Problem Set 4
ECO3121 - Fall 2024

Due 3 PM, December 11, 2024
No late submission is allowed
Please combine your answer, Stata code and requested output in **one** pdf file and upload it to Blackboard

Question 1

A researcher is interested in analyzing the effect of a free fertilizer policy on crop yield. He has a panel data set for 1000 villages in rural China over 2016 to 2019 in which the data for the average crop yield (Y_{it}) in village (entity) i in (time) year t and the indicator for whether the village participate in this free fertilizer program (X_{it}) are available.

1. He considers the following panel model:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$$

where α_i are i.i.d. unobserved random variables and α_i is correlated with X_{it}. (a) Explain how the researcher should estimate β_1 with the data he has. (b) Provide a factor that is modeled by α_i. (2 points)

> **Ans:**
> (a) Use the entity-demean estimator.
>
> (b) Any time-invariant village (entity)-level factors can be modeled by α_i, for example, the latitude and longitude of the village.

2. He considers the following panel model:

$$Y_{it} = \beta_1 X_{it} + \gamma_t + u_{it}$$

where γ_t are i.i.d. unobserved random variables and γ_t is correlated with X_{it}. (a) Explain how the researcher should estimate β_1 with the data he has. (b) Provide a factor that is modelled by γ_t. (2 points)

> **Ans:**
> (a) Use the time-demean estimator.
>
> (b) Any entity-invariant factors can be modeled by γ_t, for example, the trade war between the US and China.

3. He considers the following panel model:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \gamma_t + u_{it}$$

What is the advantage of this model compared to the above two? (2 points)

Ans:
It further controls for both any time-invariant and time-invariant fators.

4. The fixed effects estimator of β_1 in (3) can be obtained by applying two-way demean on this model. The first is entity-demean ignoring the time fixed-effects followed by the time-demean ignoring the entity-fixed effects. Assume the panel is balanced.

   Show that the order of demean is unimportant (either starting with entity-demean or time-demean doesn't matter). Give an intuitive explanation for why. (3 points)

Ans:
Define $\bar{Y}_t = \frac{1}{N}\sum_i^N Y_{it},\ \bar{Y}_i = \frac{1}{T}\sum_1^T Y_{it},\ \bar{X}_t = \frac{1}{N}\sum_i^N X_{it},\ \bar{X}_i = \frac{1}{T}\sum_1^T X_{it},$
$\bar{\alpha}_i = \frac{1}{N}\sum_1^N \alpha_i,\ \bar{\gamma}_t = \frac{1}{T}\sum_1^T \gamma_t.$

(a) Entity-demean first.

$$Y_{it} - \bar{Y}_i = \beta_1(X_{it} - \bar{X}_i) + \alpha_i - \alpha_i + \gamma_t - \bar{\gamma}_t + u_{it} - \bar{u}_i$$

Then, time-demean

$$\begin{aligned} Y_{it} - \bar{Y}_i - \frac{\sum_1^N (Y_{it} - \bar{Y}_i)}{N} &= \beta_1[X_{it} - \bar{X}_i - \frac{\sum_1^N (X_{it} - \bar{X}_i)}{N}] + \\ \gamma_t - \bar{\gamma}_t - \frac{\sum_1^N (\gamma_t - \bar{\gamma}_t)}{N} &+ u_{it} - \bar{u}_i - \frac{\sum_1^N (u_{it} - \bar{u}_i)}{N} \\ &\Rightarrow \tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it} \end{aligned}$$

(b) Time-demean first.

$$Y_{it} - \bar{Y}_t = \beta_1(X_{it} - \bar{X}_t) + \alpha_i - \bar{\alpha}_i + \gamma_t - \gamma_t + u_{it} - \bar{u}_t$$

Then, entity-demean

$$\begin{aligned} Y_{it} - \bar{Y}_t - \frac{\sum_1^T (Y_{it} - \bar{Y}_t)}{T} &= \beta_1[X_{it} - \bar{X}_t - \frac{\sum_1^T (X_{it} - \bar{X}_t)}{T}] + \\ \alpha_i - \bar{\alpha}_i - \frac{\sum_1^T (\alpha_i - \bar{\alpha}_i)}{T} &+ u_{it} - \bar{u}_t - \frac{\sum_1^T (u_{it} - \bar{u}_t)}{T} \\ &\Rightarrow \tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it} \end{aligned}$$

The key idea is that both entity- and time-demeaning remove the fixed effects (i.e., the α_i and γ_t) by centering the data around the group means (either across entities or across time). Demeaning in either order subtracts the fixed effects in different sequences, but the ultimate result is the same: the fixed effects are eliminated, and the remaining variation is purely due to the explanatory variable X_{it} which is used to estimate β_1.

Question 2

Following a national poverty alleviation program in Gansu province, in July 2017, many households in Gansu province received subsidy and financial support. Thus, this policy experiment induced a geographical allocation of subsidy that can be presumed exogenous in an income growth. Let Y_i^0 and Y_i^1 denote the consumption in village i before and after the policy intervention. Let D_i be a dummy variable that takes the value 1 if household i is in Gansu province, 0 otherwise (other provinces). We would like to estimate the treatment effect of the income change on consumption with the difference-in-difference estimator.

