**PP249: Problem Set 4**

**Due December 2nd, 2022**

As usual, you will be asked to type in your answers on Gradescope. However, please include a pdf submission that contains your work to produce those answers. (You will upload this pdf at the very end of Gradescope).

# Question 1 – Instrumental Variables [Lectures 11 & 15 ]

We are interested in evaluating the effect of an anonymous peer support program on whether frontline workers resign from their position. 1,000 frontline workers were randomly assigned either to be provided access to the anonymous peer support program or not to be admitted, with 500 assigned to each condition. Let *T*i = 1 if the individual *i* was assigned to the treatment and *T*i = 0 if they were not. One year after the program start date, the researchers acquired administrative data to verify whether the frontline workers resigned (*Y*i = 1) or not (*Y*i = 0). The workers’ participation in the program was also recorded, as well as record of any external peer support program participation. A new variable, *D*, was created with *D*i = 1 if the worker ever engaged in a peer support program and *D*i = 0 if they did not. We now have three binary variables: *D*, *T*, and *Y*, with 8 possible combinations of values.

## **Question 1.1 (see lecture)**

First, we are going to calculate the proportion of compliers, always-takers, and never-takers. The table below summarizes initial assignments and actual participation into the program:

|  |  |  |  |
| --- | --- | --- | --- |
| T (assignment) | D (actual participation) | N | P(Y=1): proportion of frontline workers resigning |
| T = 0 | D = 0 | 375 | 0.52 |
| T = 0 | D = 1 | 125 | 0.60 |
| T = 1 | D = 0 | 50 | 0.80 |
| T = 1 | D = 1 | 450 | 0.64 |

What are the possible compliance types for each of these groups? (choose all that apply)

(hint: there are at least two possible types within each group)

T = 0 and D = 0: Complier, Never-Taker, Always-Taker, Defier

T = 0 and D = 1: Complier, Never-Taker, Always-Taker, Defier

T = 1 and D = 0: Complier, Never-Taker, Always-Taker, Defier

T = 1 and D = 1: Complier, Never-Taker, Always-Taker, Defier

**Question 1.2**

Assume that there are no defiers (that is, assume monotonicity). What is the proportion of always-takers among T = 0? That is, what is P(always-taker | T = 0)? (write as a decimal between 0 and 1 on Gradescope)

Similarly, what is the proportion of never-takers among T=1? That is, what is P(never-taker | T=1)?  
(write as a decimal between 0 and 1 on Gradescope)

**Question 1.3**

Since T is randomly assigned, the proportion of always-takers is the same regardless of T=0 or T=1. Likewise, the proportion of never-takers is the same regardless of T=1 or T=0. If each person is either an always-taker, never taker, or a complier, what can we infer about the proportion of compliers?   
(write as a decimal between 0 and 1 on Gradescope)

## **Question 1.4**

Calculate the intent-to-treat effect (ITT) of program assignment (T) on resignment (Y) in percentage point terms (only write the number on GradeScope)

*(hint: ITT = E(Y|T=1) - E(Y|T=0), which is the difference in resignment probability for those who were VS were not assigned to the program)*

What do we need to assume about T (the instrument for actual participation D) for this ITT estimate to be unbiased? (choose one)

* Independence assumption (also called “Exogeneity’’)
* Strong first stage (sometimes called “relevancy’’)
* Exclusion
* Monotonicity (no defiers)

## **Question 1.5**

The first stage effect describes how effective instrument T is at influencing actual program participation D. Calculate the first stage effect (in percentage point terms. Only type in the number on Gradescope)

*(hint: FS = E(D|T=1) - E(D|T=0))*

What do we need to assume about T for this first stage estimate to be unbiased? [Hint: monotonicity defined in lecture 11 p.38]

* Independence assumption (also called ``Exogeneity’’)
* Strong first stage (sometimes called ``relevancy’’)
* Exclusion
* Monotonicity

## **Question 1.6**

Calculate the local average treatment effect (LATE) using the ITT and FS above.

**Question 1.7**

LATE is the causal effect for which of the following groups?

