ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 19-25

School District Operational Spending and Student Outcomes: Evidence from Tax Elections in Seven States

Carolyn Abott
St. John's University

Vladimir Kogan
The Ohio State University

Stéphane Lavertu
The Ohio State University

Zachary Peskowitz
Emory University

We use close tax elections to estimate the impact of school district funding increases on operational spending and student outcomes across seven states. Districts with passing levies directed new revenue toward support services and instructor salaries but did not increase teacher staffing levels. These districts eventually realized gains in student achievement and attainment. Our preferred estimates imply that increasing operational spending by $1,000 per pupil increased test scores by approximately 0.15 of a standard deviation and graduation rates by approximately 9 percentage points. There is some evidence of diminishing returns, as these effects are driven by districts below the median in spending per pupil. Based on research linking academic outcomes to earnings, we conclude that these spending increases were likely cost-effective.

VERSION: January 2020

Suggested citation: Abott, Carolyn, Vladimir Kogan, Stéphane Lavertu, and Zachary Peskowitz. (2020). School District Operational Spending and Student Outcomes: Evidence from Tax Elections in Seven States. (EdWorkingPaper: 20-25). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/mdtk-8743

School District Operational Spending and Student Outcomes: Evidence from Tax Elections in Seven States*

Carolyn Abott
Department of Government & Politics
St. John's University
abottc@stjohns.edu

Vladimir Kogan
Department of Political Science
The Ohio State University
kogan.18@osu.edu

Stéphane Lavertu†
John Glenn College of Public Affairs
The Ohio State University
lavertu.1@osu.edu

Zachary Peskowitz
Department of Political Science
Emory University
zachary.f.peskowitz@emory.edu

January 12, 2020

Abstract

We use close tax elections to estimate the impact of school district funding increases on operational spending and student outcomes across seven states. Districts with passing levies directed new revenue toward support services and instructor salaries but did not increase teacher staffing levels. These districts eventually realized gains in student achievement and attainment. Our preferred estimates imply that increasing operational spending by $1,000 per pupil increased test scores by approximately 0.15 of a standard deviation and graduation rates by approximately 9 percentage points. There is some evidence of diminishing returns, as these effects are driven by districts below the median in spending per pupil. Based on research linking academic outcomes to earnings, we conclude that these spending increases were likely cost-effective.

Keywords: student achievement, education finance, school districts, tax elections
JEL codes: H11, I28

---

*Note: We have no conflicts to declare. The Spencer Foundation provided generous funding for this study via a Lyle Spencer Research Award (201600072). We are very grateful to James Szewczyk and Matthew Troy for leading much of the data collection, and to the 35 undergraduate research assistants who scanned and hand-entered data. We also thank Chris Candelaria, Jared Knowles, Kenneth Shores, Joe Smith from TexasISD.com, and hundreds of officials for providing us with data that made this analysis possible. Finally, we are very grateful for helpful feedback from Eric Brunner, Christopher Walters, and anonymous reviewers at the Journal of Public Economics, as well as feedback from participants of the 2019 meetings of the Association for Education Finance and Policy and the Association for Public Policy Analysis and Management.

†The authors are listed in alphabetical order. Please direct communications to the corresponding author at lavertu.1@osu.edu or by mail: John Glenn College of Public Affairs, The Ohio State University, 110 Page Hall, 1810 College Rd., Columbus, OH, 43210.

1 Introduction

Researchers have long debated the returns to education spending (e.g., see Hanushek, 2003). A wave of recent research, however, provides convincing quasi-experimental evidence that increases in school district funding can improve education outcomes (Jackson, 2018a). Notably, using plausibly exogenous shocks in district revenues from state finance reforms, Jackson et al. (2016), Lafortune et al. (2018), and Candelaria and Shores (2019) find that higher revenues are associated with achievement and attainment gains—particularly among districts serving low-income students. But there may be diminishing returns to education spending. It remains an open question whether increasing spending would improve student outcomes today, after many finance reforms have dramatically increased spending per pupil and narrowed spending gaps between high- and low-income districts. Likewise, the circumstances under which spending increases have an impact may have changed.

This study estimates the impact of recent increases in education spending across diverse districts in multiple states: Arkansas, Louisiana, Michigan, Missouri, Pennsylvania, Texas, and Wisconsin. Districts in these states can raise significant revenues by levying local property taxes—and, to a lesser extent, sales and income taxes—with voter approval to exceed revenue or tax limits (see Table A1 in Appendix A for details). There is variation across states and over time in terms of what these limits are, and there is significant variation in how often districts seek voter approval, the size of tax requests, and how long tax levies are in effect. These states also employ funding formulas that affect the ultimate impact of local tax levies on district revenues. For example, under the period of study, Missouri and Pennsylvania provided funding based in part on district tax effort, and Texas and Wisconsin equalized district tax bases such that low-wealth districts effectively received a state match for local revenues they generate. Indeed, as we show below, it appears that the aggregate revenue gains from passing a local tax levy are nearly evenly divided between state and local sources.

What districts in these states have in common, however, is that they have discretion in how they spend revenue from passing a tax referendum. Unlike local bond elections tied to capital spending and restricted federal grants, these local tax levies allow districts to allocate revenues to whatever operational functions they wish. With the exception of Kogan et al. (2017) and Baron (2019)—who focus on Ohio and Wisconsin, respectively—studies that leverage close elections to estimate the impact of school funding have focused on bond referenda tied to capital spending and have found negligible impacts on achievement (e.g., see Cellini et al., 2010 and Martorell et al., 2016; but see Hong and Zimmer, 2016). Our study's focus on district-initiated efforts to generate discretionary operational funds makes it distinct.

