PLSC 30600: Problem Set 3

[YOUR NAME]

February 13, 2023

> This problem set is due at **11:59 pm on Tuesday, February 27th**.
>
> Please upload your solutions as a .pdf file saved as "Yourlastname_Yourfirstinitial_pset3.pdf". In addition, an electronic copy of your .Rmd file (saved as "Yourlastname_Yourfirstinitial_pset3.Rmd") must be submitted to the course website at the same time. We should be able to run your code without error messages. In order to receive credit, homework submissions must be substantially started and all work must be shown. Late assignments will not be accepted.

Problem 1

This problem will have you replicate and analyze the results from Moser and Voena's 2012 AER paper on the impact of the World War I "Trading with the Enemy Act" on U.S. domestic invention. The full citation is below

> Moser, P., & Voena, A. (2012). Compulsory licensing: Evidence from the trading with the enemy act. American Economic Review, 102(1), 396-427.

The premise of the study is to evaluate the effect that "compulsory licensing" policy – that is, policies that permit domestic firms to violate foreign patents and produce foreign inventions without needing to obtain a license from the owner of the foreign patent – have on domestic invention. Does access to foreign inventions make domestic firms more innovative? The authors leverage an exogenous event in U.S. licensing policy that arose from World War I – the 1917 "Trading with the Enemy Act" (TWEA) which permitted U.S. firms to violate patents owned by enemy-country firms. This had the consequence of effectively licensing all patents from German-owned firms to U.S. firms after 1918 (that is, from 1919 onward), allowing them to produce these inventions without paying for a license from the German-owned company.

The authors look specifically at domestic innovation and patent activity in the organic chemicals sector. They note that only some of the sub-classes of organic chemicals (as defined by the US Patent Office) received any compulsory licenses under the Trading with the Enemy Act while others did not. They leverage this variation in exposure to the "treatment" of compulsory licensing to implement a differences-in-differences design looking at domestic firm patent activity in each of these sub-classes (comparing sub-classes that were exposed to compulsory licensing to those that were unexposed).

The dataset is `chem_patents_maindataset.dta` – the code below will load it.

```
library(tidyverse)
# Read in the Moser and Voena (2012) dataset
chem <- haven::read_dta("chem_patents_maindataset.dta")
```

The unit of the dataset is the sub-class/year (471,120 observations) of 7248 US Patent and Trademark Office (USPTO) patent sub-classes over 65 years.

The relevant variables are

- `uspto_class` - USPTO Patent Sub-Class (unit)
- `grntyr` - Year of observation (year)

- `count_usa` - Count of patents granted to US-owned firms in the year
- `count_france` - Count of patents granted to French-owned firms in the year
- `count_for` - Count of patents granted to foreign-owned (non-US) firms in the year
- `treat` - Treatment indicator – Whether the patent sub-class received any German patents under TWEA (after 1918 when the policy went into effect) (Note that this is not an indicator for the overall treatment *group* (whether the unit *ever* received treatment) – it is only 1 after 1918 for units that receive treatment but is still 0 for those "treated" units prior to the initiation of treatment)

Part A

If you try to use a two-way fixed effects estimator on the dataset as it is, it will likely freeze up your computer as this is a *very large* dataset. We'll instead first aggregate the data in a way that will let you use a simple first-differences estimator to estimate the treatment effect.

Generate a point estimate for the average treatment effect of receiving treatment on the average annual count of US patents using a first-differences estimator with the pre-treatment "period" averaging over the years 1875-1918 and the post-treatment "period" averaging over 1919-1939. You should aggregate your data such that the outcome is the post-/pre- difference in the outcome (preferably using `tidyverse` functions like `group_by` and `summarize` or even `pivot_wider`) and each row is a USPTO patent sub-class (rather than a sub-class/year observation) and use a difference-in-means estimator with the differenced outcome.

Provide a 95% confidence interval and interpret your point estimate. Do we reject the null of no treatment effect at the $\alpha = .05$ level?

Part B

A colleague suggests that you should instead just compare the average differences in the count of US patents in the post-1918 period between exposed and unexposed sub-classes to estimate the treatment effect. Based on what we observe in the pre-1919 period, is ignorability of the treatment likely to hold under this strategy? Discuss why or why not.

Part C

The authors implement a test of their differences-in-differences identification assumptions by also estimating the effect the Trading with the Enemy Act on patents granted by French firms, which the authors note "could not license enemy patents under the TWEA." Describe what sort of a diagnostic strategy this is. What do the authors expect to find if their parallel trends assumption holds?

Estimate the effect of TWEA exposure on the count of French firm patents using a first-differences estimator and provide a a 95% confidence interval. Are the results consistent with what the authors expect if their design assumptions hold?

Part D

We might be concerned that there are differential trends in pre-treatment patenting between those sub-classes exposed to the treatment and those exposed to control. Estimate the difference in the trend in US patents between exposed and unexposed sub-classes from 1918 to 1917, 1916, 1915, and 1914 (four estimates in total: 1918-1917, 1918-1916, 1918-1915, 1918-1914). Provide a 95% robust confidence interval for each of these estimates and interpret your results. Do we reject the null that any of these differ from 0 (at $\alpha = .05$)? If the outcome trends were evolving in parallel between the treated and control groups, what would we expect these estimates to be? What do your results suggest for the validity of the parallel trends assumption?

Part E

The authors adjust for covariates out of concern for possible parallel trends violations. One possible confounder that might be driving a parallel trends violation is the overall amount of foreign patenting in the sub-class and its change over time – reflecting general technological differences that might differ between the patent sub-classes. Since the treatment does not affect the amount of foreign patenting, this is a valid control.

