

```

---
title: "GV300 Assignment 3: A/B Testing and Average Treatment Effects in Political Communications"
author: "2316631"
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
set.seed(2024)
library(ggplot2)
```

```

Overview

IMPORTANT You need to knit this file and submit the pdf to FASER as before.
 # In this assignment, you will work with a simulated A/B test run by a fictitious political party delivering messages on its official website.
 # The party wants to increase the probability that visitors sign up as supporters**.

To do this, they are going to test two homepage versions:

1. **Control homepage (Version A)**:
 Standard party messaging (policy highlights, leader photo, "Join us" button).
2. **Issue-focused homepage (Version B)**:
 Emphasizes a hot-button issue, such as immigration, important to the party's base, plus a "Join us" button.

Visitors are **randomly assigned** to Version A or Version B when they land on the website.

The outcome variable is:

- `support_signup = 1` if the website visitor signs up as a party-supporter
- `support_signup = 0` otherwise

In this assignment, you will:

1. Simulate the data for this A/B test
2. Explore and visualise the sign-up rates for each homepage
3. Estimate the Average Treatment Effect (ATE) using:
 - A difference-in-means calculation
 - A linear probability model (`lm`)
 - A logistic regression (`glm` with logit link)
4. Compare the models and interpret the results
5. Interpret heterogeneity by using an interaction
6. Reflect briefly on ethical issues in running experiments in politics

Simulating the A/B Test Data

We will simulate data for n = 1500 website visitors.

Assume:

- Baseline signup probability under the control homepage is **0.12**.
- The issue-focused homepage increases signup probability by **0.05**, to **0.17**.

Code: Simulate the experiment

```

```{r simulate-data}
TODO: Write code to:
1. Set n = 1500
2. Randomly assign each visitor to treatment (0 = control, 1 = issue-focused) with probability 0.5
3. Generate potential outcomes:
- Y0 ~ Bernoulli(0.12) for each visitor (if shown control)
- Y1 ~ Bernoulli(0.17) for each visitor (if shown treatment)
4. Create the realized outcome support_signup using:

```

```

support_signup = Y1 if treatment == 1, else Y0
5. Combine into a data frame called party_ab with columns:
visitor_id, treatment, support_signup
#
HINT: use rbinom() and ifelse()
#
1. Set sample size
n <- 1500 ## ADD n HERE

2. Random treatment assignment
treatment <- rbinom(n = n, size = 1, prob = 0.5) ## ADD COMMENT HERE

3. Potential outcomes

p_control <- 0.12
p_treated <- 0.17

Y0 <- rbinom(n = n, size = 1, prob = p_control) ## ADD COMMENT HERE
Y1 <- rbinom(n = n, size = 1, prob = p_treated) ## ADD COMMENT HERE

4. Realised outcome
afford <- ifelse(treatment == 1, Y1, Y0)

5. Data frame

party_ab <- data.frame(
 visitor_id = 1:n,
 treatment = treatment,
 support_signup = afford
)
```

## Inspect the data

```{r inspect-data}
head(party_ab)
summary(party_ab)
#How else might you inspect the data? Add here.
```

## Written question

In **2-4 sentences**, explain:

1. What does `treatment` represent in this experiment?  

ANSWER# treatment indicates whether a website visitor was shown the control homepage  

# (0) or the issue-focused homepage (1).

2. What are `Y0` and `Y1`, conceptually?  

ANSWER# Y0 and Y1 are potential outcomes: Y0 is whether a visitor would sign up  

# under the control homepage, and Y1 is whether they would sign up under the issue-focused homepage.

3. Why do we only observe one of `Y0` or `Y1` for each visitor? Why is this a problem for causality?  

ANSWER# We only observe one of Y0 or Y1 for each visitor because a person can only see one version of  

# the homepage. This is the fundamental problem of causal inference, as we never observe both outcomes for the same individual.

4. What does `support_signup` represent in terms of potential and realised outcomes?  

ANSWER# support_signup is the realised outcome: it equals Y1 for treated visitors and Y0 for control visitors.

