Lab 3: Success of Leader Assassination as a Natural Experiment

September 14, 2020

One longstanding debate in the study of international relations concerns the question of whether individual political leaders can make a difference. Some emphasize that leaders with different ideologies and personalities can significantly affect the course of a nation. Others argue that political leaders are severely constrained by historical and institutional forces. Did individuals like Hitler, Mao, Roosevelt, and Churchill make a big difference? The difficulty of empirically testing these arguments stems from the fact that the change of leadership is not random and there are many confounding factors to be adjusted for.

In this exercise, we again consider the leader assassination attempt data set. The study authors call a their study design a *natural experiment* where they argue that the success or failure of assassination attempts may be assumed to be as good as random.

This exercise is based on: Jones, Benjamin F, and Benjamin A Olken. 2009. Hit or Miss? The Effect of Assassinations on Institutions and War. (http://dx.doi.org/10.1257/mac.1.2.55) *American Economic Journal: Macroeconomics* 1(2): 5587.

Each observation of the CSV data set `leaders.csv` contains information about an assassination attempt. The variables are:

Name	Description
`country`	The name of the country
`year`	Year of assassination attempt
`leadername`	Name of leader who was targeted
`age`	Age of the targeted leader
`politybefore`	Average polity score during the 3 year period prior to the attempt
`polityafter`	Average polity score during the 3 year period after the attempt
`civilwarbefore`	1 if country is in civil war during the 3 year period prior to the attempt, or 0
`civilwarafter`	1 if country is in civil war during the 3 year period after the attempt, or 0
`interwarbefore`	1 if country is in international war during the 3 year period prior to the attempt, or 0
`interwarafter`	1 if country is in international war during the 3 year period after the attempt, or 0
`result`	Result of the assassination attempt, one of 10 categories described below
`dies`	1 if the assassination attempt is successful, 0 if not

The `polity` variable represents the so-called *polity score* from the Polity Project. The Polity Project systematically documents and quantifies the regime types of all countries in the world from 1800. The polity score is a 21-point scale ranging from -10 (hereditary monarchy) to 10 (consolidated democracy).

The `result` variable is a 10 category factor variable describing the result of each assassination attempt.

Question 0

Load the libraries and data you'll need today.

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------
## v ggplot2 3.3.2     v purrr   0.3.4
```

```
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(knitr)
leaders <- read_csv("data/leaders.csv")
```

```
## Parsed with column specification:
## cols(
##   year = col_double(),
##   country = col_character(),
##   leadername = col_character(),
##   age = col_double(),
##   politybefore = col_double(),
##   polityafter = col_double(),
##   interwarbefore = col_double(),
##   interwarafter = col_double(),
##   civilwarbefore = col_double(),
##   civilwarafter = col_double(),
##   result = col_character(),
##   treatment = col_double(),
##   dies = col_double()
## )
```

Let's first check for missing data. How many missing observations do we have for each variable in the data set?

```
leaders %>%
  summarize_all(funs(sum(is.na(.))))
```

```
## Warning: `funs()` is deprecated as of dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_warnings()` to see where this warning was generated.
```

