Homework for Chapter 18: Difference-in-Differences

*How Does It Work?*

1. In the Event Studies chapter we estimated the effect of something that occurs at a specific time by just comparing before-event to after-event, without really using a control group. What assumption is made by no-control-group event studies that we *don’t* have to make with difference-in-differences?

The assumption of predicting the counterfactual. Event studies assumes that if we have plenty of information from the before-event times and we can make an assumption that whatever was going on before would have continued doing its thing if not for the treatment (nothing else changes at the same time), then we can use that before-event data to predict what we’d expect to see without treatment, and then see how the actual outcome deviates from that prediction (the extent of the deviation is the effect of treatment). In doing this, event studies try to use before-treatment information to construct a counterfactual after-treatment untreated prediction. Difference-in-differences (DID) takes a different approach by bringing in another group that is never treated.

1. Which of the following potential back doors is controlled for by comparing the treated group to a control group? b
   1. The treated group may be following a trend, unique to the group, that would make the outcome change from before-treatment to after-treatment anyway
   2. There may be events affecting everyone that would change the outcome from before-treatment to after-treatment anyway
   3. There may be differences in typical outcome levels between the treated group and the untreated group
   4. The decision to treat the treated group, rather than some other group, may be based on factors that are related to the outcome
2. Consider a treatment and control group. Looking only at the pre-treatment period, they have exactly the same outcomes (zero gap between them in each period).
   1. Despite having exactly the same outcomes pre-treatment, it happens to be the case that parallel trends is violated for these two groups. How is this possible? Explain what it means for parallel trends to be violated in this case, or give an example of how it could be violated.

The issue does not lie in the fact that the trends for the two groups were diverging or not pre-treatment, but rather that the trends would have continued to diverge post-treatment. The difference-in-differences design is to use the change in the untreated group to represent all non-treatment changes in the treated group. That way, once we subtract the untreated groupʼs change out, all weʼre left with is the treated groupʼs change. Parallel trends is necessary for us to assume that works. If, without a treatment, the gap between the two groups would have changed from the pre-period to the post-period for any reason, or for no reason at all, then that non-treatment-related change will get mixed up with the treatment-related change, and we won’t be able to tell them apart. When we have a particular reason to believe that the untreated group would suddenly change around the time of treatment or the treated and untreated groups are not similar in many ways, even if their pre-treatment trends were the same, the parallel trends are to be violated in this case.

* 1. If we estimate the causal effect in this case using difference-in-differences, even though parallel trends is violated, how much would our effect be off by? (note you won’t be able to give a specific number)

The effect we get will be off by the difference in outcome changes between the treated and untreated groups in terms of being differently impacted by the time effect that is taken by other variables in addition to our treatment (*OtherTreatedGroupChanges*- *OtherUntreatedGroupChanges*). For example, if the difference between the groups keeps diverging along time, we’d probably overestimate the treatment effect compared to the actual effect of our treatment. If the difference between the groups keeps shrinking along time, we’d probably underestimate the treatment effect compared to the actual effect of our treatment.

1. Consider the below graph showing the average outcome for treated and control groups in the leadup to treatment (indicated by the dashed line), and also after treatment.  
   Chart, line chart

   Description automatically generated
   1. Based on the prior trend, does it seem likely that parallel trends holds in this instance?

No. Though the distance between the treated and untreated group stays roughly constant in the leadup to treatment (t5-t6) (both are trending upwards), we should still be cautious and look for more evidence for the parallel trends assumption since the trajectories for the treatment and control groups in the pre-treatment period are quite different from t0 to t5: the distribution of the treated group is multimodal (has ups and downs) while the outcome variable for the control group overall increases constantly. The trend for the treated group just before the treatment that is seemly similar to that for the untreated group is only a part of the treated group’s periodic changes and we have reasons to expect that trend will diverge from the trend of the untreated group afterwards (post-treatment).

