**Problem Set 1 & Answer Sheet**

Q1. Describe how one can design a hypothetical randomized control experiment to study the effect of having a female mathematical teacher on the math test scores of middle-school students.

1. Identify all students and teachers that satisfy the selection criteria to participate to the program.
   1. The school principals promise that students are randomly assigned at the beginning of the study
   2. After joining in classes in grade 7, school would not rearrange their classes in grade 8 and 9
   3. Teachers and parents report that the class is not assigned from students’ scores and performance
2. Using lottery method to divide the student in the school randomly into two parts. One is the classes with a female math teacher and those with a male math teacher.
3. Conduct balance test between class with a female math teacher and those with a male math teachers based on students predetermined characteristics. We do this to make sure the students are randomly assigned to different class, if so, all characteristics should be similar.
4. To estimate the effect of having a female mathematical teacher on the math test scores, we can run a regression model.

: math test score of student i in class c and school s

: dummy variable of whether student is female

: dummy variable of whether math teacher is female

: predetermined characteristics

: error term

: capture the mean math scores outcome between female and male students when having a male math teacher

: the effect of having a female math teacher

: the effect of teachers gender on students gender differences

1. Robustness checks and further identification. Since this is not related with hypothetical randomized control experiment I would not elaborate on.

Q2. Use the academic year 2013-2014 baseline survey data from the China Education Panel Survey (CEPS) to study the effect of having a female mathematical teacher on the math test scores of middle-school students.

Steps:

1. Download the data and questionnaires from Blackboard or https://ceps.ruc.edu.cn/index.php?r=index/index.

2. Merge the samples of students, parents, classes, schools. (Stata command: merge)

图片包含 文本

描述已自动生成

3. Keep those schools where the students were randomly assigned into different classes when they enrolled in Grade 7.

图片包含 图形用户界面

描述已自动生成

4. Define the treatment status: treatment=1 if the math teacher is a female.

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

5. Check the balance of some predetermined characteristics of students between the treatment group and the control group.

*Examples of the predetermined characteristics (determined before Grade 7): Gender, ethnicity group (being Han), being local residents, being the only child, father’s years of education, mother’s years of education, had preschool education, math scores at primary school, etc.*

图形用户界面, 表格

描述已自动生成

For balancing test, we can observe that all of the predetermined characteristics is statistically insignificant. F-statistics below also shows that they are jointly significant.

6. Compare the outcomes related to math study between the treatment and control groups, interpret your empirical results.

表格

描述已自动生成We know that captures the effect of having a female math teacher. According to our sample results, we can see that having a female math teacher would decrease the students’ math scores by 147%, the outcome is statistically significant. This imply that having a female math teacher can decrease students’ math test scores.

( Sorry Professor, I think I might have multicollinearity issue here, but I’ve checked for my regression and I couldn’t figure out why )

7. Check the heterogeneous effects across the gender of students, interpret your empirical findings.

表格

描述已自动生成

For heterogenous effects, we would like to know if the effects vary according to students’ gender. From the table above, we can observe the interaction of female student and female math teacher is insignificant, which suggest a female math teachers’ influence would not very likely vary among students with different gender.

Notes:

Problem set 1 will be due on Oct. 20, 2021. You should turn in three separate documents electronically: one that contains your typed answers to the problem set questions, one well-organized and well-commented Stata do file, and one Stata log file.