ESCOLA BRASILEIRA DE ECONOMIA E FINANÇAS - EPGE

Luciano Fabio Busatto Venturim

Econometrics 1 - Problem Set 5

Rio de Janeiro

4th Quarter - 2021

Question 1: Read Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania by D. Card and A. Krueger. Then follow the R tutorial by Philipp Leppert (available here) and replicate the main findings of the paper.

Question 2: This and the next question build on Patrick Kline's work on *The impact of juvenile curfew laws on arrests of youth and adults*, published at the American Law and Economics Review in 2012. Download the data file "curfews class". This file contains data on the number of arrests of youth in a panel of cities enacting juvenile curfew laws in different years.

1. Construct a variable that equals one when a city enacts a curfew law, i.e., D_{it} = (year = enacted).

2. Run an event study of log juvenile arrests on curfew enactment with city and year effects. Be sure to "bin up" the endpoints and to cluster the standard errors by city. As a normalization, leave out D_{it+1} so that all effects can be measured relative to the period before enactment.

3. Make a plot of the event study coefficients with confidence intervals. What do you conclude about the impact of the program on arrests?

4. In light of the recent developments in the difference in difference literature, mention and explain the potential threats to identification in this setting.

Solution:

3. If the enactment was in fact "random", we would expect the coefficients of the dummy variables Q_6 to Q_2 to be around zero, which means that there were little difference between the control group (formed by the ones who were never treated and the ones who had not been treated yet) and the treated group prior to the treatment, This is what we see in the image below, with the point estimates very close to zero and zero being included in all the confidence intervals. In the periods subsequent to the treatment, almost all coefficients are statistically different from zero, since for most of them, their confidence interval (at the level of 95%) do not include zero. Since the coefficients represent the difference between the treated group

and the control group, we conclude that the enactment of curfew laws reduce the number of arrests.

