Take Home Exercise 1

Replicating an Experimental Study (THE 1)

Nicolas Waser

2024-11-14

Exercise 1: *Summary of the Paper and Main Findings Provide a brief summary of the paper you are replicating. Describe the main findings, especially those related to Study 1.*

Disclaimer. For this exercise I made use of the following AI Generation tools: Deepl.com for translation purposes, the integrated Copilot tool in R Studio for generating code and text and the free versions of OpenAI's GPT-4o Mini, Anthropics Claude 3.5 Sonnet and Mistral's Codestral AI model for generating code and assisting me in solving the tasks.

The authors posit that according to the Homonationalist thesis, support for LGBT+ issues among anti-immigration nativists and right-wing populists is not entirely genuine (organic liberalism thesis) and is instead used strategically (instrumental liberalism thesis) as a tool to justify further exclusion of immigrants, especially Muslims, who are perceived to be more homophobic, than the average western citizen. The rational behind this could be summed up as: "The enemy of my enemy is my friend" or more drastically: "I hate them (Muslims) more than I hate the other group (LGBT+ people)". In order to test this hypothesis, the authors conducted two survey experiments in the UK and Spain. The dependent variable is "Support for LGBT+ inclusive education in school" and was measured by means of survey response on an order randomized scale from 0 to 11, but then for the sake of simplicity were turned into a dichotomous variable for support (0 or 1).

In the UK, for study 1, the experimental research design randomly assigned 1200 participants into a control group and a treatment group. Those in the control group were exposed to text and pictures of nondenominational people with conventional British names protesting against LGBT+ inclusive education in schools. The treatment group was exposed to the same pictures and text, but the names and pictures were changed to appear Muslim. The authors expected that individuals who harbor negative predispositions toward immigration in the treatment group, will display more socially liberal preferences toward LGBT+ education than individuals with comparable anti-immigrant sentiment in the control group. Indeed, the results showed that (with an average treatment effect of 9.7% points or a 22% increase against the baseline) respondents with an anti-immigrant disposition in the treatment group showed increased support for LGBT+ inclusive education in schools than comparable individuals in the control group. On the other hand, among those with positively predisposed toward immigration, treatment exhibits no significant effect below the 5% significance level. But overall, it can be said, that support for LGBT+ inclusive education is positively correlated with liberal attitudes towards immigrants.

In Study 2, conducted in Spain, chosen because of generally more positive attitude levels towards immigration and LGBT+ issues than in the UK, and a right-wing populist party (VOX), that unlike its counter-part in the UK (UKIP), is not following a homonationalist strategy and is more openly outspoken against LGBT+ issues. They found that regardless of immigration sentiments, treatment resulted in an 10-11% increase in support for LGBT+ inclusive education in schools. However, the effect was stronger among those with anti-immigrant sentiments (21% vs. 14%), when compared against the baseline. Furthermore, the authors also looked at shifts in citizens pride in "Western Culture". Only among individuals with nativist's views could they reasonably demonstrate increased levels of "pride in Western Culture" in the treatment group.

With these findings in mind, the authors argue that the homonationalist thesis is correct and that the increased levels of support for LGBT+ issues among the wider general public in western countries is not entirely genuine, but also in part the result of a strategic move by right-wing populists to deradicalize their image by defining themselves by "who they are not" rather than "who they are". They further speculate that as soon as the political context changes and LGBT+ rights are at stake (but absent of the perceived threat of "immigrants"), the support for LGBT+ issues among right-wing populists will decrease.

Exercise 2: *Data Preparation and Exploration Use the data file study1_data.csv to begin the replication process. Identify and describe the experimental variables (i.e., treatment and immigration attitudes) and provide visualizations of their distribution. Then, select up to four covariates (e.g., gender, age, etc.) and plot their distribution too. If necessary, clean or transform variables. Document any changes.*

From looking at the csv table for Study1 as well as looking at the study1.R script, one can deduce which variables are likely of interest for the replication. As a first step, I need to find the dependent variable (DV) support for LGBT+ inclusive education, which according to the text had a 0-11 range with a randomized scale order that was transformed to 0 and 1 for easier interpretation. Secondly, I need to find the experimental variables for treatment and immigration attitudes. The treatment variable concerns the exposure to Muslim names and pictures and is most likely operationalized as a binary variable (0,1). The variable for immigration attitudes is likely to be operationalized as an ordinal scale (0-10) or as a binary variable.