1. (2 points) Write down this difference-in-difference estimator as a function of $\{Y_i^0, Y_i^1, D_i\}$ for $i = 1, ..., n$.

   Ans:

   - The differences-in-difference estimator for the causal effect of the poverty alleviation program corresponds to β_1 in the regression,

     $$\Delta Y_i = \beta_0 + \beta_1 D_i + u_i$$

     where $\Delta Y_i = Y_i^1 - Y_i^0$.
   - Alternatively, we can write the treatment effect by $(\bar{Y}^1_{D_i=1} - \bar{Y}^0_{D_i=1}) - (\bar{Y}^1_{D_i=0} - \bar{Y}^0_{D_i=0})$, where $\bar{Y}^1_{D_i=1}$ is the sample average of consumption among households in Gansu province after the policy, and $\bar{Y}^1_{D_i=0}$ is the sample average of consumption among households in the other provinces after the policy, and the other two terms are defined similarly.

2. (2 points) What is the key assumption for this estimator to be an unbiased estimator of the treatment effect?

   Ans:
   The key assumption is the parallel trend assumption: without the program, the consumption change $\Delta Y_i = Y_i^1 - Y_i^0$ on average would have been the same for the treatment group and the control group.

3. (2 points) Suppose after the policy intervention, households in Gansu province works harder (policy irrelevant) and thus higher income and consumption are expected. Does the difference-in-difference estimator under or over estimate the treatment effect. Explain.

   Ans:
   Overestimation.
   From the perspective of the omitted variable bias formula, we know that working hard is both positively correlated with D_i (because we state that households in Gansu province work harder) and consumption (more wage income allows more consumption).
   Alternatively, we can also explain this through a violation of the parallel trend assumption. In this case, the consumption in the treatment group would have gone up even without the treatment. Attributing this change to the effect of the treatment will result in an overestimation of the effect of the poverty alleviation program.

Question 3

We are going to replicate a study conducted by Card and Krueger in 1994 that investigate the relationship between a rise in minimum wage and employment.

Economic theories have long suggested that increases in the minimum wage lead to a reduction in the employment for at least two reasons: Businesses are less likely to hire and will rather invest in other resources that are now cheaper because of wage increase. Higher salaries will induce businesses to raise their prices to compensate their greater costs; as prices increase, we expect fewer buyers, which will lead to lower demand and employment.

These theories have found mixed support but the discussion is still very much open within the policy world, as states discuss the opportunity to rise their minimum wage to help local populations to face increasing living costs. Discussions are currently occuring in New Jersey and Illinois to raise the minimum wage to 15$/hour (New york has successfully passed this same raise in 2018).

One of the first study looking at this policy problem was Card and Krueger's. They applied a difference-in-difference the design to look at two groups of fast-food restaurants: fast-food restaurants in New Jersey where the minimum wage increased from 4.25$ to 5.05$ per hour (treatment group) AND fast-food restaurants in Pennsylvania where the minimum wage did not increase (control group). They collected employment data before and after the minimum wage was approved.

Table 1: Variable Description

Variable name	Description
ID	Unique identifier for fast food
Treatment	Pre-treatment (=0) and post-treatment (=1)
Group	1 if NJ (treatment); 0 if PA (Control)
Empl	# of full time employees
C.Owned	If owned by a company (=1) or not (=0)
Hours.Opening	Number hours open per day
Soda	Price of medium soda, including tax
Fries	price of small fries, including tax
Chain	1 = BK, 2 = KFC, 3 = Roys, 4 = Wendys
SouthJ	South New Jersey
CentralJ	Central New Jersey
NorthJ	North New Jersey
PA1	Northeast suburbs of Philadelphia
PA2	Easton and other PA areas
Shore	New Jersey Shore

1. Explain why a simple comparison in employment between New Jersey and PA after the minimum wage policy may not be a good estimation for the treatment effect. (2 points)

   Ans:
   It is confounded by the unobserved differences between NJ vs. PA. (Give a specific example or draw a graph to illustrate to get full credit)

2. What is the regression model you'd like to estimate? (2 points)

Ans:
The regression model is specified as follows:

$$Empl = \beta_0 + \beta_1 Treatment + \beta_2 Group + \beta_3 Treatment * Group + \alpha_1 X + u$$

where X stands for the series of control variables.

3. What is the key assumption of the estimation method you use? (1 point)

Ans:
Key assumption: parallel trend assumption, i.e. without the treatment, the employment between the treated group and the control group was to move in the parallel trend over time.

4. The data is in the blackboard (bb) Assignment 4 file folder titled *DID_Example.csv*, and the variables are defined as in Table 1. Run the regression model you proposed in 2, and report the regression results. (3 points)
(Hint: You need first import the csv file into Stata.
Replace all "NA" with the missing value in variables, using

replace var_name = "" if var_name == "NA"

Then, convert string variables to numeric variables, by the "*destring*" command.
Please include control variables that you think necessary.)

Ans:
Stata command:
reg Empl c.Treatment c.Group c.Treatment # c.Group X
(X denotes the series of control variables)
or
gen Treatment_Group=Treatment * Group
reg Empl Treatment Group Treatment_Group X

5. Explain why this is an unbiased estimation of the treatment effect. (Hint: You can use a graph if necessary) (2 points)

Ans:
Draw the graph as in the slides, and show that the parallel trend assumption has been held.