```
## # A tibble: 1 x 13
##    year country leadername   age politybefore polityafter interwarbefore
##   <int>   <int>      <int> <int>        <int>       <int>          <int>
## 1     0       0          0     0            0           0              0
## # ... with 6 more variables: interwarafter <int>, civilwarbefore <int>,
## #   civilwarafter <int>, result <int>, treatment <int>, dies <int>
```

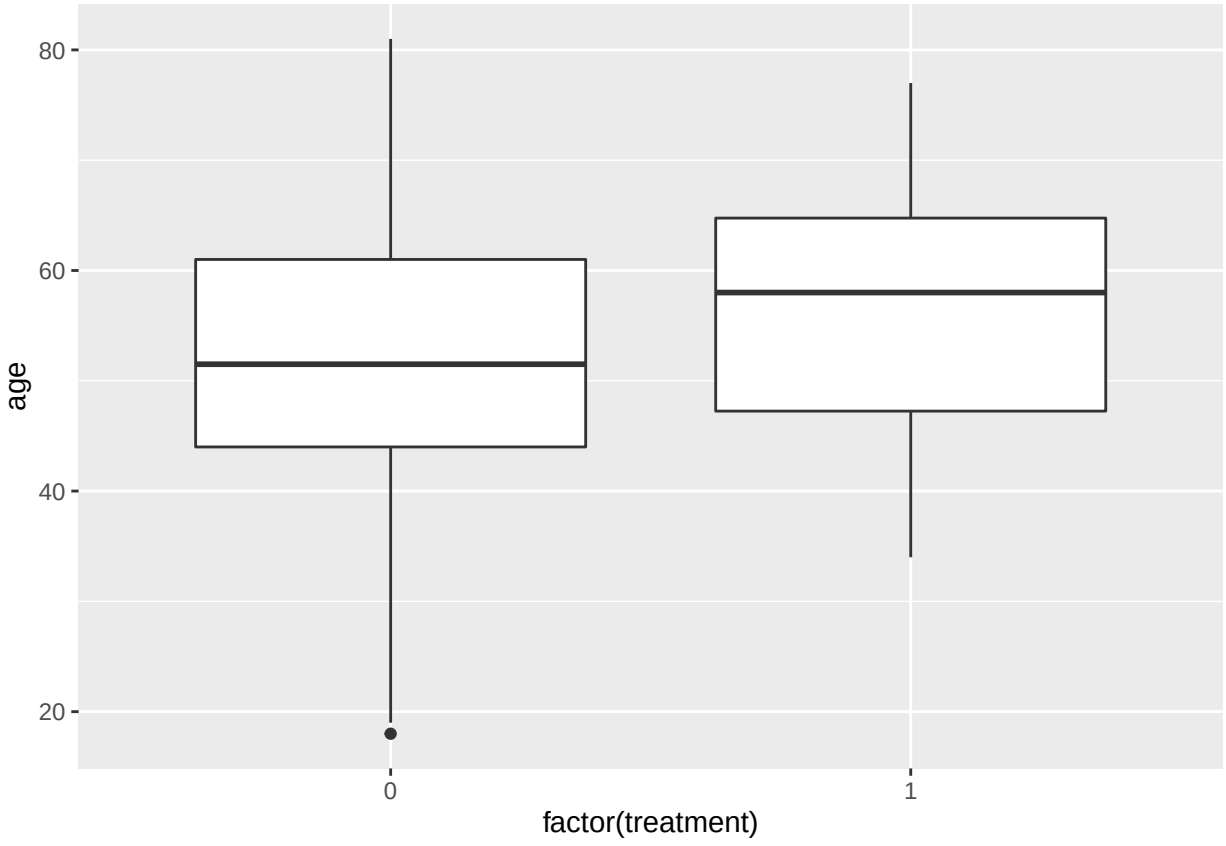
Let's now examine how we recoded `result` to `dies`:

```
table(leaders$result, leaders$dies)
```

```
##
##                                                0  1
##   dies between a day and a week                0  2
##   dies between a week and a month              0  2
##   dies within a day after the attack           0 46
##   dies, timing unknown                         0  4
##   hospitalization but no permanent disability 20  0
##   not wounded                                 96  0
##   plot stopped                                40  0
##   survives but wounded severely                1  0
##   survives, whether wounded unknown           14  0
##   wounded lightly                             25  0
```

Now we'll check for balance on pre-treatment covariates between treated countries (ie ones where the leader died in the assassination attempt), and untreated countries (where the leader survived the assassination attempt). In particular, let's examine the boxplots of treated versus untreated countries using the `ggplot2` library, startiong by looking at the `age` variable, the age of the leader experiencing the assisination attempt.

