Econ 262A: Problem Set 2

Kenny Guo

Problem 1: Difference-in-Differences Simulation

Simulate a data set with 10,000 individuals i over 10 time periods $t = 1, \ldots, 10$ in your choice of program. Construct the following variables.

Treatment timing:

$$g_i = \begin{cases} 5 & \text{if } i = 1, \ldots, 2500, \\ 7 & \text{if } i = 2501, \ldots, 5000, \\ \infty & \text{if } i = 5001, \ldots, 10000. \end{cases}$$

Treatment indicator:

$$D_{it} = \mathbf{1}[t \geq g_i].$$

Here, g_i indicates the time period when unit i receives treatment, and D_{it} is an indicator for whether unit i is treated in period t.

Outcome:

$$Y_{it} = \mathbf{1}[t \geq g_i \ \& \ g_i = 5] + 2\mathbf{1}[t \geq g_i \ \& \ g_i = 7] + \varepsilon_{it},$$

where

$$\varepsilon_{it} \sim N(0, 1).$$

The treatment has no pre-period effects, a post-period effect of 1 for the first treated group, and a post-period effect of 2 for the second treated group.

1.1

Clearly the average post-period effect for a treated unit is 1.5. What is the average post-period effect for a treated *unit-period*? In other words, calculate $\text{ATT}(g, t)$ for all post-periods and average across all treated unit-periods.

We have

$$\text{ATT}(g, t) = \begin{cases} 1, & \text{if } g = 5 \text{ and } t = 5, 6, 7, 8, 9, 10, \\ 2, & \text{if } g = 7 \text{ and } t = 7, 8, 9, 10. \end{cases}$$

and averaging the post-period effect for a treated-unit period, we get

$$\frac{2500 \cdot 6 \cdot 1 + 2500 \cdot 4 \cdot 2}{2500 \cdot 6 + 2500 \cdot 4} = \frac{25000}{35000} = 1.4.$$

Namely, since the earlier treated group spends more time being treated in the sample, it's (smaller) effect is weighted more than the later treated group.

1.2

Run a two-way fixed effects regression with unit and time fixed effects:

$$Y_{it} = \alpha_i + \theta_t + \beta D_{it} + u_{it}. \quad (1)$$

What estimate do you get for β? How does it compare to the average effects from part 1.1? Explain why.

Running the TWFE regression, we get $\widehat{\beta}_{TWFE} = 1.514$. The difference from the treated-unit-period average effect (1.4) arises because the two-way fixed effects estimator is a diff-in-diff estimator that does not weigh each treated unit-period equally. Instead, Goodman-Bacon shows it weighs treatment effects according to group sizes AND the residual variation in treatment timing after removing unit and time fixed effects. Since this variation is the same for both treated groups (4 periods treated, 6 untreated, or vice versa), the TWFE regression gives approximately equal weight to the $g = 5$ cohort effect of 1 and the $g = 7$ cohort effect of 2, yielding ~ 1.5.

1.3

Run the same regression *without the untreated group*. How do your estimates change? Explain why.

We get $\widehat{\beta}_{TWFE,-(g=\infty)} = 1.499$. Modulo noise terms, the estimate is essentially unchanged from part 1.2. Without the untreated group, identification comes only from comparisons between the early-treated cohort and the later-treated cohort. In this simulation, treatment effects are constant over time within each cohort, so more weighting on the "forbidden" comparison using the earlier-treated group for the later treated group does not introduce any dynamic treatment effect bias, so the TWFE estimate remains about 1.5, for reasons as before.

1.4

Suppose instead that the second treated group now has $g_i = 9$ instead of $g_i = 7$. Rerun the two-way fixed effects regression, including the untreated group. How do your estimates change? Explain why.

We get $\widehat{\beta}_{TWFE} = 1.387$. As Goodman-Bacon shows, the TWFE estimator weighs 2x2 comparisons by the variance of the treatment dummy within each comparison. Given our time frame of the sample and how the later treated group now only has 2/10 periods spent treated, their treatment variance has decreased relative to that of the earlier treated group, meaning their effect of 2 is weighted less, and the earlier group's 1 is weighted more, causing the dip in the estimate.

