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

David D. Liebowitz

Due: Feb. 5, 2024

A.   Assumption tests (4 points)

A1. Our identification strategy relies on the fact that California elementary schools in the bottom two deciles of performance received supplemental funding for instructional materials as a result of the Williams settlement. For our regression discontinuity approach to be successful, schools' API score (our *forcing variable*) must predict differences in the probability that schools actually receive additional funding. Figure 1 plots the probability of receipt of this funding in 2005 by 2003 API score, centered at the elementary school API cutoff score (643). In Panel A, we present all schools and in Panel B we aggregate schools in bins of 3 API scores for visual clarity. In both instances, our forcing variable (API score) is a strong (in fact, nearly perfect) predictor of receipt of Williams funding.

A2. Our modeling approach depends on our ability to project a smooth relationship between API score and student test scores, that units (schools) did not manipulate their position around the discontinuity in textbook funding and that schools immediately on one side of the discontinuity were not observably different than schools on the other side. In Figure 2, we present evidence that there is no evident bunching of schools just to the left of the discontinuous award of textbook funding. While there are slightly more schools with API scores that just qualify them for additional funding, the discrepancy is similar to other random variation throughout the distribution of ±100 API scores. Thus, we are fairly confident that bunching does not threaten our assumptions.

In Figure 3, we address the possibility that schools that received Williams funding might not have been equal in expectation to those that did not. While we can never fully address the possibility that

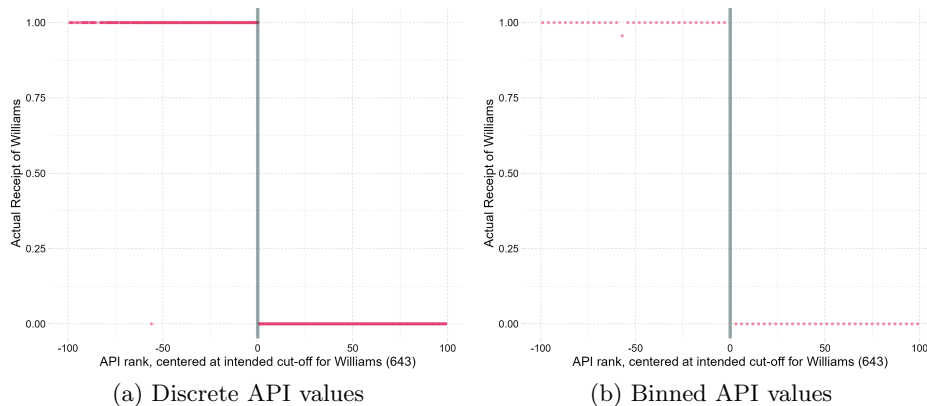(a) Discrete API values   (b) Binned API values

Figure 1: Probability of receipt of Williams funding by centered API score

unobservable attributes vary around the discontinuity, we observe no discontinuous jump in the proportion of students eligible for free- and reduced-price lunch. This provides suggestive evidence that no sorting occurred. Note that we display school characteristics for the 2002-03 school year (prior to determination of Williams eligibility). School characteristics post-cut score determination are endogenous to the receipt of funding and are therefore not valid tests for the presence of sorting.

A3. **_Optional Extension_** In Table 1, we present descriptive statistics on our analytic sample of 16,462 elementary schools. These low-performing elementary schools educate disproportionately Hispanic students from low-income backgrounds.[1]

In comparison to the full set of K-12 schools in the original Holden dataset, these schools are slightly smaller and have lower test score and API rank values. This is entirely a function of the fact that elementary schools are generally smaller in size and that their test and API scores are on a different scale. These differences, therefore, do not concern us. Elementary schools in California appear to be disproportionately Hispanic/Latino and poor, compared to middle and high schools. Pupil-teacher ratios are, surprisingly, only marginally smaller in elementary schools than in the full sample (around 20:1 for elementary compared to 21:1 for all schools).