A look at the Csv table revealed the following variables as of interest for the replication:

- Treatment variable: outcome_treat (categorical with NAs), treatment (ordinal)
- Immigration attitudes: imm_1, imm_2, (both ordinal) imm1mean (continuous, same value), immbelow, imm3 (both ordinal)

- Dependent Variable: support (ordinal), support2 (categorical)

By also looking at the study1.R script in the code chunk below, I can see that the following variables are of interest for the replication, while the rest can likely be ignored:

```
#df <-df%>%
  #mutate(treat= as.factor(treatment),
         #treatnum= as.numeric(treatment),
         #[...]
         #immbelow= as.factor(immbelow),
         #imm3= as.factor(imm3),

##Figure 3##
#model1<- glm (support ~ treat*imm_1, data=df, family="binomial")
#summ(model1, robust=TRUE)
#df$predictINT1<-predict(model1, df, type="response")

#treat <- subset(df, treatnum==1)
#control <- subset(df, treatnum==0)
#proimm <- subset(df, immbelow==0)
#noproimm <- subset(df, immbelow==1)

##FIGURE 4##
#modelsub1<- lm (support ~ treat, data=proimm)
#summ(modelsub1, robust=TRUE)
#proimm$predictedb<-predict(modelsub1, proimm)

#modelsub2<- lm (support ~ treat, data=noproimm)
#summ(modelsub2, robust=TRUE)
#noproimm$predictedb<-predict(modelsub2, noproimm)

#treatsub1 <- subset(proimm, treatnum==1)
#controlsub1 <- subset(proimm, treatnum==0)
#treatsub2 <- subset(noproimm, treatnum==1)
#controlsub2 <- subset(noproimm, treatnum==0)
```

- Treatment variable: treatment is transformed twofold -> treat (converted to cat. variable for model transformation), treatnum (converted to numeric binary variable for graph)
- Immigration attitudes: immbelow (cat., converted to binary variable for graph), imm_1 (ordinal, 0-10 with 10 standing for maximal positive attitude towards immigrants)
- Dependent Variable: support (ordinary)

From this I conclude that the following variables are of interest for the replication:

- Treatment variable operationalization: treatment (0,1)
- Immigrant attitudes operationalization: imm_1 (ordinal scale 0-10) ((or immbelow, for binary distribution Pro-/AntiImmigration))
- Dependent Variable operationalization: support (0,1)

Below I will provide visualizations of the distribution of these variables. Let's start with the treatment variable. But first, I need to load the data sets and the necessary packages. I will then transform the treatment variable for better visualization and provide a table and a bar plot of the distribution of the treatment variable.

```
study1 <-
  read.csv("/Users/nicolaswaser/New-project-GitHub-first/R/Take-Home-Ex.-1/Input Data/study1_
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.1     v tibble    3.2.1
v lubridate 1.9.3     v tidyr     1.3.1
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom
```

```
library(ggplot2)
#is.factor(study1$treatment)
study1 <-study1%>%
  mutate(treatment= as.factor(treatment))
#is.factor(study1$treatment)
```

Variable distribution in the form of a table:

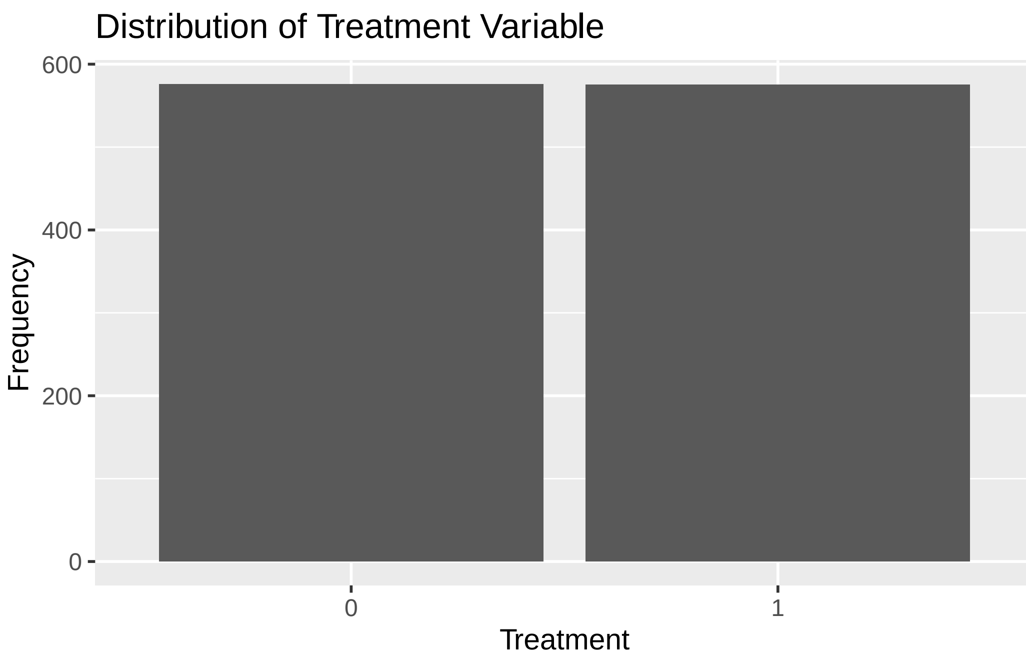
```
table(study1$treatment)
```

```

  0   1
576 575
```

Variable distribution in the form of a bar plot:

```
ggplot(study1, aes(x=treatment)) + geom_bar() +
  labs(title="Distribution of Treatment Variable",
       x="Treatment", y="Frequency")
```



Unsurprisingly, the treatment variable is distributed evenly across the two groups, as it should be in a randomized experiment.