Zachary Chance
5-28-2020

GitHub: https://github.com/Zachary-Chance1/Assignment-5-TWFE-DID.git

Assignment 5

1. Done. The table can be found under the Data subdirectory and is called hw5_ex1:

Unit	Time	Y	D	Mean Y	Mean D	Demeaned Y	Demeaned D
1	1	60	1	46.55	0.5	13.45	0.5
1	2	35	0			-11.55	-0.5
2	1	10	1			-36.55	0.5
2	2	156	0			109.45	-0.5
3	1	13	0			-33.55	-0.5
3	2	65.5	1			18.95	0.5
4	1	10	0			-36.55	-0.5
4	2	20	1			-26.55	0.5
5	1	35	0			-11.55	-0.5
5	2	61	1			14.45	0.5

2. The Coefficient δ is -15.50 with a standard error of 32.85
3. The Coefficient δ is -15.50 with a standard error of 27.31. It is the same Coefficient as 2 with a smaller Standard error. This tells me that running a regression with demeaned variables and a fixed effect for time is numerically equivalent to regressing the non-demeaned variables and controlling for unit and time fixed effects. (see Tables/TWFE_Table)
4. I summarize the value of TE in DID_dist_te in Tables. I find the mean to be 5.4. Thus, under constant treatment effects δ should be 5.4. (Also see Figures/Dist_of_te_histogram)

5. Using fem, I find the Coefficient on TREAT to be 6.3. This leads me to believe that this is biased, because δ should be 5.4 instead of 6.3, so we are at least a little biased. (See DID_Prob_5.rtf)

6. Group 4 is treated in 2004.
We use
TE × [Year − Treatdate + 1]
to get
4 × [2006 − 2004 + 1] = 12
AND
4 × [2007 − 2004 + 1] = 16
The TE should be positive in all Post-treatment Periods.

7. I get a coefficient on TREAT of −21.8153. This leads me to believe that it is biased by heterogeneity to a large degree, as it isn't even the same sign as we would expect it to be.
With dynamic treatment effects, OLS is:
$$\hat{\delta}^{DD} = \sum_{k=U} s_{kU} \hat{\delta}^{2x2}_{kU} + \sum_{l \neq U} \sum_{l>k} s_{kl} \left[\mu_{kl} \hat{\delta}^{2x2,k}_{kl} + (1-\mu_{kl}) \hat{\delta}^{2x2,l}_{kl} \right]$$
But also $\text{plim}_{n\to\infty} \hat{\delta}^{DD} = \delta^{DD} = VWATT + VWCT - \Delta ATT$
Use $ATT(2004,2007) \rightarrow -21.8153 = 16 + 0 - \text{Heterogeneity bias}$
For year = 2007, the heterogeneity bias is 15.8153 with dynamic TE (see Tables/DID_Prob_7)

8. The CS Estimator is:
$$ATT(g,t) = E\left[\left(\frac{G_g}{E[G_g]} - \frac{\frac{p(X)C}{1-p(X)}}{E\left[\frac{p(X)C}{1-p(X)}\right]}\right)(Y_t - Y_{g-1})\right]$$

9. Note: I had extreme problems getting this to work with R, so eventually I ran it inline in Stata. However, I couldn't figure out how to hook up Stata and GitHub, so I have no files to report.
Group4's ATT is 12 in 2006 and 16 in 2007. These are unbiased because they are exactly the value stated in Problem 6. That is, the expected value of the estimator is equal to the actual value of ATT for Group4 in 2006 and 2007.