---

[1]We use the term "Hispanic or Latino" (rather than Latino/a, Latinx or other nomenclature) because this is the federally recognized term used by the Office of Management and Budget/Census Bureau with which ethnicity data are harmonized at sub-national levels in the United States.
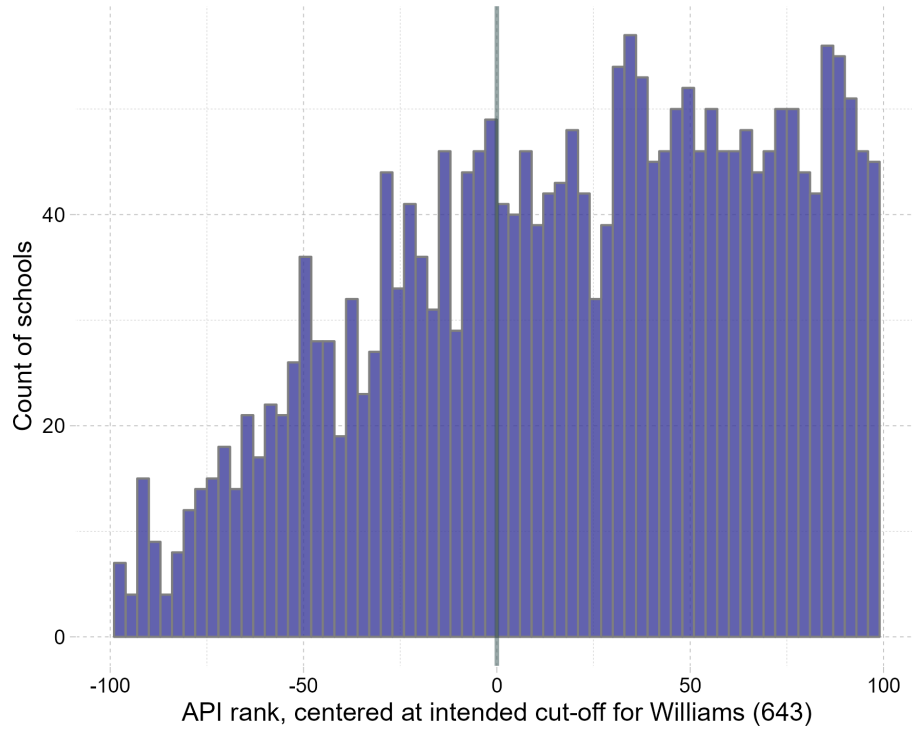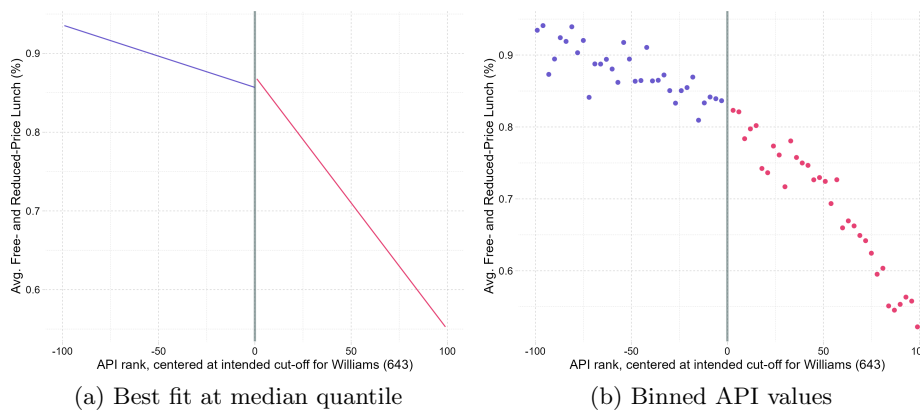
Figure 2: Count of schools by centered API score

