PS3

Kohei Takata

4/1/2022

Introduction

In this section, we will estimate the effect of unilateral divorce laws on female suicide. In the problem set link, we have provided a lightly cleaned version of their main analysis files: stevenson_wolfers_210.dta. Keep this file in a subfolder called data of your problem set Rstudio Project; this will facilitate our submission verification described at the end of this part. The data we are using is available from Justin Wolfers' website.

Variables The data for this problem set is a state-by-year panel. Observations are uniquely identified by state, year, and sex. The data has the following key variables: • st and year are the state and year variables. • sex indicates whether the outcome is observed for males or females. It is coded as 1 for males and 2 for females. • divyear is the year of unilateral divorce reform. • unilateral indicates whether unilateral divorce is legal. • suiciderate_jag is the suicide rate.

```
#data_ori <- read_dta("./data/stevenson_wolfers_210.dta")
data_ori <- read_dta("stevenson_wolfers_210.dta")
```

12. We will begin by estimating a simple 2x2 difference-in-differences regression.

a. The year in which the greatest number of states passed unilateral divorce laws was 1973. Using data on states that passed unilateral divorce laws in 1973 and those that never passed unilateral divorce laws, run a simple 2x2 DD regression to estimate the effect of unilateral divorce laws on ln(suiciderate_jag) for women, clustering standard errors at the state level. Report the estimated effect and the standard error below.

Question 12 (a) – Suggested Steps: 1. Basic data cleaning. a. Select only relevant observations. (Which gender? For what variable do we need to remove the NAs?) b. Remember we'll be looking at the log of the suicide rates! Create a new variable. c. You may want to change the class of st and year . 2. Create variables specifically for 2x2 DiD a. Create a new dataset for this regression. You will need to create 3 new variables: "treat", "post" and "treat*post" (the product of the first two). b. Treat: it should be 1 for the observations that are treated (year of unilateral divorce reform passed in 1973), and 0 for not treated (never had unilateral divorce reform in our timeline, i.e., reform passed in 2000). c. Post: year of treatment is 1973, so the post-period should be 1 for the years after 1973. d. Only keep the treated and control observations, as defined in the question. In other words, we only want to keep observations for which the reform was passed in 1973 and in 2000. 3. Run 2x2 DiD a. You can use felm. b. Remember to cluster the st. errors at the state level.

```
## data clearing
data <- data_ori %>%
        mutate(ln_suicide_rate = log(suiciderate_jag)) %>%   # take a natural log of suicide
        filter(sex == 2)                                     # filter by women
```

```r
## create a new dataset for regression
data_12a <- data %>%
            mutate(treat = ifelse(divyear == 1973, 1, 0),    # 1 if uni-divorce passed in 1973, otherwis
                   post  = ifelse(year    >  1973, 1, 0),    # 1 if the data is after 1973, otherwise 0
                   DiD   = treat*post) %>%
            filter(divyear == 1973 | divyear > 1996)         # filter states !(uni-divorce passed in 199

## regression
reg_12a <- felm(ln_suicide_rate ~ treat + post + DiD|0|0|st, data = data_12a)
#summary(reg_12a)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Fri, Apr 01, 2022 - 2:18:36 PM

Table 1: changes in Divorce rate by state

	Dependent variable:
	ln_suicide_rate
treat	0.344^{***}
	(0.120)
post	0.014
	(0.074)
DiD	-0.093
	(0.092)
Constant	-10.007^{***}
	(0.088)
Observations	528
R^2	0.158
Adjusted R^2	0.153
Residual Std. Error	0.308 (df = 524)
Note:	*p<0.1; **p<0.05; ***p<0.01
	Treatment variable is 1 if the state introduced the unilateral divorce act in 1973, and it is 0 if the state had never introduced the act by 1996. Post variable is 1 if the data was observed after 1973, otherwise, it is 0. the parenthesis below coefficients are clustered standard error by states.

13. Now we will assess whether the parallel trends assumption is reasonable in this setting by estimating an event study, pooling data from all the states

a. Consider the following event study regression specification:

$$Y_{st} = \sum_{j \neq -1} \beta_j 1 * (t - divyear_s = j) + \gamma_s + \delta_t + \epsilon_{st}$$

Interpret the coefficients.