---
## Signup Rates and Difference-in-Means ATE

We now estimate the average signup rate under each homepage version and compute the difference.

```

```

```{r group-means}
TODO: Compute the mean signup rate for control and treatment groups

mean_control <- mean(party_ab$support_signup[party_ab$treatment == 0])
mean_treated <- mean(party_ab$support_signup[party_ab$treatment == 1])

means_df <- data.frame(
 treatment = c("Control homepage", "Issue-focused homepage"),
 mean_signup = c(mean_control, mean_treated)
)```

Code: Group means and bar plot

```{r group-means-plot}
ggplot(means_df, aes(x = treatment, y = mean_signup, fill = treatment)) +
  geom_col(width = 0.6) +
  geom_text(
    aes(label = paste0(round(mean_signup * 100, 1), "%")),
    vjust = -0.5
  ) +
  scale_y_continuous(limits = c(0, 0.25), expand = expansion(mult = c(0, 0.05))) +
  scale_fill_manual(values = c("Control homepage" = "gray70",
                               "Issue-focused homepage" = "steelblue")) +
  labs(
    x = NULL,
    y = "Signup rate",
    title = "Supporter signup rate by homepage version"
  ) +
  theme_minimal() +
  theme(legend.position = "none")
```

Code: Difference-in-means ATE

```{r diff-in-means}
# TODO: Compute the difference-in-means ATE
ate_hat <- mean_treated - mean_control
```

Written questions

1. Interpret `mean_control` and `mean_treated` in plain language.
ANSWER# mean_control is the proportion of visitors who signed up when shown the standard homepage, while mean_treated is the proportion who signed up when shown the issue-focused homepage.

2. Convert `ate_hat` into percentage points . What does it mean, in words, for the party's digital strategy?
ANSWER# If ate_hat is around 0.05, this corresponds to a 5 percentage-point increase in sign-ups from using the issue-focused homepage. This suggests the issue-focused design is more effective at converting visitors into supporters.

3. Ignoring uncertainty for the moment, would this result encourage the party to adopt the issue-focused homepage? Why or why not?
ANSWER# Ignoring uncertainty, this result would encourage the party to adopt the issue-focused homepage, since it increases sign-ups without any obvious downside in this experiment.

4. Change two things about this plot to make it more interesting/readable
ANSWER# For example, the plot could use clearer colour contrasts, add confidence intervals, # include exact numeric labels, or remove unnecessary visual clutter to improve readability.
```

```

```

# Linear Probability Model (lm) vs logit (glm)

We now estimate the ATE using our workhorse regression model.

## Code: Linear probability model

```{r lm-model}
TODO: Fit a linear probability model (OLS)

lm_model <- lm(support_signup ~ treatment, data = party_ab)
summary(lm_model)
coef(lm_model)
```

## Written questions

1. What does the **intercept** in `lm_model` estimate in this context?
#ANSWER# The intercept estimates the average probability of signing up for
# visitors in the control group.

2. What does the **coefficient on `treatment`** estimate, and how does it relate to `ate_hat`?
#ANSWER# The coefficient on treatment estimates the causal effect of showing the issue-focused
# homepage on signup probability. It is equal to the difference-in-means ATE.

3. Check numerically that the `treatment` coefficient equals the difference in means from
Section 2. Are they the same?
#ANSWER# Yes, the coefficient on treatment is numerically identical to ate_hat,
# because with a binary treatment, OLS reproduces the difference in group means.

```

```

## Code: Logistic regression (logit)

```{r logit-model}
TODO: Fit a logistic regression (logit) model

logit_model <- glm(support_signup ~ treatment, data = party_ab, family = binomial(link =
"logit"))
summary(logit_model)
```

```

In the next section, we are going to go to some lengths to interpret these coefficients. Why do we have to do this?

```

#ANSWER#
# Logit coefficients are expressed in log-odds, which are not directly interpretable as
# changes in probabilities. This means we must transform the coefficients into predicted
# probabilities to understand the substantive effect of the treatment.

```

Code: Predicted probabilities under control and treatment

```

```{r logit-pred-probs}
TODO: Compute predicted probabilities from the logit model.