* 1. If we estimate difference-in-differences anyway, are we likely to overestimate the actual causal effect, underestimate it, or get it right on average?

We are likely to underestimate the actual causal effect because the dependent variable for an ideal control group would probably have gone through a periodic decline after t6, just as the treated group without the treatment (the counterfactual) under parallel trends assumption, and then rise again after around t7, which should have led to larger distances between the outcome values for the treated and control groups than we actually get now with the not-so-ideal untreated group (as illustrated on the graph).

1. In mid-2020, during the COVID-19 pandemic, different countries pursued different courses of action. Some locked down fully, imposing harsh penalties to most people for leaving the house outside certain proscribed times. Some were looser and only suggested staying at home, and some had hardly any restrictions at all. You notice that COVID rates tend to spike dramatically in different countries at seemingly-random times, and want to know if certain restrictions helped.  
   From March through May 2020, US and Canada COVID case rates followed similar trends (US rates were higher, but the trends were similar). You want to look at the effect of COVID restrictions enacted in Canada in late May 2020 on case rates. Is DID, with the US as a control group, a good way to estimate this effect? If not, what concerns would you have about this research design?

I don’t think DID is a good way to estimate this effect with the US as a control group, or at least, we need to do more (e.g., a placebo test) to provide more evidence to let us believe the parallel trend holds. Though the treated (Canada) and untreated (US) groups had similar trajectories for the dependent variable (COVID case rates) before treatment (COVID restrictions), it is relatively a short time window for us to state that the two groups have similar long-term trend. The infectivity of a pandemic could have made the COVID case rates spike in countries with different attributes. Canada and US are still different groups in many ways, such as different healthcare systems and weather, which could all cause differences in the COVID case rates in the long run.

1. Consider the below table of mean outcomes, and calculate the difference-in-difference effect of treatment. Write out the equation you used to calculate it (i.e. show how the four numbers in the table are combined to get the estimate)

|  |  |  |
| --- | --- | --- |
|  | Before | After |
| Treated | 5 | 9 |
| Untreated | 6 | 7.5 |

(9-5)-(7.5-6)=2.5

*How is it Performed?*

1. You are planning to estimate whether voter-protection laws increase voter turnout. You note that, in 2015, a lot of new voter-protection laws were enacted in some provinces but not in others. Conveniently, no new laws were enacted in 2012, 2014, or 2016, so you decide to use 2012 and 2014 as your “before” periods and 2016 as “after”.
   1. Which of the following best describes what you’d want to regress state-and-year level “voter turnout” measures on? iv.
      1. An indicator for whether the state is treated, and an indicator for whether the year is 2016. + an interaction term, an indicator for being in the treated group AND in the post-treatment period 2016
      2. ~~A set of fixed effects for state, and a set of fixed effects for year.~~
      3. An indicator for whether the state is treated, ~~a set of fixed effects for year~~ an indicator for whether the year is 2016 (AfterTreatment), and an indicator for whether the state is currently treated. = βTreatmentGroup\*AfterTreatment
      4. A set of fixed effects for state, and for year, and an interaction between “is 2016” and “is a treated state”.= an indicator for being in the treated group AND in the post-treatment period
      5. This design should not be estimated using a regression.
   2. Unless you chose the final option in the previous question, specify which coefficient in that regression would give you the DID estimate.

The coefficient on the interaction term between “is 2016” and “is a treated state” would give us the DID estimate.

1. You are looking at a difference-in-difference design to estimate the effect of providing laptops to school children on their test scores. Look at the below regression output, in which “Treated” is an indicator that the school received laptops in 2008 as part of a new program (the untreated group did not receive any laptops until years after the sample window for this study ended), and “After” is an indicator for being after the year 2008.  
     