Specifically, we employ data on over 3,000 district tax referenda held between 2000 and 2015 to estimate the impact of local tax levies on spending and student outcomes using a regression discontinuity design.[1] Our analysis reveals that, on average, districts in which referenda passed spent \$400-\$500 more annually per student through 5-7 years after the election. These funds went toward support services and instructor salaries, but there was no increase in teacher staffing levels. Specifically, levy passage led district spending on instructor salaries to increase steadily with a peak of around \$4,000 more per teacher after five years—approximately 7 percent of the average instructor salary in our sample. This overall increase in district spending is associated with a steady climb in average test scores in math and English language arts, as well as an increase in four-year, cohort-based graduation rates. By 5-7 years after a tax referendum, districts in which tax levies passed had achievement gains of approximately 0.1 of a standard deviation and gains in graduation rates of 3-4 percentage points. Specifically, our preferred estimates imply that increasing operational spending by \$1,000 per pupil increased test scores by approximately 0.15 of a standard deviation and graduation rates by approximately 9 percentage points.

These results are comparable to those from Lafortune et al. (2018) and Candelaria

[1] Although we employ data on over 3,000 district elections, the effective sample of tax levies includes just over 940 referenda across 560 unique districts in which tax referenda were within 20 percentage points of passage and for which we observe achievement or graduation rates within five years of the election.

and Shores (2019), and they imply returns comparable to those of typical educational interventions (see Kraft, 2018). As we discuss below, if one monetizes the returns to achievement based on increases in students' earnings, our estimates yield benefit-cost ratios above 1—particularly for districts that spent below the sample median of $10,893 per pupil (in 2012 dollars). Importantly, the main results are robust to variation in model specification, variation in the sample bandwidth around the referendum passage threshold, and the inclusion of controls to account for potential changes in student composition. Thus, this study provides convincing evidence that district funding increases continue to have a pronounced impact on student achievement and attainment, even after decades of state finance reforms.

We also conducted heterogeneity analyses to explore a few important topics. First, we examined whether spending increases are more beneficial in districts serving disadvantaged student populations. It is unclear whether there would be such returns in wealthier districts, as their students should be less dependent on formal schooling to build their knowledge and skills. Indeed, there is some evidence that capital spending only has an impact on test scores among districts serving low-income students (Rauscher, 2019), and recent estimates of spending's impact are based primarily on finance reforms focused on low-income districts. Second, we tested whether achievement and attainment effects are larger in low-spending districts, to get a sense for whether there are diminishing returns. Finally, following Brunner et al. (2019), we tested whether spending and achievement effects differed depending on states' collective-bargaining laws, as one of the most salient debates in school finance concerns whether collective bargaining prevents districts from allocating resources toward their most productive uses (e.g., see Chubb and Moe 1990; Hoxby 1996).

These additional analyses are merely suggestive but yield some interesting results. First, as we note above, districts that spent below our sample median prior to putting a tax levy on the ballot realized greater achievement and attainment gains than those that spent above the median, even though absolute revenue and spending increases are comparable between high- and low-spending districts. Second, we find no statistically significant differences in spending, achievement, or attainment effects between districts with child poverty rates that are above our sample median (18.2 percent) and those with poverty rates below the median. Third, we find that elections held early in our sample period (prior to 2008) drive our spending and achievement results. Further analysis indicates that this is because referenda held during the Great Recession do not yield substantial revenues. Overall, there is evidence of diminishing returns to spending—increasing funding for high-spending districts is less likely to yield improvements in graduation rates and achievement on math and English language arts exams—and returns to spending are present regardless of district poverty rates and the extent to which districts serve children who qualify for free or reduced-price lunches.

Finally, we found no significant differences in spending and outcomes between districts in states with strong collective bargaining laws (Michigan, Pennsylvania, and Wisconsin) and those in states with weak or no collective bargaining (Arkansas, Louisiana, Missouri, and Texas) during most years of our study. Power issues are perhaps largely to blame, but the lack of differences may also be due to our analytic sample. Our sample consists entirely of districts that chose to raise additional revenues through a local tax referendum, in spite of the significant costs associated with doing so (e.g., the cost of administering an election). Moreover, our research design focuses on districts in which there was sufficient (or nearly sufficient) public support for increasing taxes to fund district operations, and the effects we detect are concentrated in low-spending districts. Thus, the districts in our sample likely had a relatively clear need for operational funds and likely differed significantly from those that responded to the influx of revenues from state finance reforms by paying off debt and lowering local tax rates (see Brunner et al., 2019).

That our results are driven by districts that likely had a great need for resources suggests that the achievement and attainment effects may be greater than they would have been for the average district. Nevertheless, this study provides convincing evidence that spending increases can have a significant impact on student achievement and attainment. To the extent that these results are generalizable to districts outside of our effective analytic sample, they support recent arguments that school finance policy could realize further gains by targeting funding toward low-spending districts. In the following sections, we describe our data and empirical strategy, review the results, and offer some concluding thoughts.

2 Data

First, we describe our tax election data—an original contribution of this study. Second, we describe the publicly available data we use to examine school district finances, staffing, and student composition, achievement, and attainment. Third, we describe the variables we use in the analysis.