* Complier
* Never-Taker
* Always-Taker
* Defier

What assumptions are necessary for LATE to be a valid estimate of causal effects? (select all that apply) *[Hint: SUTVA defined in lecture 12 p.6]*

* Independence assumption (also called ``Exogeneity’’)
* Strong first stage (sometimes called ``relevancy’’)
* Exclusion
* Monotonicity
* SUTVA (i.e. no spillovers)

# Question 2 – Event Study Designs [Lectures 18, 19 & 21]

A longstanding question in education policy is the impact of unionization on teacher salaries. A series of papers examine state-level adoption of mandatory collective bargaining laws for teachers, which essentially mandate that school districts recognize local teachers’ unions (see, for instance, [this terrific paper](https://e50c00f1-c826-48c4-98ce-bef51415dedc.filesusr.com/ugd/a763a0_3703edea889f4b19bd0686a293369b28.pdf) from Agustina Paglayan). Below is a map of states adopting this type of policy over time, where states in gray never adopt mandatory collective bargaining laws:



The plots below give event study estimates for two important outcomes: overall education expenditures and teacher salary (all in logs of 2010 $’s).[[1]](#footnote-1)



## **Question 2.1**

How do I interpret the estimates to the left of zero? To the right? Discuss the conditions necessary for these estimates to be valid (i.e. what are identifying assumptions?)

**Question 2.2**

An economist looks at these figures and makes the following comment: “Estimates for teacher salary look reasonable, but estimated causal effects on expenditures don’t seem valid”.

Why is the economist concerned about those estimates? *(hint: pre-trends)*

**Question 2.3**

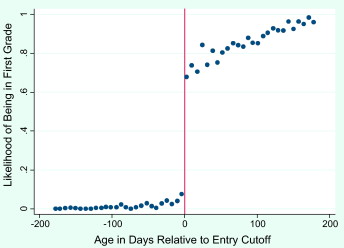
A seminal study in this area estimated that the impact of mandatory collective bargaining laws increased teacher salaries by 12 percent (i.e. +0.12 log points). Are estimates in these figures consistent with that conclusion? Why or why not?

# Question 3 –Regression Discontinuity Designs [Lecture 23]

Several researchers have attempted to use school starting ages as a source of regression discontinuities. Typically, states set calendar dates (e.g., September 1) and decree that all students who are 5 by September 1 of a given calendar year should start kindergarten that year.

Imagine that you have data on all children born in 1992. You compute the number of days that each child was born before or after the school entry cutoff – e.g., a child born on August 15 would be assigned a +16, as she was born 16 days before the cutoff, while a child born on September 15 would receive a -15. You then measure the grade that each child was in during the 1998-1999 school year. A child with a relative birthdate of 0 would have turned six on September 1, 1998, so we expect that children with relative birthdates greater than zero should be in first grade in this year while those with relative birthdates less than zero should be in kindergarten.

You divide the sample into bins of seven relative birthdates (so that one bin contains all children with relative birthdates of 0 to 6, another 7 to 13, and so on; other bins contain -1 to -7, -8 to -14, and so on). You then compute the fraction of children in each bin who are in first grade in 1998-1999. The graph below shows the results:



**Question 3.1**

Explain what this figure means. What can we infer about the enforcement of the state starting age law? What can you infer about the proportion of compliers, defiers, always takers, and never takers for the policy of encouraging students to start kindergarten by the year in which they will be 5 on September 1? Assume that there are no defiers in this context (i.e. assume monotonicity)

**Question 3.2**

Suppose that a researcher gathers data from the same cohort in 2017, the year that the individuals turned 25, and shows that the average number of years of completed education is noticeably higher for those with positive relative ages than for those with negative relative ages. Describe how this fact could be used to identify the effect of completed education on earnings as an adult.

**Question 3.3**

Sorting around the discontinuity is a recognized threat to a regression discontinuity design. How likely is this to be a problem here? How could you tell if it were?

# Question 4

For question 4, you can pick which of the two below exercises to do – 1) either further conceptual review – i.e., does not involve a data exercise; or 2) a data exercise, specific utilizing our canonical difference-in-difference. It’s up to you: **please only turn in one** – we won’t grade both! Not to say that you can’t do both for practice if you’d like to – feel free! The solutions will have answers to both exercises.

