POL304 Assignment 1

Talia Fabregas

January 25, 2024

The code and data associated with the completion of this assignment can be accessed via my GitHub repository. The assignment was completed using R Programming Language and the tidyverse Wickham et al. (2019), janitor Firke (2023) and dplyr Wickham et al. (2023) for sub-setting data, and knitr Xie (2014) for the tables. The analysis is based on the research paper titled "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya" Lyall (2009).

```
## loading packages
# install.packages("tidyverse")
# install.packages("janitor")
# install.packages("dplyr")
# install.packages("knitr")
library(tidyverse)
library(janitor)
library(dplyr)
library(knitr)

## load the data
chechen_data <- read_csv("chechen.csv")

# time-varying confounder is controlled, as only 90 days before and after shelling are obs

## a sample from the data
kable(head(chechen_data))
```

village	groznyy	fire	deaths	preattack	postattack
Elistanzhi	0	0	NA	4	3
Malye Shuani	0	1	0	0	1
Belgatoy	0	1	34	1	0
Oktya'brskoe	0	0	NA	0	0

village	groznyy	fire	deaths	preattack	postattack
Chiri-Yurt	0	0	NA	4	5
Gansolchu	0	1	0	0	0

Question 1

The research question that this article aims to explore is:

What is the relationship between a state's use of indiscriminate violence and insurgent attacks?

Question 2

It is widely theorized and agreed upon amongst academics that indiscriminate violence is counterproductive in the reduction of insurgent violence because it gives insurgents the ability to convince the civilian population that the risks of joining them and undermining the state is worth it, might offer them protection, and appears less dangerous than non-participation, despite the potential for serious consequences Lyall (2009) .

However, this might not be the case if states use indiscriminate violence when the insurgency is weak and unable to convince civilians that joining them offers any protection from violence. This leads to the theory that a state's use of indiscriminate violence can reduce insurgency attacks by depleting the insurgency's war resources and leading civilians to believe that the insurgency does not offer them any safety and is responsible for provoking the state's violent response, which can promote collective action against the insurgency between civilians and the state.

Question 3

The central research hypothesis is:

A state's use of indiscriminate violence can reduce insurgent attacks under certain circumstances.

Question 4

Table 2 shows the number of villages that were shelled at least once by Russian Artillery and the number of villages that were not shelled during the period of interest. A total of 125 unique villages appear in the data set. Of those 125 villages, 75 were shelled at least once, and 50 were not shelled.

```r
# count how many unique villages in the data set
num_villages <- length(unique(chechen_data$village))

# subset to only include the villages that were shelled at least once
shelled_observations<- chechen_data |>
  filter(fire == 1)

# count the number of unique villages that were shelled
num_shelled = length(unique(shelled_observations$village))

# get the number of unique villages that were not shelled
num_not_shelled = num_villages - num_shelled

# create a data frame
q4 <- data.frame(VillagesShelled = num_shelled,
                 VillagesNotShelled = num_not_shelled)

# print the table for reference
kable(q4)
```

Table 2: The Number of Villages in the Data Set that Were and Were Not Shelled

VillagesShelled	VillagesNotShelled
75	50

Question 5

Based on the results of the cross-sectional analysis conducted, I would conclude that indiscriminate violence reduces insurgent attacks. Table 3 shows the average number of insurgent attacks for shelled vs. non-shelled villages, and we can see that the difference in the average number of insurgent attacks between the shelled and non-shelled villages is ~0.55.

```r
# Subset the data into observations that do not describe a shelled village
not_shelled_observations <- chechen_data |>
  filter(fire == 0)

# Already have the shelled observations from Question 4

# shelled villages
Y_shelled_after = mean(shelled_observations$postattack)
```

```
# non-shelled villages
Y_nonshelled_after = mean(not_shelled_observations$postattack)

q5 <- data.frame(
  MeanShelledAfter = Y_shelled_after,
  MeanNonShelledAfter = Y_nonshelled_after,
  # compute the mean number of insurgent attacks after the fires
  SATE = Y_shelled_after - Y_nonshelled_after
)

kable(q5)
```

Table 3: The Average Number of Insurgent Attacks for Shelled vs. Non-Shelled Villages

