Problem Set 3

Your Name - Net ID - Section Number

Due Dec 1, 2023

This homework must be turned in on Brightspace by Dec. 1, 2023. It must be your own work, and your own work only – you must not copy anyone's work, or allow anyone to copy yours. This extends to writing code. You may consult with others, but when you write up, you must do so alone.

Your homework submission must be written and submitted using Rmarkdown. No handwritten solutions will be accepted. **No zip files will be accepted. Make sure we can read each line of code in the pdf document.** You should submit the following:

1. A compiled PDF file named yourNetID_solutions.pdf containing your solutions to the problems.

2. A .Rmd file containing the code and text used to produce your compiled pdf named your-NetID_solutions.Rmd.

Note that math can be typeset in Rmarkdown in the same way as Latex. Please make sure your answers are clearly structured in the Rmarkdown file:

1. Label each question part

2. Do not include written answers as code comments.

3. The code used to obtain the answer for each question part should accompany the written answer. Comment your code!

Question 1 (Total: 100)

Does US military assistance strenghen or further weaken fragile and conflict-affected foreign governments? Aid may bolster state capacity and suppress violence from nonstate actors such as paramilitary groups. On the other hand, aid may be diverted to those same violent groups. To answer the question, Dube and Naidu (2015)(https://www.journals.uchicago.edu/doi/10.1086/679021?mobileUi=0) leverage changes in the allocation of US military aid to Colombian military bases. They test whether Colombian municipailites in which military bases are located have more or less paramilitary violence when the level of U.S. miliary aid increases, relative to Colombian municipalities in which miliary bases are not located.

For this problem, you will need the 'bases_replication_file.dta' file. The variables you will need are:

- parattq - DV here is paramilitary attacks
- bases6 - indicator variable whether or not there is a base in the municipality
- lrmilnar col - (logged) U.S. military and narcotics aid to Colombia
- bases6xlrmilnar col - the treatment i.e., the interaction between the level of U.S. military and narcotics aid and whether or not there is a base in the municipality
- lnnewpop - is log of population

Part a (60 points)

The treatment in this case is a continuous 'intensity' variable that changes over time. The authors use the interaction between the level of U.S. military and narcotics aid and whether a base exists in a municipality. How many units are in the 'control' group (no bases)? Does the bases variable change over time or is it a unit-constant factor? How about the logged military aid variable, does it change across units for a given year? What do the authors seem to be assuming about how military aid is allocated?

```r
library(tidyverse)
library(haven)
library(estimatr) # for lm with robust se : ?lm_robust()

# Load bases data
bases <- haven::read_dta("bases_replication_final.dta")

# How many observations are in the ``no bases group"


## How about each of them?


## How many municipalities do we have

```

Part b (20 points)

The authors use a common empirical strategy called two-way fixed effects to estimate the average treatment effect of military aid. The model they estimate includes fixed effects for both time periods and units (and includes logged population as an additional covariate):

$$Y_{it} = \gamma_t + \alpha_i + \tau D_{it} + \beta X_{it} + \epsilon_{it}$$

What assumptions are the authors making in order to identify the treatment effect of military aid?

Part c (20 points)

Using the two-way fixed effects estimator, estimate the effect of U.S. military and narcotics aid on the number of paramilitary attacks, including log of population as a covariate. The two sets of fixed effects are for municipality (municipality) and year (year). Cluster your standard errors at the unit level (see the cluster argument in lm_robust. Report a 95% confidence interval for your estimate and interpret your results.

```
#?lm_robust (set se_type to "CR0")
# Fit Regression using lm_robust

# Summarize
```