# Question 4A – Conceptual Review [Lectures 18, 19 & 21]

Below are a series of conceptual questions regarding program evaluation. For each prompt, select the most appropriate answer and include 1-2 sentences to justify your choice.

Note: Assume when we say DiD here, we're talking about the standard 2x2 (two time points for treatment and control group each, one pre-period and one post-period for each group)

**Question 4A.1**

Let’s say in 2027, federal law now states marijuana becomes legal for recreational use. You’re interested in focusing on the impact of this federal change on the states in which recreational marijuana was not legal (so all of these states are experiencing the policy change in 2027). You want to measure the impact of this change on state revenue. You have access to yearly data from 2020-2034.

* Pre-Post
* Interrupted Time Series
* Difference-in-Difference
* Fixed Effects/Event Study

**Question 4A.2**

Some hypothetical local government institution wanted to help invigorate their employees, so they considered multiple options before landing on “let’s just buy donuts for everyone.” Let’s pretend we observed their level of happiness at two time points: once before knowing about/consuming the donuts, and once immediately after. Assume all employees had donut-access (and these donuts were super-donuts, so okay to be safely and happily consumed by anyone with gluten or other allergies), and now we want to measure whether their employees’ happiness improved post-donut.

* Pre-Post
* Interrupted Time Series
* Difference-in-Difference
* Fixed Effects/Event Study

**Question 4A.3**

Let's say that state A implements a cigarette tax and state B does not. Our outcome of interest is cigarette sales in the state for that year. We have county level data from each state, but we only have data from the year prior and after the policy change. They are both in the same region, and for the sake of argument, we can say that Professor Trelawney informs us that our parallel trends assumption holds through working that divination skill.

* Pre-Post
* Interrupted Time Series
* Difference-in-Difference
* Fixed Effects/Event Study

**Question 4A.4** [Lectures 11 & 15]

In a setting with two-sided noncompliance, I am the effect of random assignment on those who, if assigned to treatment, will participate in the program and, if not assigned to treatment, will *not* participate in the program. Who am I?

1. Intention to Treat (ITT)
2. Local Average Treatment Effect (LATE)
3. Treatment on the Treated (TOT)
4. B and C

**Question 4A.5** *[Lectures 23]*

True or false: one of the key assumptions of regression discontinuity is that our running variable is continuous at the threshold.

1. True
2. False

**Question 4A.6** *[Lectures 11 & 15]*

What is the name of the assumption that states that an instrument should be related to the outcome only “through” the instrumented variable?

1. Monotonicity
2. SUTVA (i.e. no spillovers)
3. Strong first stage
4. Exclusion restriction

# Question 4B: Data Exercise for Diff-in-Diff [Lectures 18, 19 & 21].

**If you already answered Question 4A, there is no need to turn in this question. If you plan on doing this exercise (Question 4B), then there is no need to turn in Question 4A. Only choose one to submit** between Question 4A or 4B!

For this exercise, you’ll be analyzing data from the study, *Community-based Crisis Response: Evidence from Sierra Leone’s Ebola Outbreak[[2]](#footnote-2),* from Christensen and colleagues (2020). For their study, the authors analyzed data on recent outbreaks of Ebola in West Africa, focusing on the impact of Community Care Centers on total and confirmed cases.

You’ll have an Excel workbook, *community\_care.xlsx* that will have two sheets that you will need to use for this section. The first sheet is called “plots” which will be the sheet you use to make your line plots. The second sheet is called “analysis” which you will use to analyze data and replicate the results from [this paper](https://escholarship.org/content/qt6p15s2p7/qt6p15s2p7.pdf)!

## **Question 4B.1**

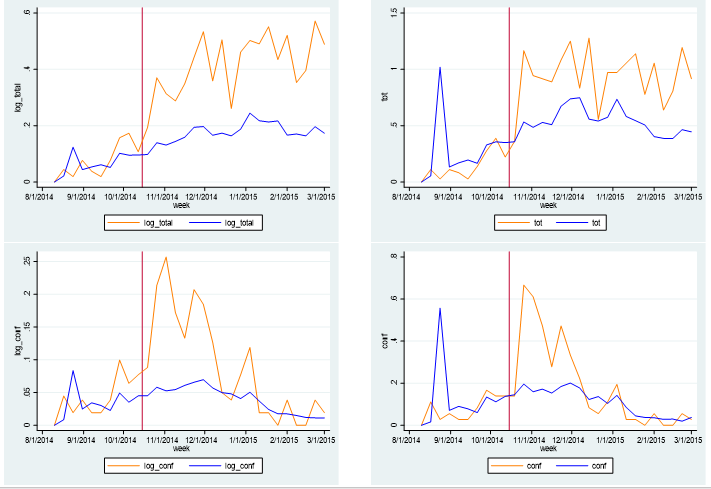
Construct four plots for each of the following outcomes, each having two lines for each group (one for treatment and one for control). Include a vertical line indicating the treatment date (Oct 15, 2014):

