EDLD 650: Data Analysis and Replication Exercise (DARE) 2

David D. Liebowitz

Due: Jan. 31, 2022

In this project, you will replicate and extend analysis from Holden (2016). In the dataset EDLD_650_CA_school_es.dta, you will find the variables listed in Table 1. The dataset includes school years 2003-2009. Without binning the forcing variable, your figures will look slightly different than those in Holden. This is fine and will not affect the substantive conclusions. If you would like to set yourself a data management challenge, attempt to mirror the figures exactly! Pay careful attention throughout to which years you should be using in your analyses.

A. Assumption tests (4 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.

A1. Create a figure that describes whether the forcing variable predicts the question variable of interest. Specifically, did a school with an Academic Performance Index (API) of 643 or lower *in 2003* receive additional instructional material funding *in 2005*? Present the figure and associated write-up as you would report these in an academic paper in 2-3 sentences.

A2. Is there evidence of schools manipulating their placement around the discontinuity? Present at least two figures demonstrating (a) whether there is evidence that schools attempted to receive an API score that would have made them eligible to receive additional funding; and (b) whether there is evidence that schools that did and did not receive Williams funding were different in observable ways. These figures should present characteristics of schools that are exogenous to the receipt of Williams funding (*think about what would be exogenous in*

Table 1: Variable definitions

Variable name	Description
cds	Unique school id
sname	School name
distid	District id
year	School year (spring)
readingscore	Mean reading scaled score
mathscore	Mean math scaled score
api_rank	Academic performance rank in 2003; determines eligibility for Williams
yrs_teach	Average years of teaching experience
yrs_dist	Average years of teaching experience in district
pct_ai	Percent of American Indian students
pct_as	Percent of Asian students
pct_pi	Percent of Pacific Islander students
pct_fi	Percent Filipino students
pct_hi	Percent Hispanic students
pct_aa	Percent African American students
pct_wh	Percent White students
pct_other	Percent students w/ missing or mult responses
percentfrl	Percent free- and reduced-price lunch
total	Total school enrollment
classsize	Average class size
average_score	Avg. standardized math/reading score by school
api_mean	Avg. standardized math/reading score by API rank
norm	Centered API score
ind	Intended recipient of Williams based on 2003 API score (0/1)
ind_norm	Interaction of centered API (norm) and intent to treat (ind)
receive_williams	Actual recipient of Williams in 2005 (0/1, NA if year!=2005)

this case). Present the figures and associated write-up as you would report these in an academic paper in 1-2 paragraphs.

A3. ***Optional Extension*** Construct a table of summary statistics. What are the sample characteristics of the subset of elementary schools in our dataset? What is different about our data from the full data used in Holden (2016) and how might that affect the interpretation of results? Present the table and associated write-up as you would report these in an academic paper in 1-2 paragraphs.

B. 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. For these tasks, think about what bandwidth is the correct one to present for your main results.

B1. Did the receipt of additional funds for instructional materials improve test score outcomes for elementary students in California? Construct a figure that presents graphical evidence in support of your answer to this question. Pay close attention to the bandwidth of analysis Holden selects. Either use the same bandwidth or justify a different selection. Write up your description of this figure in one (1) paragraph as you would for an academic paper. If you have not completed the previously mentioned data management tasks to bin the forcing variable, explain how and why your figure looks different from the original figure in 2-3 sentences.

B2. In a regression framework, formally test whether the receipt of additional funds for instructional materials improved test score outcomes for elementary students in California. Present these results in a table and associated 1-2 paragraph write-up as you would in an academic paper.

B3. Write a discussion paragraph in which you present the substantive conclusions (and limitations) of your results about the effects of added textbook funding for California elementary school students.

B4. ***Optional Extension*** Present a series of robustness checks to the main results you have found. Consider varying year, functional form and bandwidth of your estimates.