2.1 Tax Election Data

We collected information on school district tax and bond referenda held from 2000 to 2015 across 20 states. In this study, we focus on tax referenda that enable districts to raise discretionary operational funds. Among states from which we collected local tax election data, we omit Oklahoma because various state requirements effectively render levy votes meaningless, as districts have no choice but to approve the maximum allowable rates (OSDE, 2017, p. 2); Nebraska, Maryland, South Carolina, and Virginia because there are few observations; Indiana because our data go back to 2008 only; and Illinois because we lack vote totals. Because public records across many counties make it difficult to determine vote totals in districts that cross county lines, we also purged the sample of referenda for which this is a problem.[2] Finally, we omitted 2007-08 referenda from Pennsylvania because the vast majority were related to a state initiative to shift some of the tax burden from local property taxes to local

[2]Specifically, because our records requests in Arkansas, Michigan, Missouri, and Pennsylvania were at the county level and some districts span county boundaries, we removed observations from districts that span boundaries because it was unclear whether a given county's totals reported all votes across the district or only those votes from the portion of the district in that county. Additionally, numerous election results (predominantly in Arkansas) indicated implausible vote totals and outcome margins. To address this issue, we removed tax measures from the sample if the vote total was under 1 percent of the voting age population. Because we focus our analysis in a neighborhood around the cutoff, elections associated with implausible margins should not affect the main analysis.

income taxes (which all but three of 501 districts rejected). Ultimately, we ended up with a sample of over 3,000 tax elections held from 2000 to 2015 across seven states: Arkansas, Louisiana, Michigan, Missouri, Pennsylvania, Texas, and Wisconsin.

[Insert Table 1 about here.]

Table 1 presents descriptive statistics for the full sample of 3,070 referenda and, in brackets, descriptive statistics for the sample within 20 percentage points of the passage threshold—which is a bandwidth that we feature in the analysis. As Table 1 reveals, restricting the bandwidth to increasingly close elections leads the passage rate to approach 50 percent in all states, which is the passage threshold for all but a handful of Missouri districts where a tax levy would result in an effective tax rate that exceeds 6 percent (see Table A1 in Appendix A). Indeed, as our density tests below confirm, near the threshold we have about as many passing and failing referenda. Note that we lack data from Texas prior to 2006, as there were no tax ratification elections yet.

2.2 School District Data

School district data are from public sources. We obtained school district finance, staffing, and student data from the National Center for Education Statistics' Common Core of Data (CCD). The achievement data are from the Stanford Education Data Archive (see Reardon et al., 2018). In particular, we obtained SEDA's 2009-2015 student achievement data for grades 3-8 in mathematics and English/Language Arts (ELA) standardized at the cohort-subject-grade level. These estimates are standardized nationally using NAEP data and can be used to estimate effect sizes in grade-specific standard deviations. In the analysis below, we use a single district-level average of these standardized scores, pooling across both subjects. Note that focusing on grades 3-8 means that eight years after a referendum, we observe outcomes for students who experienced greater spending for all (or nearly all) years of formal schooling.[3] Finally, we obtained four-year, cohort-based graduation rates from the websites

[3] Another advantage of focusing on grade 3-8 achievement data is that some districts in our sample (e.g., some in Michigan, Missouri, and Wisconsin) do not include high school grades.

of state education agencies. Most states adopted these graduation rate calculations after the U.S. Department of Education issued Title I regulations requiring them in 2008. Although this measure is unavailable until as late as 2011 in some states, like the SEDA achievement estimates, they have the significant benefit of being comparable across states.

[Insert Table 2 about here.]

For each tax referendum, we merged in data for the corresponding district from 3 years prior to the election to 9 years after the election, and then stacked these observations. Thus, multiple referenda from a single district enter as separate observations. (Below, we discuss the implications of including multiple proposals per district.) Table 2 presents descriptive statistics corresponding to referenda within 20 percentage points of the passage threshold and for which we observe achievement data in at least one of the first five post-election years, as this is the effective analytic sample in much of the analysis below. Statistics for the CCD variables are based on the year preceding the corresponding tax election, whereas SEDA standardized test score estimates (averaged across math and ELA) and graduation rates capture averages over the first five post-election years.

Table 2 reveals that the 566 unique districts in our restricted sample (those within 20 percentage points of the passage threshold and for which we observe achievement in the fifth post-election year) spent roughly around the national average per pupil ($11,185 in 2012 dollars) and that we observe more districts with at least one passing referendum (429) as compared to districts with at least one failing referendum (235). It is also worth noting that test score estimates are highly correlated with the pre-election percentage of district students who qualify for free or reduced-price lunches (FRL). This correlation is suggested by the statistics in Table 2, as states in which districts have the lowest percentage of FRL students (Missouri, Pennsylvania, and Wisconsin) have districts with the highest average student test scores—placing them around 0.15 of a standard deviation above average. States in which districts have the highest percentage of FRL students (Arkansas and Louisiana) have the lowest test scores—placing them around 0.2 of a standard deviation below average.

There is a similar relationship between achievement and spending per pupil. Finally, the table also illustrates the diversity of districts in our sample. Because we pool across such diverse districts, it is particularly important that we focus on districts close to the passage threshold and that we reduce the influence of outliers by scaling our variables appropriately. We also follow Brunner et al. (2019), Gordon (2004), and Lafortune et al. (2018) and remove observations for districts with enrollments below 250, because of the noise in finance (and achievement) data for these districts.

2.3 Primary Outcome Variables

The analysis below focuses on how sustained increases in district spending translate to salaries, staffing levels, and student outcomes through up to seven post-election years—about the amount of time recent studies have indicated that it takes to detect the impact of increased revenues on achievement and attainment (e.g., see Candelaria and Shores (2019) and Lafortune et al. (2018)). (We also find that revenue effects taper off and that the sample of referenda with test-score observations dwindles beginning in the seventh post-election year.) Because we are interested in the cumulative effects of increased spending—in addition to its immediate effects—some of the analysis below examines average increases in inputs across a number of post-election years. Specifically, for each tax measure i, much of the analysis focuses on the average of expenditure or staffing levels (Y) from post-election years $k = 0$ through $k = M$, where $M \in \{5, 7\}$:

$$Z_i^M = \frac{1}{M+1} \sum_{k=0}^{M} Y_i^k \tag{1}$$

We focus primarily on the analysis of averages in expenditures and staffing through year 5 (i.e., $Z_i^5 = \frac{1}{6} \sum_{k=0}^{5} Y_i^k$), but we also examine average effects through year 7 as a sensitivity check.[4]

[4] We include the school year in which elections occur ($k = 0$) because districts with passing or failing levies early in that school year might adjust expenditures or staffing before the school year is over. Thus, we scale the measure by the number of post-election years plus 1 ($M + 1$).