MeanShelledAfter	MeanNonShelledAfter	SATE
1.496855	2.050314	-0.5534591

Question 6

As seen in Table 4, the average number of insurgent attacks in the pre-shelling period (a pre-treatment variable for each of the treatment and control groups) for observations describing a shelling and for observations that do not are nearly identical, at 2.11 and 2.15, respectively. There is very little difference between this pre-treatment variable in the treatment and control groups, which suggests that the comparison in the previous question exhibits strong internal validity. Therefore, it is reasonable to believe that the treatment and control groups are identical on average in terms of all confounders (Chyzh, Lecture 2, Slide 15).

```
# Compute the mean number of insurgent attacks for observations that describe a shelled vi
# and observations that do not during only the pre-shelling periods
avg_pre_shelled <- mean(shelled_observations$preattack)
avg_pre_not_shelled <- mean(not_shelled_observations$preattack)

# Create a data frame
q6 <- data.frame(
  Shelled = avg_pre_shelled,
  NotShelled = avg_pre_not_shelled
)

# Print the table using knitr
```

```r
kable(q6)
```

Table 4: Average Number of Insurgent Attacks during Pre-Shelling Periods in Shelled Village vs Non-Shelled Village Observations

Shelled	NotShelled
2.113208	2.150943

Question 7

Among observations in the data set that describe villages that were shelled, the number of insurgent attacks decreased after being fired on. Table 5 shows the mean number of insurgent attacks in the pre- and post-shelling periods for observations that were shelled; the mean number of insurgent attacks decreased from ~2.11 before being fired on to ~1.5 after being fired on.

```r
# Compare villages that were shelled before and after
q7 <- data.frame(
  mean_before=mean(shelled_observations$preattack),
  mean_after=mean(shelled_observations$postattack)
)

# Print the table for reference
kable(q7)
```

Table 5: Mean Number of Insurgent Attacks in Villages that Were Shelled, Before and After Being Fired On

mean_before	mean_after
2.113208	1.496855

Question 8

This analysis supports the claim that indiscriminate violence reduces insurgency attacks. Table 6 shows the mean difference in the difference between insurgency attacks in the pre- and post-shelling periods for villages that were shelled and for villages that were not shelled. The mean difference between the pre- and post-shelling periods for villages that were shelled is ~0.616, while the mean difference in the pre- and post-shelling periods for villages that were not shelled is ~0.1. Therefore, the difference-in-difference analysis shows that villages that were shelled saw a greater reduction in insurgency attacks between the pre- and post-shelling periods than villages that were not shelled, which supports the claim that indiscriminate violence reduces insurgency attacks.

```r
# mean shelled before - mean shelled after
shelled_difference = mean(shelled_observations$preattack) -
  mean(shelled_observations$postattack)

# mean not shelled before - mean not shelled after
not_shelled_difference = mean(not_shelled_observations$preattack) -
  mean(not_shelled_observations$postattack)

# Compute the sample average treatment effect for the treated
satt = shelled_difference - not_shelled_difference

# create a data frame with these numbers
q8 <- data.frame(
  MeanDiffShelled = shelled_difference,
  MeanDiffNotShelled = not_shelled_difference,
  SATT = satt
)

# print the table to use for reference
kable(q8)
```

Table 6: Mean Difference in Difference Between Pre- and Post-Shelling Villages for Villages that were Shelled vs Villages that were not

MeanDiffShelled	MeanDiffNotShelled	SATT
0.6163522	0.1006289	0.5157233

Question 9

Which of the three designs–cross-sectional, temporal, or difference-in-difference–is the most appropriate, given whether or not the data meet the assumptions for each?

A cross-sectional design assumes that the treatment and control groups are comparable. In question 5 (shown in Table 3) we found that the average number of insurgent attacks in the 90 days before is nearly identical for both the treatment group and the control group. This demonstrates that both groups are comparable, as they had extremely similar average levels of insurgent violence beforehand.

A temporal design assumes that there is no time-varying confounding. This data set only counts the number of insurgent attacks in the 90 days immediately preceding and the 90 days immediately following the attack for both the treatment group (shelled village observations) and the control group (non-shelled village observations). The exact same periods of time are examined for both the treatment and control group, therefore the data does meet the assumptions for a temporal design.

A difference-in-difference design assumes that there is a parallel time trend, which covers the assumption that there is no group-specific or time-varying confounding. This data set meets the assumptions for this, as shown in question 5, Table 3, with the nearly identical levels of insurgent violence in the 90 days before the attack, as well as the controlled observation period that only takes into account the 90 days immediately following the attack. Therefore, difference-in-difference is the most appropriate for this data set.

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Lyall, Jason. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya." *Journal of Conflict Resolution* 53 (3).

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Xie, Yihui. 2014. *Knitr: A Comprehensive Tool for Reproducible Research in r.*