1) total cases (*total* in dataset)

2) confirmed cases (*confirmed* in dataset)

3) the natural logarithm of total cases (*log\_total* in dataset)

4) the natural logarithm of confirmed cases (*log\_conf* in dataset)



## **Question 4B.2**

We will estimate several difference-in-difference models to replicate results in the paper. Import the “analysis” sheet into Stata or R, and follow the instructions below to implement DiD regressions.

Fill in the table below with DiD estimates and their standard errors.

* There are four outcomes: total cases (*total*), confirmed cases (*confirmed*), log total cases (*log\_total*), and log confirmed cases (*log\_conf*)
* For each outcome, you will estimate two specifications:
  + Basic DiD specification (regress outcomes on treatment dummy, post dummy, treatment X post interaction)
  + DiD specification with fixed effects for section\_code and week.
  + For both specifications, cluster standard errors at the section\_code level
* *Hint: Think carefully about collinearity between treat and section\_code, and between post and week. Should the treatment dummy and post dummy be in the specification when including section\_code FE and week FE? (note: FE is the same as including a set of dummies)*
* Stata Tips:
  + Use “xtset” and “xtreg” to run regressions with fixed effects.
  + Cluster standard errors are specified using “vce(cluster var).”
  + Google these commands to learn how to use them
* R Tips:
  + You can use the following libraries: lfe, estimatr, sandwich, lmtest.
  + You’ll also need tidyverse, lubridate, and readxl.
  + For basic DiD specifications with clustered SE, you can use lm() with vcovCL() and coeftest()
  + For DiD with fixed effects, you can use felm()

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Outcome**  **(Specification)** | **Total Cases  (No Fixed Effects)** | **Total Cases  (Fixed Effects)** | **Confirmed Cases  (No Fixed Effects)** | **Confirmed Cases  (Fixed Effects)** |
| *Estimate* |  |  |  |  |
| *Standard Error (clustered on section\_code)* |  |  |  |  |
| **Outcome**  **(Specification)** | **Log Total Cases (No Fixed Effects)** | **Log Total Cases (Fixed Effects)** | **Log Confirmed Cases  (No Fixed Effects)** | **Log Confirmed Cases  (Fixed Effects)** |
| *Estimate* |  |  |  |  |
| *Standard Error (clustered on section\_code)* |  |  |  |  |

**Question 4B.3**

Upload your Stata or R code for the plots and regression table here (file upload on Grade Scope)

## **Question 4B.4**

Interpret your above results – does it seem that locations with Community Care Centers had differential reporting of cases over time, compared to locations without Community Care Centers?

## **Question 4B.5**

It’s important to check our placebo estimates! The implementation date was October 15th. In a subsequent analysis, the authors estimated an effect of +4.8% for October 7th (i.e. a week before the implementation). In other words, the DiD estimate for the pre-treatment period was 0.048, where we estimate differences between CCC and non-CCC locations in the weeks leading up to the implementation of CCCs. What does this mean?

1. These are essentially the “repeated 2x2 difference-in-differences” estimators discussed in lecture, but this type of event study plot is standard regardless of the specific estimation method. [↑](#footnote-ref-1)
2. Christensen, Darin, Dube, Oeindrila, Haushofer, Johannes, Siddiqi, Bilal, and Voors, Maarten. Data and Code for: Community-Based Crisis Response: Evidence from Sierra Leone’s Ebola Outbreak. Nashville, TN: American Economic Association [publisher], 2020. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2020-10-08. https://doi.org/10.3886/E120764V1 [↑](#footnote-ref-2)