Unlike the expenditure and staffing variables, the district-cohort achievement estimates and graduation rates already capture the accumulated impact of greater resources since the election. In particular, because SEDA achievement estimates are based on test scores in grades 3-8, focusing on post-election years 5 or 7 means that we are focusing on achievement effects among students who experienced differential annual expenditures levels for most or all of their formal schooling. That is why the analysis below estimates achievement and attainment effects for the year in which we observe it, as opposed to using an average across years.

[Insert Table 3 about here.]

Table 3 provides a descriptive look at our primary outcome variables through five post-election years for referenda within 20 percentage points of the passage threshold and for which we observe test scores five years after the election. We have these variables for up to 946 referenda across 566 unique school districts. Expenditures per pupil are in 2012 dollars, and staffing variables capture student counts per full-time-equivalent (FTE) employee. We observe expenditure and staffing variables for all five post-election years for elections held in 2011 or earlier, as the CCD finance data are available through the 2015-16 school year. There are very few elections for which we observe achievement and graduation data for all five post-election years. We have achievement data across all K-8 or K-12 districts for school years 2009-2015, and we have graduation data across all K-12 and high school districts for school years 2011-2017 (though there are some states that enable us to observe graduation rates as early as 2008). Nevertheless, for the purposes of these descriptive statistics, we provide an average of achievement and attainment across the first five post-election years.

Finally, to increase precision and as a sensitivity check, some of the analyses below include lags of the dependent variables one and three years prior to the election. Unfortunately, we observe pre-election test scores and graduation rates for too few referenda to permit such a sensitivity analysis of our achievement and attainment estimates. To conduct such a sensitivity analysis, we instead created a proxy for drop-out rates by dividing 12th grade district enrollments by 8th grade enrollments four years prior and subtracting that value from 1. This variable is highly (and, of course, negatively) correlated with diploma counts divided by 8th grade enrollments (ρ=-0.92)—which Heckman and Lafontaine (2010) validate but which is available for only a very small subset of our referenda. Importantly, pre-election observations of our dropout proxy are (negatively) correlated with post-election test scores (-0.47) and 4-year cohort-based graduation rates (-0.49). They also remain significant predictors of post-election outcomes if we control for baseline student characteristics (fraction FRL, fraction Black, and fraction Hispanic). Thus, using our dropout measure as a proxy for lagged academic performance allows us to explore the sensitivity of our estimates.[5]

3 Empirical Strategy

We use a regression discontinuity design to estimate the impact of passing (as opposed to failing to pass) a local tax levy on school district spending, staffing, achievement, and attainment. Specifically, we estimate variants of the following OLS model:

$$Y_i = \tau Pass_i + f(Pass_i, Vote_i) + \mathbf{X}_i'\gamma + \epsilon_i \tag{2}$$

The outcome Y for tax measure i is a function of $Pass_i$, indicating whether (1) or not (0) a tax referendum passed; $f(Pass_i, Vote_i)$, which is a linear or quadratic function of the centered vote variable, $Vote_i$, with and without an interaction with $Pass_i$; and $\mathbf{X}$, which captures the following baseline district covariates: logged operational expenditures per student one and three years prior to the election, logged student counts one and three years prior to the election, and student demographics observed one year prior to the election (the fraction of

[5]Unfortunately, there are quite a few missing values of our dropout rate proxy due primarily to missing or implausible by-grade student counts. We exclude (i.e., code as missing) observations of dropout rates if 12th grade enrollments are twice as large as 8th grade enrollments. Removing these outliers yields a roughly normal distribution—in part because there remain implausible values, as dropout rates are sometimes negative. We do not bottom-code dropout rates to avoid introducing measurement error in districts that experience enrollment increases or declines.

students who are Black, the fraction of students who are Hispanic, and the fraction of students eligible for free or reduced-price lunches).[6] The baseline demographic variables are highly correlated (0.8-0.9) with achievement observed between 0 and 7 years after the election. We also include academic year fixed effects (which we roll into $\mathbf{X}$ to simplify the notation) to account for the calendar years in which we observe Y. As we show below, these baseline covariates and year fixed effects significantly reduce the noise in our estimates. Finally, note that in some specifications below, the outcome Y is an average across five or seven post-election years (i.e., Z_i^5 or Z_i^7 from equation 1), whereas in others it is an outcome observed in a single year relative to the election (from $k = -3$ to $k = 9$).

For greater precision, and to examine the sensitivity of our results, in some specifications we also include as covariates lagged values of the dependent variables observed one and three years prior to the election. Because we cannot observe student achievement or graduation rates prior to elections held in 2010 or earlier (the majority of our observations), we use instead the proxy for dropouts that we describe above (1 minus 12th grade enrollments divided by 8th grade enrollments four years prior). We further test the sensitivity of our estimates by varying the bandwidth of the centered vote variable (from a sample within 10 percentage points of the passage threshold to a sample within 20 percentage points) and modeling the vote in favor of passage using both linear and quadratic specifications. Bandwidth selection is based on Calonico et al.'s (2014) mean-squared-error (MSE) optimal bandwidth procedure. For most of our models, their procedure generates a bandwidth of approximately 10 percentage points from the passage threshold. But it yields a bandwidth that approaches 20 percentage points for certain analyses. To facilitate exposition, we select a bandwidth capturing the entire range (either 10 or 20 percentage points) across all expenditure, staffing, achievement, and attainment models.[7] In addition to capturing the general range of optimal bandwidths, this has the added benefit of making the sample referenda (and districts) com-

[6]We include student characteristics only for the year prior to the election due to missing values of these variables in early years. Including a second lag does not change estimated effects.