```

```

p_hat_control <- predict(
 logit_model,
 newdata = data.frame(treatment = 0),
 type = "response"
)

```

# WHAT DOES THE ABOVE DO?

```

#ANSWER#
#'Above' uses the fitted logistic regression model to predict the probability
that a visitor signs up if they are shown the control homepage (treatment = 0).
Setting type = "response" converts the model's log-odds output into an interpretable
probability between 0 and 1.

```

```

p_hat_treated <- predict(

```

```

logit_model,
newdata = data.frame(treatment = 1),
type = "response"
)

ADD COMMENT HERE. WHAT DOES THE ABOVE DO?
#ANSWER#
'Above' uses the fitted logistic regression model to predict the probability that
a visitor signs up when shown the issue-focused (treatment) homepage (treatment = 1).
The argument type = "response" converts the predicted log-odds into a probability
between 0 and 1, making the result easy to interpret.

p_hat_control
p_hat_treated

ate_logit <- p_hat_treated - p_hat_control
ate_logit

WHAT DOES THE ABOVE DO?

#ANSWER#
'Above' first displays the predicted probability of signing up under the control
homepage (p_hat_control) and under the issue-focused homepage (p_hat_treated).
It then calculates ate_logit by subtracting the control probability from the treatment
probability, which gives the average treatment effect implied by the logistic regression
model in probability terms.
```

## Written questions

1. Interpret `p_hat_control` and `p_hat_treated` in plain language.
#ANSWER# p_hat_control is the predicted probability that a visitor signs up when
# shown the control homepage, while p_hat_treated is the predicted probability under the issue-focused homepage.
2. Compare `ate_logit` to `ate_hat` and the `treatment` coefficient from `lm_model`. Are they
similar or different? Why might they be close in this simple A/B test?
#ANSWER# ate_logit is very similar to ate_hat and the OLS treatment coefficient because the
treatment effect is modest and the model is simple. In such cases,
# linear and nonlinear models often produce very similar estimated effects.

```
---```
Comparing lm and logit visually

We now visualise how the two models fit the data.

Code: Plot with fitted probabilities

```{r plot-fitted, fig.width=6, fig.height=4}

# Prepare a data frame for group-level summaries
group_df <- data.frame(
  treatment = c(0, 1),
  mean_signup = c(mean_control, mean_treated),
  model = "LM group means"
)

logit_df <- data.frame(
  treatment = c(0, 1),
  mean_signup = c(p_hat_control, p_hat_treated),
  model = "Logit predicted probs"
)

# Main ggplot
ggplot(party_ab, aes(x = factor(treatment), y = support_signup)) +
  geom_jitter(
    width = 0.1,
    height = 0,

```

```

alpha = 0.1,
color = "black"
) +
geom_line(
  data = group_df,
  aes(x = factor(treatment), y = mean_signup, group = 1, color = model),
  linewidth = 1
) +
geom_point(
  data = group_df,
  aes(x = factor(treatment), y = mean_signup, color = model),
  size = 3
) +
geom_point(
  data = logit_df,
  aes(x = factor(treatment), y = mean_signup, color = model),
  shape = 17,
  size = 3
) +
scale_x_discrete(labels = c("Control", "Treatment")) +
scale_y_continuous(limits = c(0, 1)) +
scale_color_manual(
  values = c(
    "LM group means" = "blue",
    "Logit predicted probs" = "red"
  )
) +
labs(
  x = "Treatment",
  y = "Support signup (0/1)",
  title = "Observed outcomes and fitted probabilities",
  color = NULL
) +
theme_minimal()
```

```

#INSERT VIS 2 HERE!!!!#  
## Written questions

You have been hired as an analyst for the political party and you have to present these findings to your employer. The people that you are giving the presentation to do not have much quantitative experience. How might you show the information above in a way that is easy to understand? Make sure that you can answer:

1. How different are the predicted sign-up probabilities from the two models for each group?  
#ANSWER# The predicted sign-up probabilities from the two models are very close for both the control and treatment groups.
2. Does either model seem clearly inappropriate here, or do they both tell a similar story?  
#ANSWER# Neither model appears inappropriate here; both tell a consistent story that the issue-focused homepage increases sign-ups.
3. Which model do you prefer and why?  
#ANSWER# The linear probability model may be preferred for its simplicity and ease of interpretation,  
# while the logit model is theoretically better suited for binary outcomes. In this case, both are acceptable.