1.5

Construct $\text{ATT}(g, t)$ for each cohort and treatment date using the Callaway-Sant'Anna method and using only the untreated group as a control. Aggregate these estimates by averaging $\text{ATT}(g, t)$ across all treated unit-periods, i.e. weighting each treated observation equally. How does this estimate compare to your answer in part 1.1 and to the estimates from parts 1.2–1.4?

The estimated average cohort-time treatment effects $\text{ATT}(g, t)$ using the untreated group as a control are in Table 1. Averaging across all treated unit-periods, weighting each equally, gives an ATT of ~ 1.256. The Callaway-Sant'Anna estiamate differs from the TWFE estimate from 1.4 (1.387) because TWFE uses weights based on the variance of the treatment dummy (and can

g <dbl>	t <int>	ATT_hat <dbl>
5	5	1.0004602
5	6	0.9804877
5	7	1.0129950
5	8	1.0014043
5	9	0.9929985
5	10	0.9960845
9	9	2.0341511
9	10	2.0333843

Figure 1: Average cohort-time treatment effects

compare treated cohorts to other treated cohorts). The Callaway-Sant'Anna estimator instead uses only untreated units as controls, and importantly, weighs each treated unit-period equally so more overall weight is placed on the effect of 1, recovering the treated-unit-period weighted average effect. The estimate differs from 1.1, 1,2, and 1.3 (1.4, ~ 1.5, and ~ 1.5), because now the later treated group spends less time in the sample being treated, so its effect of 2 receives less weighting across unit-periods.

1.6

Return the second treated group to $g_i = 7$. Let the outcome Y_{it} now have the following form:

$$Y_{it} = \mathbf{1}[t = g_i] + 2\mathbf{1}[t = g_i + 1] + 3\mathbf{1}[t \geq g_i + 2] + \varepsilon_{it},$$

where

$$\varepsilon_{it} \sim N(0, 1).$$

The treatment now has:

- no pre-period effects,
- an effect of 1 in the first treated period,
- an effect of 2 in the second treated period, and
- an effect of 3 in all subsequent periods.

The average post-period effect for the first treated group is 2.5, and the average post-period effect for the second treated group is 2.25.

Run a two-way fixed effects regression like Equation (1). What estimate do you get for β? How does it compare to the average post-period effects?

We get $\widehat{\beta}_{TWFE} = 2.210$. This is below the treated-unit-period average effect,

$$\frac{6(2.5) + 4(2.25)}{6 + 4} = 2.4.$$

Treatment effects are now dynamic, so we must be wary of the forbidden comparisons in the TWFE regression. For the early-treated group, the treatment effect rises from 1 in the first treated period, to 2 in the second treated period, and to 3 in later periods. Thus, once the later-treated group becomes treated, the early-treated group is not a valid control group because its own treatment effect is still evolving over time. The TWFE estimator nevertheless includes 2×2 comparisons that use the early-treated group as a control for the later-treated group. These comparisons bias the estimate downward relative to the average post-period treatment effect.

1.7

Rerun the same regression without the untreated group. What happens to your estimate?

We get $\widehat{\beta}_{TWFE} = 1.124$, so it gets biased even more downwards. Intuitively, with no more untreated group to use as a control (which will not experience the dynamic treatment effects), TWFE places all weighting on the 2x2 comparisons between treated groups. In particular, more weight is placed on the forbidden comparison with the early-treated group as a control, and since dynamic effects are increasing, this biases the estimate downwards.

1.8

Rerun the same regression, but include the untreated group again. Give the untreated group a weight of 1000 in the regression and the treated groups a weight of 1. What is your estimate for β, and how does it compare to the average post-period effects?

We get $\widehat{\beta}_{TWFE} \approx 2.391$. This is much closer to the treated-unit-period average 2.4 than the unweighted TWFE estimate (2.210). The TWFE regression puts more weight on treated-versus-untreated comparisons and less on from forbidden comparisons between already-treated and later-treated cohorts, so we mitigate the dynamic trends bias. However, the estimate is not exactly equal to 2.4 because TWFE still uses variance-based weights rather than weighting each treated unit-period equally. Since the two cohorts have symmetric treatment variation, TWFE places roughly equal weight on these two cohort-level post-period effects of 2.5 and 2.25 (roughly 2.375) So, the weighted TWFE estimate is close 2.375.