(a) Best fit at median quantile

(b) Binned API values

Figure 3: Free- and reduced-price lunch proportion by centered API score

Table 1: Descriptive Statistics

Statistic	N	Mean	St. Dev.
Reading Score	16,462	324.70	14.51
Math Score	16,462	341.08	20.74
API rank in 2003	16,462	664.13	49.52
Total enrollment	16,462	631.69	259.92
Pct. Hisp.	16,462	65.81	24.66
Pct. White	16,462	15.72	18.32
Prop. FRL	16,462	0.76	0.18
Avg. teach. exp	16,462	12.83	3.14
Avg. class size	16,462	19.85	7.86

Notes: This table presents school-year means and standard deviations for elementary schools from 2003-2009.

B. Replication and Extension 6 points)

B1. California elementary schools that barely qualified to receive additional funding for instructional materials as a result of the Williams settlement appear to experience improved student test score outcomes as a result. In Figure 4, we present graphical evidence of a discontinuity in mean reading/mathematics performance around the cutoff for Williams funding. Panel A plots discrete API values in a bandwidth of ±50, an arbitrary distance from the cutoff, aligned to the visual display in Holden (2016). Casual visual inspection suggests a sharp increase in the magnitude of the relationship between textbook funding and student outcomes for slightly higher performing schools (i.e., schools right below the cutoff). It also is suggestive of a secular trend that may vary on either side of the funding cutoff.

Panel B plots results from API values binned in groups of three. The binning approach reduces random noise and visual clutter. This has the benefit of allowing us to expand the bandwidth of visual analyis to better understand the relationship between the forcing variable and the outcome. Panel B provides evidence of a linear relationship to the right of the discontinuity and potential curvilinearity to the left. We would want to explore the extent to which these functional forms fit our data in a formal regression framework.[2]

---

[2]Note that this binned figure differs in an important visual way from Figure 5, Panel B in Holden (2016) even though both construct bins in which API scores are rounded to the nearest 3. His binning function places a set of schools with an API score of 0 (and therefore

4

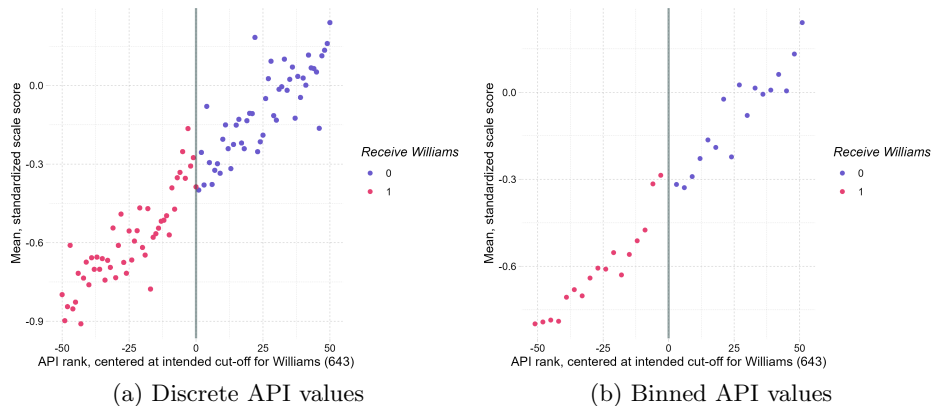(a) Discrete API values   (b) Binned API values

Figure 4: Effect of textbook funding on student achievement in elementary schools

B2. We formalize these graphical results by estimating a series of Ordinary Least Squares models in a regression discontinuity framework. Specifically, we estimate:

$$\text{TESTSCORE}_{it} = \alpha + \mathbb{1}\left(APISCORE_{i2003} \leq 643\right)\delta + \\ f\left(APISCORE_{i2003}\right) + u_{it} \tag{1}$$

where TESTSCORE_{it} is an average of standardized math and reading scores at the school level, the API score (the measure used to allocate differential funding) for elementary school i is measured in 2003. The causal parameter of interest in Equation 1 is the coefficient on the indicator for whether a school was, in fact, eligible for receipt of additional textbook funding (δ). The function $f(\cdot)$ is a flexible, continuous function for which we can vary the functional form and bandwidth.