---

# Heterogeneous Effects by Age or Interest (Interaction)

Now, we are going to add simulated covariates. The two variables we will make are age and high political interest.

Assume visitors are between 18 and 75, and age is roughly distributed normal around 40

```

```{r het-effects}
party_ab$age <- round(pmin(pmax(rnorm(nrow(party_ab), mean = 40, sd = 12), 18), 75)) # ADD COMMENT. WHAT DOES THIS DO?

```

```
# High political interest: 1 = high interest, 0 = low interest (about 40% high)
party_ab$high_interest <- rbinom(nrow(party_ab), size = 1, prob = 0.4)

head(party_ab[, c("visitor_id", "treatment", "support_signup", "age", "high_interest")])
```

Next, we are going to run an interaction with treatment and interest

```
```{r het-interest}
lm_het_interest <- lm(support_signup ~ treatment * high_interest, data = party_ab)
summary(lm_het_interest)
```

Call:

```
lm(formula = support_signup ~ treatment * high_interest, data = party_ab)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1610	-0.1597	-0.1364	-0.1061	0.8939

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.10609	0.01644	6.454	1.46e-10 ***
treatment	0.05364	0.02307	2.325	0.0202 *
high_interest	0.03027	0.02567	1.179	0.2385
treatment:high_interest	-0.02905	0.03648	-0.796	0.4260

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.346 on 1496 degrees of freedom

Multiple R-squared: 0.004553, Adjusted R-squared: 0.002557

F-statistic: 2.281 on 3 and 1496 DF, p-value: 0.07756

Questions:

1. What is the implied treatment effect by ``interest`` subgroup. Note that interest is coded 0,1

#ANSWER# The treatment effect for low-interest visitors (interest = 0) is given by the treatment coefficient,

# while for high-interest visitors it is the sum of the treatment and treatment:high\_interest coefficients.

2. What is the treatment effect when age = 0? Is this a meaningful baseline?

#ANSWER# The treatment effect when age = 0 is not meaningful, since age 0 is outside the support of the data.

# This highlights a limitation of interpreting regression intercepts mechanically.

3. Which variable gives you information of how the treatment effect changes for each additional year of age? How do you interpret that?

#ANSWER# The coefficient on the interaction between treatment and age shows how the treatment effect changes with each additional year of age.

# A positive coefficient would imply stronger effects among older visitors.

4. Can you plot the treatment effects at specific ages (e.g., 25, 40, 60)?

#ANSWER# Yes, this can be done by calculating predicted values at ages such as 25, 40, and 60 and plotting them separately.

# Ethical Reflection on Political A/B Testing

So far, we have treated this like a standard A/B test. But here the product is a political message, and the outcome is political support.

In less than 200 words, reflect on \*\*ethical considerations\*\* that arise when a political party runs A/B tests like this on its website visitors.

You don't need a right answer; the goal is to think critically and connect the statistical tools (ATEs, models) with their real-world political and ethical implications.

#ANSWER#

# Political A/B testing raises distinct ethical concerns because the object being optimised is not a neutral consumer choice, but political attitudes and behaviour.

# Website visitors are typically unaware that they are participating in experiments designed to influence their political support, which challenges principles of informed consent and transparency.

# Unlike commercial advertising, political messaging plays a direct role in shaping democratic participation, making covert experimentation more ethically sensitive.

# There is also a risk that optimisation prioritises messages that maximise engagement or sign-ups rather than those that promote accurate information or deliberative debate.

# If emotionally charged or polarising issues systematically outperform more balanced messages, A/B testing may incentivise parties to amplify division rather than democratic understanding.

# From a distributional perspective, these techniques can further skew political influence if certain demographic or interest groups are more responsive to targeted messaging.

# Although the statistical methods themselves are neutral, their strategic use can subtly reshape political discourse.

# Analysts therefore have an ethical responsibility to evaluate not only the effectiveness of political interventions, but also their broader implications for democratic norms and accountability.