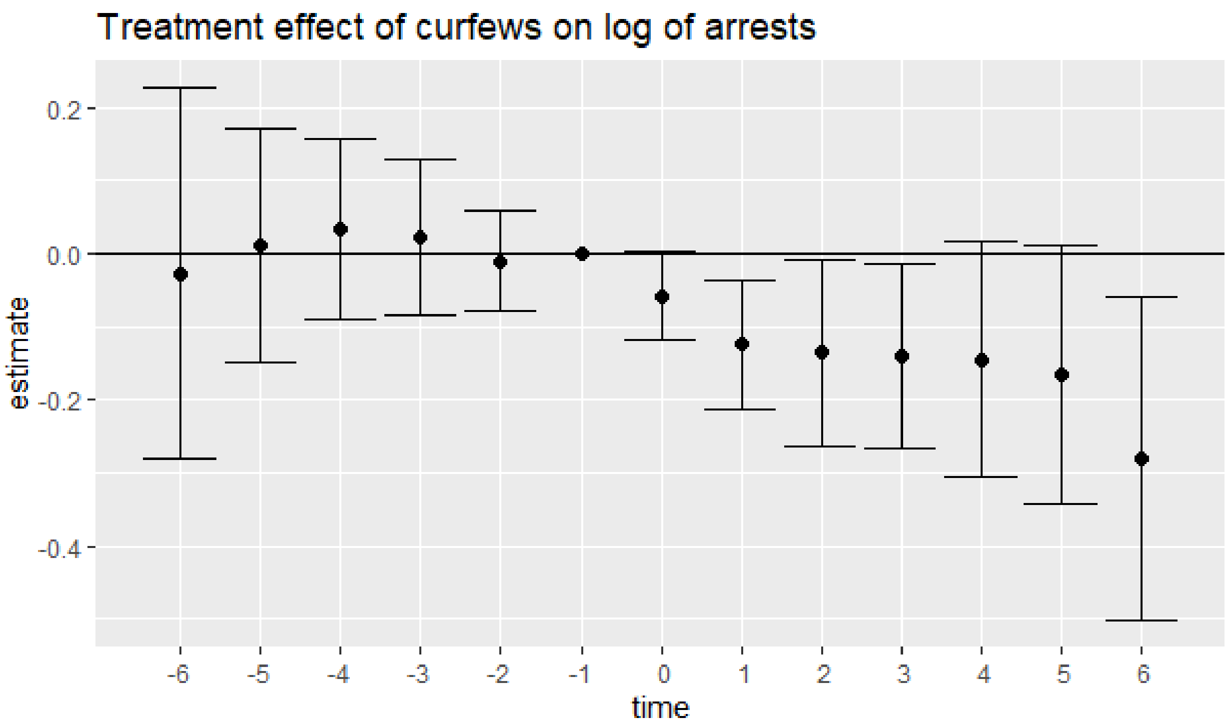

Figura 1: Treatment Effect

4. When different groups are treated at different times and the treatment effect might vary over time, we might have a problem estimating the true treatment effect using a two-way error components model. The first problem may create bias in the estimator if strict exogeneity fails, which can happen specially if there is anticipation of the treatment,or when the parallel trend assumption does not hold (the error v_t is not the same for the treated group and the control group). When the the treatment effect vary over time, what we get from the the previous model as an estimate of the treatment effect is in fact an weighted average of all 2x2 DiD estimators possible in the model (2 periods and 2 groups), and this weighted average is of difficult interpretation, so we cannot know what is in fact being estimated by the model.

Question 3: Having looked at the event study you decide to try parameterizing the treatment effect to obey a simple partial adjustment model (we studied this type of models last week) of the form:

$$Y_{it} = \alpha_i + \gamma_t + \delta Y_{it-1} + \beta_0 D_{it} + \beta_1 D_{it-1} + \epsilon_{it}$$

1. Take first differences of the above model in order to eliminate the fixed effect and run ΔY_{it} on ΔY_{it-1}, ΔD_{it}, ΔD_{it-1}, and year dummies. Interpret your results.

2. Arrelano and Bond (1991) suggests instrumenting for ΔY_{it-1}. Try using Y_{t-2} as an instrument. Now try using ΔY_{it-2} as an instrument. Now use them both. How do the results change)

Solution:

1. Here, I interpreted D_{it} as the variable D_0 created in question 2, i.e, $D_{it} = D^0_{it} = I(t = e_i)$, where e_i is the date of treatment, and $D_{it-1} = D^1_{it} = I(t = e_i + 1)$, i.e, D_{it-1} indicates that the event occurred at the previous period. We then take the first difference and run the regression asked. The image 2 shows the point estimates and cluster-robust standard errors

```
                     Estimate  Std. Error t value  Pr(>|t|)
(Intercept)       -0.02444873  0.02594099 -0.9425 0.3461456
lnarrests_diff_lag -0.14043721  0.12850996 -1.0928 0.2747029
Q_0_diff          -0.00971403  0.02446676 -0.3970 0.6914183
Q1_diff           -0.03196168  0.02726464 -1.1723 0.2413265
```

Figura 2: Two-way error component fixed effects

Note that the estimated effect is negative, i.e., the enactment of the curfew laws have a decreasing effect on the log of arrests, and that this effects is gets bigger over time. However, none of the estimates are statistically significant.

2. First, we use only the 2-lagged value of the variable lnarrests (Y_{it-2})) as an instrument for ΔY_{it-1}. In this case we have the following estimates and (cluster) standard errors

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0237675	0.0484824	0.490	0.6241
lnarrests_diff_lag	0.7482765	0.3859356	1.939	0.0528
Q_0_diff	-0.0503111	0.0376044	-1.338	0.1812
Q1_diff	-0.0534364	0.0314080	-1.701	0.0891

Figura 3: Lag IV

As we can see in image 3, now the estimated coefficient for the lagged dependent variable is positive and very close to one. The coefficients one the treatment dummy variables remain negative but they have a bigger absolute value, meaning that the effect of treatment is bigger.

Then, we estimate the model using the 2-lagged value of the first difference of the dependent variable (ΔY_{it-2}).

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-0.047535	0.023979	-1.982	0.0477	*
lnarrests_diff_lag	0.368953	0.280949	1.313	0.1894	
Q_0_diff	-0.033644	0.031656	-1.063	0.2881	
Q1_diff	-0.043610	0.028841	-1.512	0.1308	

Figura 4: Diff IV

As we see in the image 4, the point estimates change (the treatment effect decreases), but now none of the estimates are significant at the usual levels.

Finally, we use both instruments. In the image 5, we can see that the point estimates almost do not change relative to the previous model, but the estimate of the coefficient of the lagged depend variable becomes significant at the 90% level of significance. This might indicate that the variable ΔY_{it-2} is a weak instrument that generates inflated standard errors.

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-0.046275	0.024949	-1.855	0.0639	.
lnarrests_diff_lag	0.443807	0.265760	1.670	0.0952	.
Q_0_diff	-0.037095	0.032017	-1.159	0.2469	
Q1_diff	-0.045387	0.028914	-1.570	0.1168	

Figura 5: Both IV