Create a variable for the change between the post- and pre-treatment count of foreign patents in the USPTO subclass. Use the Abadie (2005) weighting estimator to estimate the ATT under a conditional parallel trends assumption (fit a model for the propensity score using this covariate). Do we reject the null of no treatment effect at the $\alpha = .05$ level? Compare your results to your estimate from Question A and discuss why they might differ.

Problem 2

Consider the standard instrumental variables set-up in the treatment non-compliance setting, Z_i denotes the assignment to treatment, D_i is the actual treatment that is taken and Y_i is the outcome of interest. Assume that all of the instrumental variables assumptions hold (treatment is ignorable, exclusion restriction, non-zero first stage, monotonicity/no-defiers).

The specific example you'll examine in this problem is the JOBS II randomized trial which evaluated the effect of a job training program on reducing depression among those who recently experienced job loss (among other health outcomes). This specific dataset on 502 "high-risk" individuals from this experiment comes from:

> Little, Roderick J., and Linda HY Yau. "Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model." Psychological Methods 3.2 (1998): 147.

The code below will load the data into R

```
## Load the JOBS II dataset
jobs2 <- read_table2("wjobs.tab", na= ".")
## Make the treatment and instrument variable
jobs2$Z <- jobs2$Tx
jobs2$D <- as.integer(jobs2$Tx == 1 & jobs2$c1 == 1)
```

The relevant columns are

- `Z` - Assignment to the job training program
- `D` - Actual participation in job training
- `depress` - Change in depression from baseline (higher values = more depression)
- `risk` - Pre-treatment mental health risk score
- `educ` - Number of years of education completed
- `age` - Age
- `single` - Marital status: single

Part A

Examine the data and explain why monotonicity **is guaranteed** to hold for this particular design.

Part B

Using the Wald estimator, estimate the local average treatment effect of participation in the job training program among the compliers. Assuming asymptotic normality, provide a 95% confidence interval and conduct a hypothesis test for the null of no LATE at $\alpha = .05$.

Part C

Estimate the first stage effect of assignment to training on participation in the program. Would you consider assignment to the job program to be a strong instrument for participation?

Part D

Which subset of your sample consists exclusively of never-takers (think in terms of combinations of values of Z_i and D_i)? Suppose you were to take the mean of one of your covariates among this group – explain why this would identify the mean of the covariate among the never-takers.

Part E

Using the group you identified in D, examine the pre-treatment covariate averages among the never-treated respondents. Compare these to the overall means in the sample. Discuss whether there are any notable differences and what this would imply for the representativeness of the LATE from Part B.

Problem 3

In "The Long-Term Impact of Mobilization and Repression on Political Trust" (2021), Desposato et. al. examine the impact of exposure to the 1989 student protest movement in China on present day trust in government. The paper uses a survey of Beijing residents who would have been enrolled in college around the time of the protests. The researchers compare the present day attitudes of those who began college in 1985-1988 and would have had direct personal exposure to the movement to those who enrolled in college in the fall of 1989 and would not have been exposed due to the post-Tiananmen crackdown.

To address the possibility of cohort-specific differences, the researchers use a fuzzy regression discontinuity design to estimate the effect of exposure vs. non-exposure by leveraging the fact that enrollment decisions are partly driven by a cut-off date. Students born after September 1, 1970 would be on track to enroll in the fall of 1989 (or later) and miss exposure to the protest movement while those born just before would have enrolled in the fall of 1988 (or earlier) and have been exposed to the movement. However, exposure in this case is **fuzzy** as not all students who would be eligible to enroll based on the birthday cut-off would have enrolled as expected (some students may start early, some may delay, etc...).

The paper claims that individuals who attended college during the protest movement exhibit less trust in the central government.

The original source for this data is

> Desposato, Scott W., Gang Wang, and Jason Y. Wu. "The long-term impact of mobilization and repression on political trust." *Comparative Political Studies* 54.14 (2021): 2447-2474.

The code below will read in the dataset in as `protest`. Please note that you will have to drop any observations with missing data in the relevant variables.

```r
load("DWW_CPS.RData")
protest <- data
```

The relevant variables are:

- `X_c`: The centered difference between respondent's birthday and the enrollment cut-off of *September 1, 1970* (in years, continuous). Values greater than zero denote birthdays above the cut-off. Values below zero denote birthdays below the cut-off.
- `T`: Exposure to the 1989 student protest movement.
- `TrustCent`: Trust in the Central Government (0-10 index) High values denote greater trust
- `TrustProv`: Trust in the Provincial Government (0-10 index) High values denote greater trust
- `TrustLocal`: Trust in the Local Government (0-10 index) High values denote greater trust

Part A

Estimate the "intent-to-treat" effect of being born just before vs. just after the September 1, 1970 birthday cut-off on trust in the central government using a local linear regression with a uniform kernel. Use a bandwidth of $h = 2$ to the left and right of the cut-point. Provide a 95% confidence interval and discuss your results. Compare your regression discontinuity estimates to the simple difference in average central government trust between respondents exposed to the protest movement and respondents who were not exposed to the protest movement.

Part B

Estimate the first stage effect of being born just before vs. just after the September 1, 1970 birthday cut-off on exposure to the 1989 protest movement using a local linear regression with a uniform kernel. Use a bandwidth of $h = 2$ to the left and right of the cut-point. Provide a 95% confidence interval and discuss your results. Is the birthday cut-off a strong instrument for exposure to the protest movement?

Part C

Now use a bandwidth of $h = 4$. Provide a 95% confidence interval and compare your results to Part B.

Part D

Make two binned scatterplots to visualize the first stage and overlay the regression estimates - one for your results from Part B and one for your results from Part C. Compare the two plots and discuss which of the two regressions provides a better estimate of the CEF and why.

Do you think this fuzzy RD is a valid design for identifying the impact of exposure to the 1989 student protest movement on present-day attitudes?