EDLD 650: Data Analysis and Replication Exercise (DARE) 1

David D. Liebowitz

Due: Jan. 17, 2022

You will replicate and extend analysis from Liebowitz, Porter and Bragg. In the file EDLD_650_DARE_1.csv, you will find the variables listed in Table 1. These data have been aggregated to the state level and altered to preserve confidentiality; thus your results will differ (slightly) from those in the paper. Please make sure that your graph axes are labeled, the variables in your tables have comprehensible labels, and your table notes are complete.

Table 1: Variable definitions

Variable name	Description
school_year	School year (fall)
state_id	State id (numeric)
state_abbrev	State name (char)
eval_year	Year of evaluation reform
class_remove_year	Year of reform to removal from class
suspension_year	Year of reform to suspension
PBIS	Schools in state successfully implementing PBIS
enroll	Total enrollment (#)
FRPL_percent	% students receiving free- or reduced-price lunch
enroll_WHITE	Enrollment of White students (#)
enroll_BLACK	Enrollment of Black students (#)
enroll_HISP	Enrollment of Hispanic students (#)
enroll_AM	Enrollment of AI/AK students (#)
enroll_ASIAN	Enrollment of Asian/PI students (#)
enroll_OTHER	Enrollment of other multi-racial students (#)
ODR_class	Per-500 stu. per-day rate of classroom-originating ODRs
ODR_other	Per-500 stu. per-day rate of other-location-originating ODRs
ODR_subjective	Per-500 stu. per-day rate of classroom-originating "subjective" ODRs
ODR_objective	Per-500 stu. per-day rate of classroom-originating "objective" ODRs

A. Data Management Tasks (1 point)

For these tasks, no write up is required. The code you submit will be sufficient.

A1. Convert the raw counts of enrollment by race/ethnicity into percentages (i.e., divide the enrollment count for each ethno-racial category by total enrollment). For programming efficiency, can you use a function to do this task?

A2. Generate dichotomous policy predictor variables that take the value of 1 in state-year observations in which the policy is in place. Call them **eval**, **class_remove** and **suspension**. They should take the value of 0 in years during which these policies were *not* in place. Also, generate a running time variable (**run_time**) that reflects how far or close the state-year observation is from the implementation of higher-stakes teacher evaluation and a variable that permits the effects of the evaluation policy to vary (linearly) over time (**evalXyear**). How will you deal with states that never implement evaluation? Do that too.

B. Understanding the Data and Descriptive Statistics (3 points)

For the following tasks, give your best attempt at completing the analysis and write-up. If you are unable to conduct the programming or analysis, describe what you are attempting to do and what your results would mean.

B1. Inspect your data. What sorts of missingness exist within the data file? What sorts of missingness should concern you? Which do not? In this assignment, please restrict your sample to state-years with non-missing outcomes.

B2. Graphically display the distribution of the outcome data. What do you notice about the distribution of outcomes? Are there any actions, transformations or sensitivity tests you would like to conduct based on this evidence?

B3. What is the analytic sample from which you will draw your inferences? To what population are you drawing these inferences? For this analytic sample, reproduce Column 1 of Table 1 from Liebowitz, Porter & Bragg (2022) to create a summary of descriptive statistics for the following data elements. All of these statistics (except for state-year and year enrollment) should be weighted by the state-year population:

- Mean state-year enrollment

- Mean year enrollment

- % low-income (FRPL)

- % Am. Indian/Alask. Native

- % Asian/PI

- % Black

- % Hispanic

- % White

- % state-year observations in which PBIS was successfully implemented

- Classroom ODR rate

- Other location ODR rate

- Subjective-Classroom ODR rate

- Objective-Classroom ODR rate

Describe the characteristics of your sample as you would report these statistics in an academic paper. How are the characteristics of the sample you will be using for this replication exercise different from the sample in Liebowitz, Porter & Bragg (2022)? How, if at all, do you anticipate this will affect your results?

B4. ***Optional Extension*** Plot the average classroom (**ODR_class**) and classroom-subjective ODRs (**ODR_subjective**) by how close the state-year observation is to the implementation of the teacher evaluation policy ***for the states that implemented evaluation reform***. (*Note: this is similar to Figure 2 in the original paper*). What do you notice about the raw outcome data plotted against the secular trend? Are there any actions, transformations or sensitivity tests you would like to conduct based on this evidence? Why do we stress plotting these raw averages only for states that implemented evaluation reform? How would including these states alter the interpretation of this figure?

C.   Replication and Extension 6 points)

For the following tasks, give your best attempt at completing the analysis and write-up. If you are unable to conduct the programming or analysis, describe what you are attempting to do and what your results would mean.

C1. Estimate the effects of the introduction of higher-stakes teacher evaluation reforms on Office Disciplinary Referrals. In one of your models, assume that the effects are constant and in another relax this assumption to allow the effects to differ (linearly) over time. Present these difference-in-differences estimates in a table and the associated write-up as you would report these results in an academic paper. Do you notice any important differences in these results and those reported in the original paper? If so, how would you consider addressing them (*it is not necessary at this point for you to actually conduct the analysis, just describe approaches you might take*)?

C2. Liebowitz et al. (2022) conduct a broad set of robustness checks. For this DARE assignment, you will conduct **two (2)**. First test whether the main results you present in Question C1 are robust to the introduction of potentially simultaneous discipline policy reforms. Present the table and associated write-up as you would report these results in an academic paper. Then select an additional robustness check (either from the paper or not) and present evidence on whether your findings are sensitive to this test.

C3. Write a discussion paragraph in which you present the substantive conclusions of your results about the effects of the introduction of higher-stakes teacher evaluation on ODRs.

C4. ***Optional Extension*** Use an event-study approach to this difference-in-differences research design to estimate the effects of the introduction of higher-stakes teacher evaluation reforms on Office Disciplinary Referrals (ODRs). Present these findings in an event-study graph. Present the figure and associated write-up as you would report these results in an academic paper. Do you notice any important differences in these results and those reported in the original paper? If so, how would you consider addressing them (*At this point, it is not necessary for you to actually conduct the analysis. Just describe approaches you might take.*)?

C5. ***Optional Extension*** Use one (or more) approaches to present the extent to which the successful implementation of Positive Behavioral Intervention and Supports (PBIS) framework moderating the effects of the introduction of higher-stakes teacher evaluation policies. Present these difference-in-differences estimates and associated write-up as you would report these results in an academic paper. Do you notice any important differences in these results and those reported in the original paper? If so, how would you consider addressing them?