EC 587 HW3

Erik Andersen

October 2023

Question 1

a)

Below is the table from the pooled regression.

Variable	Coefficient	SE
Treatment	1.909594	0.5449885

This cannot be interpreted as a causal effect because it is just a naive regression. So the interpretation is that being in the treatment group is associated with an increase of about 2 candidates on the ballot. It is significant.

b)

Below are the estimated means for Seattle before and after the intervention

Variable	Coefficient
Before	3.076923
After	8.10989

We see that the cross-time comparison is associated with a much larger increase in number of candidates than the naive regression. Below are the same estimates for the comparison cities.

Variable	Coefficient
Before	3.4776955
After	3.892523

The difference between Seattle's difference and the comparison cities difference is 2.66. So Seattle increased an extra 2.6 candidates per election more than the comparison cities. This is the diff-in-diff estimate, so if all the standard assumptions hold, this is the causal effect of the policy.

c)

Below is the comparison between Seattle and the control cities before the intervention in 2017.

Variable	Coefficient
Control	3.483982
Seattle	4.795354

In 2017, Seattle had about one extra candidate per election. Below are the means after the intervention.

Variable	Coefficient
Control	3.888271
Seattle	8.040663

After the intervention, Seattle had about 4 more candidates than the control cities. The difference between these two is 2.84, which is qualitatively the same as the estimate from part c. The point estimate is slightly different because of the controls present. This again is the diff-in-diff estimator, so we can interpret this as the causal effect.

d)

The estimate from the simple diff-in-diff is 2.84. This is identical to part c, and qualitatively the same as part b. They are all estimating the same effect, so this is expected.

e)

The effect estimated from the two-way fixed effects framework is 3.23, which matches the value in the paper.

f)

For brevity's sake, I don't report the regression table here, but see the attached pdf for interest. The coefficient of interest is the interaction between the time variable, and a dummy variable for being in Seattle. If there are parallel trends, this coefficient shouldn't be statistically different from 0, which is saying that the control cities and Seattle had the same trend pre-intervention. The coefficient without city fixed effects is 0.02 with associated p-value 0.83. This gives us the expected outcome of pre-treatment parallel trends. Similarly with city fixed effects, the estimate and associated p-value are 0.0006 and 3.62e-03. This is significant, but the difference in slope is not qualitatively meaningful. Either way, we have evidence of parallel pre-trends.

g)

The coefficient of interest here is the interaction between a dummy for being post intervention and Seattle. The results are reported below.

Variable	Coefficient	se
PostxSeattle	3.449166	0.942599

The results match exactly the estimates in the paper.

h)

Below are the results for the same regression as part g, but estimated only with comparison cities in Washington.

Variable	Coefficient	se
PostxSeattle	3.377484	0.803842

The effect is qualitatively the same as when using all the cities. Below is the same, but with only California cities.

Variable	Coefficient	se
PostxSeattle	3.471152	0.942289

Again, we have basically the same result.

Question 2

b)

In this section, we run a Hausman test on the results from parts d and e. In the Hausman test, we assume that the simple diff-in-diffs regression is consistant and efficient. We then compare this to the two-way fixed effect model from part e. The Null of the test is that the estimate from the two-way model is consistent and efficient. The alternative hypothesis is that there are systematic differences between the two test, and because we assume the simple diff-in-diffs is consistent and efficient, this implies the two-way model is not-efficient. The test returns a chisquared test statistic of 5.5 with implied p-value of 0.3504. We fail to reject the null here, so we cannot reject that two-way fixed effects are efficient here.

c)

In this test, we are seeing if there is a significant difference between the treatment effect in the simple versus two-way fixed effects models. If there is a difference, this implies that one of the models is misspecified. The null of the test is that the two treatment effects are the same. The alternative is that they are different which implies the two-way model is misspecified since we assume the simple diff-in-diff is correct. The test statistic is 0.0917 which has associated p-value of 0.7621. We again fail to reject the null of the estimates being different.

Question 3

a)

The estimate on the balanced panel is given below.

Variable	Coefficient	se
Treatment	3.1471380	0.5398269

We get a larger effect than in part d above. The result is closer to the two-way fixed effect from part e. The standard error is smaller, but I think this is due to a different regression method calculating standard errors in a slightly different way.

b)

Below is the estimate.

Variable	Coefficient	se
Treatment	3.58461	0.390617

Again, the estimate is higher than part e, and the standard error is smaller.

c)

See the pdf for a pretty graph of the weights. San Diego, Tacoma, Spokane, and Oakland in descending order get the most weights in the full synthetic control. This is a bit weird. I was expecting San Francisco, but it's hardly given any weight. San Diego seems very strange as the top city to me.

d)

When using only Washington unsurprisingly, Tacoma and Spokane get the most weights. This makes sense intuitively, and it makes sense the same cities were selected when using all the cities and when using only Washington cities.

e)

San Diego, Long Beach and Fresno are the top weighted cities using only California. This is surprising for a bunch of reasons. Once again, San Francisco doesn't show up, and in fact no bay area cities are given any significant weight. Its also surprising that Oakland doesn't show up with a high weight here despite being the fourth highest weighted city when using the whole sample.

f)

The treatment effects for each of the cities acting as a placebo are shown in the table below. Seattle was omitted from the sample

City	Treatment
Bellevue	-0.80490472
Everett	0.74340100
Fresno	0.39988461
Kent	0.10871953
Long Beach	0.03236447
Los Angeles	0.12500044
Oakland	1.68856365
Sacramento	-0.81663585
San Diego	1.42376659
San Francisco	0.35308083
San Jose	-0.77738658
Spokane	-0.01924154
Tacoma	0.35998186
Vancouver	0.40646735

We see that none of the placebo estimates are close in magnitude to Seattle's. This gives more credence to Seattle's effect being a true effect because none of the other cities experienced a similar jump to Seattle in its coefficient.