[7]Using a bandwidth under 10 leads to volatile results.

parable as we link changes in expenditures and staffing inputs with those of achievement and attainment outcomes.

We also weight all regressions by student counts, such that the estimated effects capture average effects across students as opposed to across districts. In addition to providing impact estimates that are not weighted in favor of small districts, we do this because the finance and achievement data become very noisy as districts get smaller and because Reardon et al. (2018) recommend the use of weights to account for variability in the precision of district test score estimates. Because the inverse of the squared errors of test score estimates are highly correlated with district enrollments ($\rho > 0.90$), weighting achievement models by student counts makes the expenditure, staffing, achievement, and attainment estimates comparable while also enabling us to account for error in the test-score measure. Indeed, as we show below, estimated achievement effects are similar if we use precision weights instead of enrollment weights.[8]

Finally, it is important to note how we handle districts for which we observe multiple referenda. Because we stack elections, we include as separate observations data from the same district if it has multiple referenda in our sample. In particular, 50 percent of districts that pass a referendum pass a subsequent referendum within the next five years, and 55 percent of districts in which a referendum fails subsequently pass a referendum within the next five years. At the vote threshold for passage, districts in which a referendum failed are 50 percentage points more likely to pass a referendum within the next five years than districts in which a referendum passed. In our main analysis below, we do not attempt to account for such "non-compliance" among districts with referenda that initially fail. Instead, we report intent-to-treat (ITT) estimates and cluster errors at the district level to account for our inclusion of multiple elections for some districts.

We focus on estimating ITT effects because generating treatment-on-the-treated

[8] As we show in Table D5 of appendix D, the impact of tax passage on graduation rates is the same whether or not we use weights, but finance and achievement effects are more pronounced with the inclusion of weights. These findings are consistent with the notion that there is significant noise in finance and achievement measures among smaller districts.

(TOT) estimates—by scaling both the spending and achievement estimates by the difference in compliance rates between districts with passing and failing levies—should lead to the same implied returns to spending as directly comparing ITT spending and achievement estimates (i.e., the denominators would cancel out). Indeed, the implied returns are nearly identical (but imprecisely estimated) if we use initial levy passage as an instrument for an endogenous predictor capturing the proportion of post-election years that follow levy passage, such that non-complying districts in which an initial levy fails but a subsequent levy passes are coded as partially "treated" (see Table D9 in Appendix D). Cellini et al. (2010) employ a TOT estimator that one can use in a dynamic framework. We do not employ a dynamic framework because we observe student achievement data for a limited number of years (2009-2015), which precludes us from calculating within-district changes in outcomes between a pre-election baseline and post-election years 5-7 for most elections. Indeed, the district elections on which we base our year 1 estimates are almost completely different from those on which we base our year 7 estimates.

3.1 Validity of RD Design

The parameter τ may be interpreted as the causal impact of passing a tax referendum for districts at the vote threshold necessary for passage if potential outcomes are continuous through that threshold. As is common in the literature, we examine the plausibility of this assumption by testing for pre-treatment imbalances in covariates, as well as testing for a discontinuity in the density of the running variable at the cutoff.

First, we tested whether there are any imbalances in the levels of observed district characteristics one year prior to the election, as well as whether there are differences in trends in the years leading up to the tax election (specifically, a difference between levels one year before the election and three years before). The results in Table B1 in Appendix B compare pre-election levels and trends using linear and quadratic specifications of the vote variables, as well as bandwidth restrictions of 10 and 20 percentage points in the centered vote share in favor of passage. The first three columns reveal four statistically significant results across 99 tests of differences in levels one year prior to the election: benefits per pupil are lower among passing referenda (-\$327, $p<0.1$) for the linear model restricted to a bandwidth of 10 percentage points; there are 0.66 fewer students per teacher in the linear model with a bandwidth of 10 ($p<0.05$); there are 0.6 fewer students per teacher in the model featuring a quadratic specification and a bandwidth of 20 ($p<0.1$); and there are 104 more students per instructional aide in the model featuring a quadratic specification and a bandwidth of 20 ($p<0.1$).

These imbalances appear minor, and a joint hypothesis test using a seemingly unrelated regressions (SUR) model confirms that these differences are not statistically significant when we use a linear specification of the running variable and a bandwidth of 10 percentage points ($p=0.168$), and if we use a quadratic specification and a bandwidth of 20 percentage points ($p=0.2114$). However, the model featuring a quadratic specification and a bandwidth of 10 percentage points—the model without a single significant result across 33 tests in Table B1—yields a significant result for the joint hypothesis test ($p=0.0221$). An analysis of trends yields similar results. Although in this case all three sets of models yield insignificant results for the joint hypothesis test ($p=0.69$, $p=0.54$, and $p=0.91$), the model with the quadratic specification and a bandwidth of 10 percentage points once again shows the most imbalance and reveals potentially problematic pre-election trends in student composition: an increase in the fraction of FRL students and a decline in the fraction of students designated as "limited English proficient" (LEP).

The imbalances we find when using a bandwidth of 10 percentage points and a quadratic polynomial of the vote are consistent with our finding that models using a linear specification of the vote share yield unstable estimates if we use a bandwidth below 10, which is generally the bottom end of the optimal bandwidth using Calonico et al.'s (2014) procedure. It appears that, within this narrow bandwidth, the quadratic polynomial leads to some over-fitting. Thus, in the analysis below, we focus on the two sets of models that yield covariate balance: those with a linear specification of the vote variable and a bandwidth of 10, and a model with a quadratic specification of the vote variable and a bandwidth of 20. As we note above, this captures the range of optimal bandwidths using Calonico et al.'s (2014) procedure, and the results from these models also generally characterize the range of results from using bandwidths within this range.