In **??**, we present a taxonomy of regression discontinuity models of the effect of receipt of textbook funding on test scores.[3] All models test the effect for school years 2005 to 2009. Models 1, 2 and 4 estimate results in a bandwidth of 19.099 API rank score using the optimal bandwidth selection procedure specified in Calonico, Cattaneo and Titiunik

---

eligible for Williams) into the first bin to the right of the discontinuty (incorrectly). He corrects this by setting their bin value to missing, and so they do not contribute to the plot at all. Instead, we assign schools in our zero-rounded bin with API values of -1 or 0 to the -3 bin and schools with API of 1 to the 3 bin. This seems like a more sensible choice. The effect on the visualized discontinuity is dramatic and highlights the importance of seemingly small choices in data management procedures.

[3]Note that these estimates are slightly different from those presented in Holden (2016) as he corrects his standard errors for potential homoskedasticity and weights estimates slightly differently than in standard OLS, though these differences are marginal.

(2014). Model 3 extends the bandwidth to 25 API rank scores. We select the latter range given the evidence that Holden (2016) presents in Figure 8 on the relative stability of coefficient estimates through that bandwidth. Model 1, which imposes a common secular trend on both sides of the discontinuity implies that the receipt of Williams funding improves student test scores by 0.18 school-level standard deviations (σ). As suggested by graphical evidence, the highest performing schools that receive Williams funding experience the greatest benefit (Model 2). API rank scores are centered on zero and scores below zero are eligible for aid; as a result these coefficients require some interpretation. We predict that schools exactly at the discontinuity (API of 643) will have mean test scores 0.16σ higher than those that just exceed the eligibility threshold. We assign schools further from the cutoff negative API scores. Thus the coefficient on the interaction between Williams receipt and API rank reflects how much *less* benefit schools one API score point further away from the discontinuity receive from textbook funding. As an illustrative example, we predict a school with an API score of 638 (centered score of -5) would experience a 0.09σ benefit from Williams eligibility ($0.163 + 0.014 \times (-5) = 0.093$).

Extending the bandwidth to 25 API points in either direction (Model 3) results in somewhat attenuated effects (0.11σ), but still statistically different than zero and substantively similar in magnitude to our main estimates. Finally, a quadratic specification for the secular trend returns a nearly identical main effect of textbook funding receipt. For reasons of simplicity and completeness, we adopt Model 2 as our preferred specification.

B3. Our analysis suggests that increased textbook funding in California between 2005 and 2009 resulted in improved mathematics and reading test scores at the elementary school level. We estimate that schools that were just barely eligible to receive an additional \$96.90 per student in instructional material funding experienced between a 0.10 and 0.18 standard deviation unit improvement in math and reading test scores.

B4. ***Optional Extension*** Incorporated into B2 above.

Table 2: Regression discontinuity estimates of the effects of instructional material funding on average math/reading test scores

	Linear, same slope	Linear, diff slope	Linear (+/- 25 API)	Quadratic
	(1)	(2)	(3)	(4)
Receive Williams	0.180***	0.163***	0.106**	0.171***
	(0.047)	(0.047)	(0.040)	(0.047)
API Rank	0.018***	0.011***	0.013***	
	(0.002)	(0.003)	(0.001)	
Receive Williams × API Rank		0.014**		
		(0.004)		
API Rank				10.165***
				(1.205)
API Rank sq.				−2.007**
				(0.618)
Observations	2,680	2,680	3,510	2,680
R^2	0.043	0.047	0.056	0.047

Notes: *p<0.05, **p<0.01, ***p<0.001. Cells report coefficients and associated standard errors. Estimates pool years from 2005 to 2009.