   Using the table, fill in the blanks in the sentence “\_Assuming that parallel trends in the test scores hold for the treated and untreated schools within the time window we looked at\_, the effect of laptops on test scores was \_ 5.034 (introduction of laptops in some schools saw an increase in students’ test scores that was 5.034 unit larger in treated schools than it was in the untreated schools)\_\_, and this effect (**was**/~~was not~~) statistically significant at the 95% level.”

|  |  |
| --- | --- |
|  | **Test Scores** |
| (Intercept) | 80.342\*\*\* |
|  | (0.501) |
| After | 3.369\*\*\* |
|  | (0.696) |
| Treated | 4.116\*\*\* |
|  | (0.718) |
| After× Treated | 5.034\*\*\* |
|  | (0.993) |
| Num.Obs. | 1523 |
|  | 0.188 |
| Standard errors in parentheses.  + p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001 | |

1. A standard “prior trends” test might estimate a regression using the model (only using data from before-treatment), where is a time variable, is an indicator for being in the treated group, and is an outcome variable, and look for a large/significant estimate of . Explain why this test is performed, and specifically what it shows.

This statistical test is performed to see if the prior trends are different, i.e., whether the treated and untreated groups already had differing trends in the leadup to the period where treatment occurred, and if so, how much different. The interaction term allows the time trend to be different for each group. A test of = 0 provides information on whether the trends are different. Finding that =0 is unlikely shows that the trends are different. A large/significant estimate of shows the difference in outcomes between the treated and untreated groups could change from before treatment to after treatment for other reason besides treatment, and we don’t know the actual treatment effect from the DID estimate.

1. Consider the below graph with estimates from a dynamic difference-in-differences model for a treatment that occurs between periods 4 and 5, with 95% confidence intervals shown.  
   Chart, line chart

   Description automatically generated
   1. What about this graph might make us concerned about our identification assumptions?

The pre-treatment effects are not close to 0 until period 2. The confidence interval for period 1 is entirely above zero. The confidence intervals from period 1 to period 3 all have parts that are above zero. This is not ideal. We should further look at if the deviations are fairly small in their actual value. We also see three positive effects for the three periods after treatment goes into effect. The impact appears to be immediate but inconsistent. We see three positive effects for the periods after treatment goes into effect. The confidence intervals for the first two post-treatment periods are above zero. The confidence interval of the difference-in-differences estimate includes 0 around period 7. The treatments become gradually more effective over the first period after treatment, and less effective since then.

* 1. Ignoring any concerns we have, what would we say is the effect of treatment on Y in this case? (note the height of the line in period 5 is about 3, in period 6 is about 1, and in period 7 is about .5).

(3+1+0.5)/3=1.5

Ignoring any concerns we have, the treatment effect on Y in this case is 1.5.

1. Chapter 18.2.5 points out a problem with two-way fixed effects in cases where treatment is not all assigned at the same time, but rather different groups get treated at different times (a “rollout” design). In these designs, two-way fixed effects treats “already-treated” units, who were treated in earlier periods, as “control” units, as though they hadn’t gotten treated at all. However, there’s nothing theoretically wrong about using an already-treated unit as a control; the DID assumptions don’t require that the control group be untreated, just that the gap between treated and control doesn’t change when the treated group’s treatment goes into effect. Why are we so concerned, then, about using an already-treated group as a control? You can answer generally, or use as an example a DID with only two groups – an already-treated group and a newly-treated group. (hint: to do the example, try assuming the treatment only has an effect for the single period after treatment, and the already-treated group is treated exactly one period before the treated group)

Theoretically, we can use a treated group as the comparison, but if the effect itself is dynamic or the treatment effect varies across groups, by using two-way fixed effects we would set the estimation up in a way so that parallel trends won’t hold. If we have a treatment effect that gets stronger over time, for example, then the “treated comparison group” should be trending upwards over time in a way that the “just-now treated group” shouldn’t. Parallel trends breaks and the identification fails. The problems can get even stranger when it comes to things like the dynamic treatment effects estimator. The effects in the different periods start contaminating each other.

Coding (which includes any how-the-pros-do-it questions)