We also tested for bunching at the passage threshold via a McCrary (2009) test of the density of the running variable (the vote share in favor of tax increases). As the figures in Appendix C reveal, we are unable to reject the null of a continuous density for the full sample (p=0.56). The density plots by state reveal a statistically significant discontinuity for Michigan referenda (p=0.0431), but it appears that there are more districts that just *fail* to pass a tax increase than districts that just succeed in passing a tax.[9] Although there is some suggestive evidence of bunching to the right of the threshold for some states (e.g., Texas and Wisconsin), those differences do not approach conventional levels of statistical significance (p>0.25).

Overall, there are minor differences in pre-election levels and trends for the vote-share specification and bandwidth combinations we feature in the main body of the paper (although the minor differences in baseline student-teacher ratios are worth keeping in mind for the analysis below). Because the covariate balance tests generally yield results consistent with the notion that the pre-election covariates we include in the regressions are continuous across the vote threshold needed for passage, including these covariates should introduce no assumptions about the functional form of the underlying regression function (Calonico et al., 2018). However, if the few imbalances we detect are indeed evidence of a violation of RD assumptions, a causal interpretation of our estimates depends on the assumption that referendum passage is independent conditional on the baseline covariates we include.

[9]Hong and Zimmer (2016) find a greater density on the right side of the threshold for Michigan bond referenda. Indeed, if we combine bond and tax referenda, the density of the running variable appears almost perfectly smooth across the threshold.

4 Results

First, we take an initial look at the impact of tax levies on overall revenues, expenditures, achievement, and attainment for each post-election year and using a sample that includes all referenda held from 2000 through 2015. Second, we focus on average affects through five and seven post-election years using a wider range of expenditure and staffing variables and the sample of referenda for which we observe achievement five or seven years after the election. Third, we test for effect heterogeneity according to district poverty levels, spending levels, the timing of the election, and whether districts are in states with strong or weak collective-bargaining laws. Throughout, we test the sensitivity of the results to model specification and the bandwidth for selecting the sample.

4.1 Analysis by year relative to the election

Figure 1 displays changes in logged revenues per pupil for referenda within 20 percentage points of the passage threshold and for which we observe achievement after the election. The results are from a model with a quadratic specification of the vote and no additional covariates, and the dependent variable is the difference in logged revenues between various pre- and post-election years as compared to the year prior to the election. The figure reveals an average increase in revenues of around 10 percent near the threshold (regardless of the tax referendum's outcome) from three years prior to the election to nine years after the election. However, the figure also reveals that at the passage threshold, revenue growth in districts with a passing referendum (those to the right of the center line) is about 1 percent greater in the year of the election and 4-5 percent greater three years after the election. This revenue advantage disappears completely nine years after the election. As we discuss above, this is primarily because districts in which referenda initially failed were likely to pass a subsequent referendum within the next few years. Thus, their spending eventually catches up. The results are similar using a linear specification and a bandwidth of 10 percentage points around the passage threshold.[10]

[Insert Figure 1 about here.]

Table 4 presents yearly estimates of the impact of tax levy passage on operational expenditures per pupil, student achievement, and graduation rates. These estimates are based on the model in equation 2 that includes baseline covariates and enrollment weights, which are not included in the models reported in Figure 1. For now, we focus on operational (or "current") expenditures, which do not include capital outlays. They include gross salaries and benefits for all district employees—including temporary employees (e.g., substitute teachers)—as well as supplies and purchased services. As Table 4 reveals, increases in these expenditures track increases in revenues, and the results are the same whether we use a linear specification of the vote share variable and a bandwidth of 10 percentage points, or a quadratic specification and a bandwidth of 20 percentage points. Districts in which tax levies passed spent around $600 more per pupil by the fourth year after the election, and this spending differential declines from that point on.

[Insert Table 4 about here.]

Table 4 also reveals that achievement effects follow these spending increases. Although these effects do not become statistically significant until the fifth post-election year in the model featuring a quadratic specification, the magnitudes of estimated effects are comparable between the two samples and model specifications. Five years after the election, we generally find achievement effects of over 0.1 of a standard deviation—even if we use test-score precision weights instead of enrollment weights—and these effects climb to almost 0.2 standard deviations by the seventh year. We find similar effects for graduation rates, which are over 3 percentage points higher by the sixth post-election year.[11]

[10]Unlike the results we feature in the remainder of the paper, these estimates are from models that do not include weights for student counts. As we show in the appendix using a panel model that includes election fixed effects—as in Cellini et al. (2010)—as well as weights for student counts, in some years changes in revenues are as much as 10 percent higher in districts with passing referenda (see Table D2 in appendix D). Additionally, as Table D1 in Appendix D indicates, it appears that there is an even split between local and state revenue.

[11]We seldom observe pre-election achievement estimates and graduate rates from which to calculate changes from the year prior to an election. Thus, producing reduced-form RD plots for these measures yields a lot of

The estimates for later years should be viewed with caution. As we show in Table 5, passing a tax levy may be associated with changes in student composition. There seems to be a steady decline in the fraction of students who qualify for free or reduced-price lunches, and this decline is statistically significant in year 7. There is also an immediate and persistent jump in the fraction of students who are designated as "limited English proficient" (LEP) and modest evidence of a decline in students with "individual education plans" (IEPs). Such results could be due to changes in programming instead of actual changes in student composition. For example, districts with greater resources may be more inclined to identify LEP students, as programs related to these populations can be costly to administer (e.g., they may require more staffing). On the other hand, these findings might indeed reflect changes in student composition—changes which could increase average achievement without districts having improved their ability to educate students. This is particularly true of changes in the fraction of FRL students, as this measure is often considered a proxy for child poverty.

[Insert Table 5 about here.]

