_click <- 0.60
# The number of sponsored clicks (B)
spons_week <- 12681
# Additional clicks from sponsored ads (A)
add_spons <- diff_treat - diff_contr
add_spons <- abs((exp(add_spons) - 1))
income <- add_spons*spons_week*conversion_rate*margin
total_cost <- spons_week*cost_per_click
ROI <- ((income - total_cost)/total_cost)* 100
ROI

## [1] 182.4132
```

Recommendations

1. One recommendation would be to run a controlled tests as this test was not controlled entirely. By running a controlled randomized test they can better understand the impact of ads.
2. Try running ads at a specific time. Branded ads might work better at certain times of the day, like in lunchtime or in the evening where people have more time to browse and shop.Therefore, the impact of ads should be measured on certain time of the day.
3. We would still recommend using the ads. While ads may still be effective, the ROI is likely lower than initially assumed. Instead of removing them altogether, the focus should shift to optimizing ad strategy and budget based on more realistic performance expectations.