Assignment 5 - Econometrics II

Floris Holstege, Stanislav Avdeev

1

This approach is a difference-in-difference approach, which can be written in OLS form as:

$$
\begin{aligned}
y_{g1} - y_{g0} &= (a_1 - a_0) + \delta D_g + (U_{g1} - U_{g0}) \\
&= B_0 + \delta D_g + U_g
\end{aligned}
$$

Where y_{g1}, y_{g0} is the dependent variable (grade point average) at after the first year ($t = 0$), and after the second year ($t = 1$). B_0 is the trend and δ is the effect of treatment D_g, in this case providing housing if one has a GPA higher than 8. U_g are the errrors of this model.

We know that OLS is consistent as long as D_g and U_g are uncorrelated. In order for this to hold, the treatment needs to be independent of the outcome variable. That is not the case here, as the outcome variable consists partially of first year grades (y_{g0}), which also determine whether or not someone receives the treatment. Hence, we would not expect this approach to give a consistent estimator.

2

I)

A bivariate model of sex ratio and % out-of-wedlock births, using only observations from the year before the war, finds no statistically significant effect from sex ratio on % out-of-wedlock births (robust standard errors), suggesting that position of men in the marriage market does not improve with a reduction in de sex ratio.

Table 1:

	Dependent variable:
	% of out-of-wedlock births
Sex Ratio	−0.089
	(5.050)
Constant	6.772
	(5.786)
Observations	87
Adjusted R^2	−0.012
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

	Before the war	After the war	Difference
Above the median	5.09	6.15	1.06
Below the median	7.96	8.45	0.49
Difference-in-Difference			0.57

However, there are some obvious problems with this approach. The problem with this approach is that there might be department-specific attributes that could mean the the % of births out of wedlock are structually higher or lower. Using OLS, while ignoring these department-specific effects results in biased estimations.

We can improve upon this approach by using a difference-in-difference approach, with the military mortality rate as a treatment. In the French army, the vast majority of armed forces were men. A high military mortality rate should thus decrease the ratio of men to women, increasing the position of men in the wedding market after the war, subsequently leading (if you believe the argument made in the exercise) to more out-of-wedlock births.

If we use difference-in-difference, we avoid the previously mentioned problem, if one believes that the department-specific attributes are consistent over the two periods.

II)

III)

We run the following model

$$\%of out-of-wedlockbirths = \beta_0 + \beta_1(Post \cdot \text{Military mortality}) + \beta_2 Post + U$$

In which the dependent variable is the % of out-of-wedlock births, "Post" is a dummy variable, with value 1 if this dependent variable was measured after the war, and 0 if before.

Before we make any conclusions, how should we interpret the coefficients β_1, β_2?. For β_1, the interaction term is simply the effect of military mortality on % of out-of-wedlock births, if this % was measured after the war. This is appropriate, because if the % was measured before the war, we would not yet know the potential effect of the military mortality rate. For β_2, this simply measures the (average) difference between % of out-of-wedlock births before and after the war. This is important, because it ensures that any effect attributed to the 'Military mortality' variable is not due to a general increase in out-of-wedlock births after the war.

Table 2 shows the results for the abovementioned model, with robust standard erros, and with and without department dummies. When we do not apply department dummies, the model shows that the 'post' variable is statistically significant, indicating that in general the % of out-of-wedlock births increased by 9.4% after the war. But the interaction variable between 'post' and military mortality is also statistically significant, indicating that one % increase in the mortality rate on average led to an 0.5% decrease in the out-of-wedlock births, suggesting that the position of males in the wedding market only worsened as the sex ratio decreased.

IV)

Once we include dummies for each department, the results are reversed: the 'post' variable is still statistically significant, but now indicates that the % of out-of-wedlock births decreased by 1.7% after the war, and the interaction variable between 'post' and military mortality indicates that one % increase in the mortality rate on average led to an 0.15% increase in the out-of-wedlock birth, suggesting that the position of males in the wedding market only worsened as the sex ratio increased.

Table 2:

	Dependent variable:	
	illeg	
	(No department dummies)	(With department dummies)
post_mortality	-0.515^{***}	0.148^{***}
	(0.151)	(0.049)
post	9.376^{***}	-1.740^{**}
	(2.689)	(0.872)
Constant	6.672^{***}	6.360^{***}
	(0.380)	(0.071)
Observations	174	174
Adjusted R^2	0.085	0.958
Note:		*p<0.1; **p<0.05; ***p<0.01

How is this possible? When we use a dummy for each department, we assume that there is a fixed effect (both before and after the war) per department that affects the the % of out-of-wedlock births. We prefer this specification, because it seems plausible that certain aspects of department, such as socio-economic circumstances and social and cultural norms in particular areas, should affect the rate of out-of-wedlock births.

V)

The key assumption we make is the so-called ‘common trend’ assumption, which states that the difference between the control and treatment (existent because of the different make-up of the control and treatment group) remains constant over time. In our case, we assume that the trend in the difference in out of wedlock births between departments that ended up having with high and low military mortality rates continued.

How can we test if this assumption is reasonable? One of the easiest and most intuitive ways is to simply plot the difference in out of wedlock births between departments over time, and see if the difference is constant or rather time-varying. But we cannot use this test in this case, since we only have data for two moments (before and after the war), instead of several years leading up to the war.