We consider the extent to which these potential changes in student composition account for our achievement results by including post-election fractions of FRL, Black, and Hispanic students as controls in our models.[12] As we show in Table 4, including these controls attenuates the estimated achievement effects, which are now 0.08-0.09 standard deviations in years 5 and 6, and no longer reach conventional statistical significance in year 7. On the other hand, controlling for student demographics increases the point estimates and statistical significance for models of graduation rates. Overall, like Lafortune et al. (2018)—who also find that increased revenues led to a decline in FRL students in some specifications—we conclude that a decline in student poverty is unlikely to explain our achievement effects. On average, increasing spending by around $500 per pupil implies achievement effects of around 0.1 of a standard deviation and an increase in graduation rates of 3-4 percentage points. The

noise. Nevertheless, we present these plots in Figure E1 of Appendix E.

[12]Estimates are similar if we control for IEP and LEP students, but our sample size becomes small because of the significant number of missing values for these measures.

following analyses provide evidence that strengthens this conclusion.

4.2 Average annual inputs through 5 and 7 post-election years

We now turn to examining changes in inputs that might explain these achievement and attainment effects. Because the effects of increased spending seem to accumulate to some extent, we present the results of models that estimate effects on average annual expenditures and staffing *through* five or seven years after the election (see equation 1 above), whereas we present estimates of the achievement and attainment effects for the latest year (5 or 7) of the relevant range. To maximize precision, in all subsequent analyses we include lagged values of the dependent variables one and three years prior to the election. (As we discuss above, the "lag" for achievement and attainment is actually our proxy for dropout rates: 1 minus 12th grade enrollments divided by 8th grade enrollments four years prior.) The only estimates that are substantively affected by the inclusion of lags are those related to student-teacher ratios (see Table D7 in appendix D). It appears that we find a significant effect of levy passage on student-teacher ratios in models without lags because of the pre-treatment imbalance we discuss above (see Table B1 in appendix B).

[Insert Table 6 about here.]

Table 6 reveals that the average increase in annual expenditures is around $400-$500 per pupil through 5-7 years after the election, and that this additional spending is more or less evenly divided between instruction and support services. On the other hand, there are no significant changes in capital outlays. The results also indicate that these expenditures went almost entirely toward higher salaries. Although there is suggestive evidence that districts added support staff, there are no significant impacts on student-teacher ratios. As Table 7 illustrates, whereas there are increases in staff FTEs to account for increases in overall salary-related spending, there are no changes in teacher FTEs. Salaries per instructor increase by around $4,000 by the fifth post-election year, which suggests that teachers may have gotten raises. Unfortunately, the data do not enable us to discern whether existing teachers received raises or districts hired new teachers at higher wages (e.g., replacing substitutes with permanent instructors).[13]

[Insert Table 7 about here.]

Finally, Table 6 reveals that achievement estimates are more precise and that attainment estimates are larger with the inclusion of baseline values of our proxy for district dropout rates. As in Table 4, the table also reveals that attainment estimates take longer to reach statistical significance than achievement estimates, although this may be due to smaller sample sizes for this measure. Table 8 tests the sensitivity of these estimates to the inclusion of post-election student demographic variables. Both achievement and attainment effects decline in magnitude but are more precisely estimated. The input variables remain qualitatively similar. Because the more conservative estimates in Table 8 should account for potential changes in student composition, we consider these our preferred estimates of spending's achievement and attainment effects.

[Insert Table 8 about here.]

4.3 Effect heterogeneity

We examine the heterogeneity of these effects along several important dimensions. First, we are interested in whether spending increases are more beneficial in districts serving disadvantaged student populations. Recent estimates of spending's impact are based on finance reforms focused on low-income districts, so it is unclear whether there would be such returns in wealthier districts. There are reasons to believe that this impact would not extend to wealthier districts, as higher income students may be less dependent on formal schooling to build academic knowledge and skills. Second, we are interested in the related question of whether there are diminishing returns to school district spending. In particular, we are interested in whether the effects are larger in low-spending districts, and whether the re-

[13] Salaries include gross payments to employees, including temporary staff (e.g., substitute teachers) and payments for additional duties (e.g., coaching). We found no impact on benefits. Support services include student support services, instructional staff support, general administration, school administration, operations and maintenance, student transportation, and other support services (see Cornman et al., 2018).

turns to spending decline during our period of study (2000-2015). Finally, one of the most salient school finance debates concerns whether collective bargaining prevents the allocation of resources toward their most productive uses. For example, Hoxby (1996), Chubb and Moe (1990), and others argue that teachers unions are able to capture revenues and direct them toward higher salaries and better working conditions for their membership, at the expense of student educational outcomes. Indeed, Brunner et al. (2019) find that districts in strong collective-bargaining states responded to increased revenues in very different ways than districts in states with weak or no collective bargaining.

To test for these effects, we compare spending and educational outcomes between districts that are above or below our sample median in terms of poverty rates among 5-17 year-olds (according to the American Community Survey) and above or below the median in per-pupil spending levels in the year prior to the election.[14] We also compare estimates among referenda held prior to 2008 and those held in 2008 or later. Essentially, this involves comparing referenda held in the wake of the Great Recession (2008-2010) and those held prior to it, as our analysis is limited to districts for which we observe achievement in the fifth post-election year for calendar years 2009 and 2015. Finally, following Brunner et al. (2019), we compare budgeting and outcomes between districts in states with strong collective bargaining laws (Michigan, Pennsylvania, and Wisconsin) and states with weak or no collective bargaining laws (Arkansas, Louisiana, Missouri, and Texas).