```
leaders %>%
  ggplot(aes(x = factor(treatment), y = age)) +
  geom_boxplot()
```



What are the median ages of the two groups?

What are the 1st quartile, 3rd quartile, IQR, and range for age across the two groups? Are there any outliers in the plots? Are the distributions of age symmetric or skewed? If skewed, which direction?

What are the mean and standard deviations of age for treated countries (ie ones where the leader died in the assassination attempt), and untreated countries (where the leader survived the assassination attempt).

How do the two groups compare across the age? Are you concerned about covariate balance on age? Why may covariate balance (or lack of balance) on age matter?

Question 1

Compare covariate balance using a boxplot across the `politybefore` variable. Measure the spread of the policy score by creating a summary table that contains mean, standard deviation, median, IQR, Min, and Max.

Describe the boxplot and discuss what it tells you about covariate balance and why it matters.

Note: in Lab 1 we calculated the distribution of this variable using the `summary` command: today we are comparing the distributions visually and considering measures of central tendency.

Your Answer here:

The medians, max and min (range) are the same for both the box plots, but the average for the two boxpl

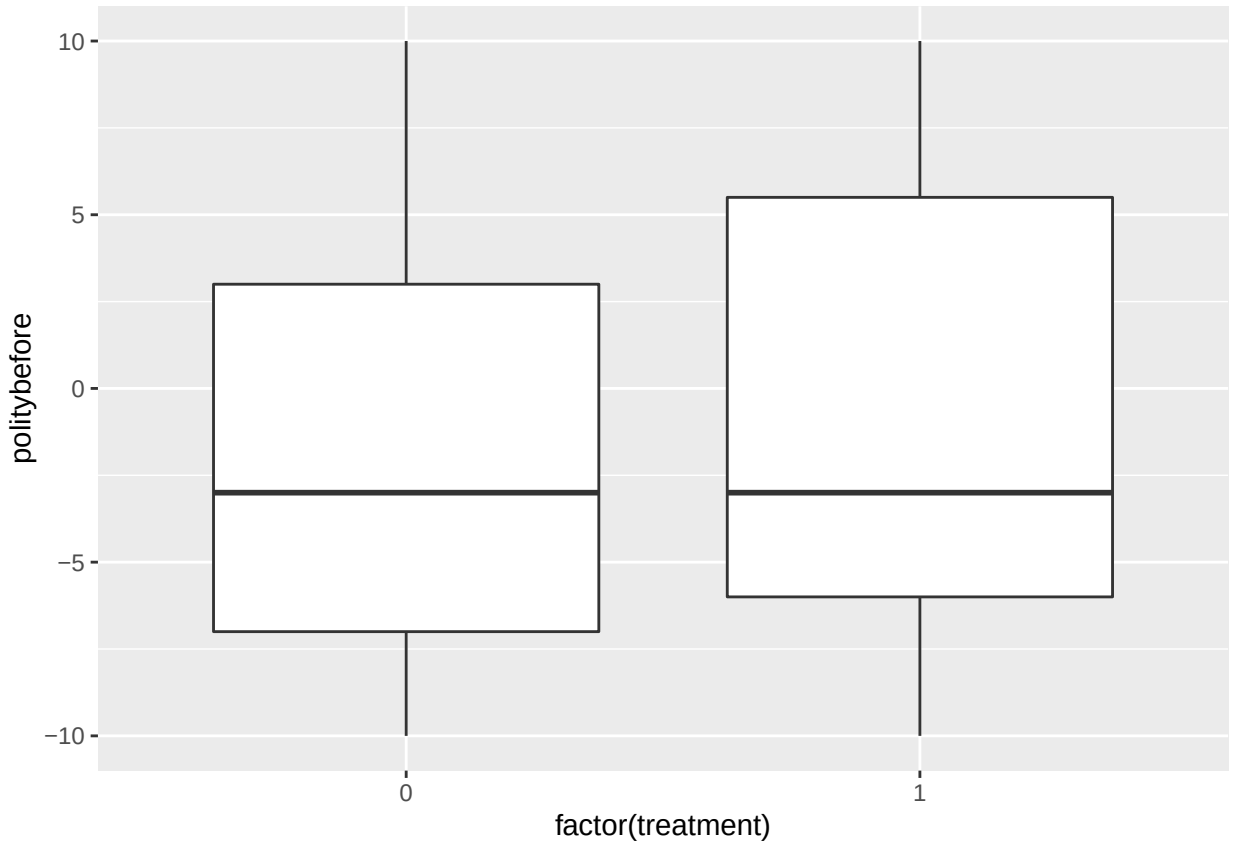
```
dt_sum <- leaders %>%
  group_by(treatment) %>%
  mutate(mean_polBefore = mean(politybefore),
         sd_polBefore = sd(politybefore),
         median_polBefore = median(politybefore),
         iqr_polBefore = IQR(politybefore),
         min_polBefore = min(politybefore),
         max_polBefore = max(politybefore)) %>%
  distinct(mean_polBefore, sd_polBefore, median_polBefore,iqr_polBefore, min_polBefore, max_polBefore)

  kable(dt_sum, caption = "Summary Statistics for Polity Before Scores Across Treatment Groups")
```

Table 2: Summary Statistics for Polity Before Scores Across Treatment Groups

treatment	mean_polBefore	sd_polBefore	median_polBefore	iqr_polBefore	min_polBefore	max_polBefore
0	-1.7431973	6.417421	-3	10.0	-10	10
1	-0.7037037	6.541808	-3	11.5	-10	10

