

## ECON 523: Program Evaluation for International Development

### Empirical Exercise 3

In this exercise, we're going to analyze data from Ignaz Semmelweis' handwashing intervention in the maternity hospital in Vienna. The data come from Semmelweis' (1861) book, and some helpful person put them on Wikipedia.

Before you begin, save a new do file containing the Stata code below.

```
clear all
set scheme s1mono
set more off
set seed 314159

** change working directory as appropriate to where you want to save
cd "C:\Users\pj\Dropbox\ECON-523\topics\3-DD1\stata"

** load data
use E3-semmelweis-vienna-by-wing.dta
drop if Year<1840
```

Extend your do file as you answer the following questions:

1. Use the following reshape command to convert your data into a panel data set containing a variable **Rate** and a variable **clinic** that indicates whether an observation comes from Clinic 1 (doctors) or Clinic 2 (midwives). How many observations are there in the data set now? How many from each clinic?

```
reshape long Rate, j(clinic) i(Year)
```

2. Generate a **post** variable equal to one for years after the handwashing policy was implemented (and zero otherwise) and a **treatment** variable equal to one for the doctors' wing (and zero otherwise).
3. Generate the interaction term you need to estimate a difference-in-differences model in a regression framework.
4. Use the **labelvariable** command to give your variables short, easy to interpret labels.
5. Implement difference-in-differences in an OLS regression framework. Use the command **eststoclear** immediately before your **regression** command, and then use the command **eststo** (estimates store) immediately after. This will save your results.

6. You can use the `esttab` command to make a table of your regression results. Try it by typing `esttab` in the command window. The command `esttabusingclinic-regs.rtf` will save your table as a word document. Look through the `esttab` options to make your table look more professional: report standard errors rather than t-statistics in parentheses below your coefficients. Have your variable labels appear in place of variable names, and make sure your first column is wide enough to accommodate the labels you have given your variables. Make the column with your regression coefficients say OLS at the top using `esttab`'s `mtitle` option.
7. Take a screenshot of your finished regression table and upload it to gradescope.
8. Which coefficient in the regression table (i.e. the coefficient on which variable) is the difference-in-differences estimate of the treatment effect of handwashing on maternal mortality?
9. Which regression coefficient is the estimate of the degree of selection bias?
10. Which regression coefficient is the estimate of the time trend in the absence of treatment?