```
leaders %>%
  ggplot(aes(x = factor(treatment), y = politybefore)) +
  geom_boxplot()
```


10
5
0
−5
−10
politybefore
0
1
factor(treatment)

Question 2

Compare covariate balance across `interwarbefore` and `civilwarbefore` by calculating the mean within each group (we won't make boxplots here because these are dichotomous variables). What do these numbers represent? Hint: you may wish to use the `group_by` and `summarize` commands.

```
dt_war_sum <- leaders %>%
          group_by(treatment) %>%
          mutate(mean_intWarBefore = mean(interwarbefore),
                 mean_civWarBefore = mean(civilwarbefore)) %>%
          distinct(mean_intWarBefore, mean_civWarBefore)

kable(dt_war_sum, caption = "Covariate balance - prior war instances")
```

Table 3: Covariate balance - prior war instances

treatment	mean_intWarBefore	mean_civWarBefore
0	0.1785714	0.2346939
1	0.2222222	0.1481481

Your Answer here:

```
These numbers represent the two categories, 'interwarbefore` and `civilwarbefore` and these categories r

International war -
we see covariate imbalance here through the difference in means. The proportion of countries with a leac

Civil war -
we see covariate imbalance here through the difference in means. The proportion of countries with a leac
```

Question 3

```
mean(leaders$polityafter[leaders$treatment == 1]) - mean(leaders$politybefore[leaders$treatment == 1])
```

```
## [1] -0.05864198
```

What is the difference in mean polity scores before and after an assassination attempt among countries with successful assassination attempts (ie where `dies == 1`)?

What study design is this? In this study design how is the average missing counterfactual for the intervention countries estimated?

Using this study design, what is the answer to the specific causal question? What confounder(s) do you think are most likely to bias this estimate of the causal effect?

Your Answer here:

-0.058 is the difference in polity scores before and after the assassination attempts.

This study design here is a pre post study. In this study design, the average missing counterfactual is calculated by estimating the mean polity score before the intervention occurred. Using this study design, the answer to the specific causal question is the estimate we have just calculated since it gives the estimated causal effect of the treatment on the same population. The confounder most likely to bias this estimate is the age since it is the one variable that differs among participants from the same population.

Question 4

What is the difference in mean post-assassination attempt polity score between treated and control countries? What study design is this? In this study design how is the average missing counterfactual for the intervention countries estimated? Using this study design, what is the answer to the specific causal question? What confounder(s) are most likely to bias this estimate of the causal effect?

```
mean(leaders$polityafter[leaders$treatment == 1]) - mean(leaders$polityafter[leaders$treatment == 0])
```

```
## [1] 1.132212
```

Your Answer here:

This is a post-only study design. In this study design, the average missing counterfactual for the intervention can be estimated by calculating the mean polity after score for countries that did not experience the intervention (i.e. where leaders survived the assassination attempt). The answer to the specific causal question is whether the intervention occurring results in different outcomes for the intervention and control group. The confounders most likely to bias this estimate of the causal effect are `interwarbefore` and `civilwarbefore`.

Question 5

Now calculate the difference-in-differences estimate; ie, the average difference in polity scores before and after an assination attempt for the treated observations minus the average difference in polity scores before and after an assination attempt for the control observations. Using a difference-in-differences study design what are the potential outcomes? Using this study design, what is the answer to the specific causal question? How does this estimate compare to your previous results?

```
(mean(leaders$polityafter[leaders$treatment == 1]) - mean(leaders$politybefore[leaders$treatment == 1]))
```

```
## [1] 0.09271857
```

Your Answer here:

The potential outcomes for the difference in difference: • What the change in polity would be for the country if the leader was assassinated? • What the change in polity would be for the country if the leader survived the assassination attempt?

```
The impact of the assination on polity, on average the assisination group was 9.27% higher.
The difference in the scores is greater in the Difference in Difference
study than the the Pre-Post and the Post study designs.
```