Our limited sample size is an issue. For example, removing referenda from any one state renders insignificant at least one of our spending, achievement, or attainment results. The effect is particularly pronounced if we remove Michigan, Texas, or Wisconsin, which contribute the greatest number of referenda to the sample (see Table D8 in appendix D). Nevertheless, we estimate variants of the following model, which interacts the passage indicator and polynomial of the vote in equation 2 with an indicator demarcating the categories

[14] We obtain similar results if we use the fraction of FRL students as a proxy for student poverty, as opposed to estimates of childhood poverty from the American Community Survey.

we wish to compare:

$$Y_i = \tau^P Pass_i + \tau^{AP} Above_i * Pass_i + Above_i * f(Pass_i, Vote_i) + f(Pass_i, Vote_i) + \lambda Above_i + \mathbf{X}_i' \gamma + \epsilon_i \quad (3)$$

$Above_i$ indicates whether or not a referendum was held in a district that is above the sample median in terms of spending or in a district above the median in terms of poverty. We interact this indicator with the indicator for referendum passage ($Pass$), as well as the polynomial for the percentage of the vote in favor of passage.[15] Thus, τ^P captures the effect of referendum passage for districts below the median in spending or poverty, and τ^{AP} is the difference in the estimated effect for districts above the median. We estimate similar regressions using an indicator of whether (1) or not (0) a district is in a state with strong collective bargaining (CB) and whether the tax election was held late in our sample (*late*)—that is, between 2008-2010 (1) or earlier (0). Once again, we report standard errors clustered by district.

[Insert Table 9 about here.]

The results in Table 9 reveal no statistically significant differences in spending or academic outcomes between districts with a relatively high proportion of students in poverty (above 18.2 percent) and those with a lower proportion of students in poverty. The lack of significant differences is perhaps partly due to our lack of statistical power, but the results clearly indicate that relatively high-income districts realized achievement gains that correspond to their heightened spending on salaries. Thus, we have good evidence that achievement gains are not limited to relatively low-income districts. On the other hand, although Table 9 reveals similar spending effects in both high- and low-spending districts (above or below $10,892 per pupil in 2012 dollars), achievement effects occur only in districts that were below the median in per pupil spending prior to the election—a difference that is statistically

[15] In our presentation of results, we focus on models that employ a linear specification of the vote variable. Although the results are qualitatively similar using a quadratic specification, they are also quite volatile.

significant (p<0.1).[16] Thus, we have evidence suggesting that there are diminishing returns to school district spending, regardless of district poverty levels. These results are robust to variations in bandwidths and model specifications.

[Insert Table 10 about here.]

Table 10 reveals no statistically significant differences in estimated effects based on the strength of a state's collective-bargaining laws. Again, as we note above, this is likely due in part to our lack of statistical power. But these results also may be due to our sample. The districts in our sample sought out additional revenues and their voters approved (or nearly approved) higher taxes, which suggests they may have had clear operational needs. That our achievement effects are driven by low-spending districts also supports this theory.

Finally, Table 10 reveals that our spending results are driven by referenda held prior to the Great Recession.[17] Further analysis reveals that elections held in the wake of the Great Recession had no significant impact on inflation-adjusted revenues, which corresponds to null achievement effects during that period (though the difference in achievement effects is not statistically significant). The results are consistent with districts requesting less money during this period. Further analysis suggests that this explanation is plausible. If we broaden the sample to include referenda between 2010 and 2015 (for which we do not observe achievement five years later) the revenue effects of levy passage are just as they were prior to the Recession.

5 Implied Returns to District Operational Spending

The most straightforward approach to calculating the returns to spending is to assume that operational spending's impact is immediate and not cumulative, such that the spending effects in a given year map only to the achievement and attainment effects in that same year. However, achievement and attainment are the product of years of cognitive and behavioral

[16]Further analysis reveals that revenues increased by about $600 per pupil in both sets of districts.

[17]As we show in Table D6 of appendix D, an analysis that focuses on the first referendum that each district passes yields results similar to those we find for early elections. Specifically, referendum passage is associated with an increase of around $700 in per pupil spending, increases in test scores of around 0.15 of a standard deviation, and increases in graduation rates of over 4 percentage points.

development. Interventions in early grades can have an impact many years later. For example, in order to get a diploma, one must take time to complete the necessary high school coursework. Whether one completes this coursework, in turn, depends on cognitive and behavioral outcomes in earlier grades. Similarly, improvements in a child's understanding of mathematics taught in grade 3 can facilitate improvements in grade 8 mathematics, so grade 8 test scores may not fully reflect improvements in a district's education production until those 3rd grade students reach the 8th grade. Thus, because achievement and attainment are the result of cumulative processes, there will likely be some delay before graduation rates and test scores in grades 3-8 reflect the full impact of changes in district practices.

Achievement and attainment effects should also take time to set in because the impact of spending on district practices is seldom instantaneous. For example, in their analysis of school finance reforms, Lafortune et al. (2018) found immediate spending increases that persisted, but achievement gains occurred gradually. This delay may have been due in part to districts directing a large portion of new revenue toward capital expenditures. Some capital expenditures could have an almost immediate impact on student learning (e.g., purchasing air conditioners), but construction projects take time to complete. Operational expenditures, on the other hand, could conceivably lead to immediate impacts because adding new staff—or simply retaining current, experienced staff—can have a significant impact on student learning. Indeed, studies focused on budget cuts related to the Great Recession—during which time staff retention had a pronounced effect—are notable for the immediate spending impacts they detect (see Jackson et al., 2018; Kogan et al, 2017; Lavertu and St. Clair, 2018). Nevertheless, we should still expect delays related to capacity-building (e.g., the multiple years it takes new teachers to hit their stride) in addition to delays related to the cumulative and long-term nature of cognitive and behavioral development.

With these concerns in mind, we sought to calculate the returns to spending by averaging annual spending effects across post-election years and linking this multi-year average to test scores and graduation rates in a terminal year that is sufficiently far into the future that cumulative learning and district capacity-building processes will have played out. Fortunately, our results indicate that both spending and achievement effects have plateaued by post-election years 5-7, which we see as evidence that these can serve as appropriate terminal years.[18] Consider our results in Table 8, which we prefer because the models account for possible post-election student sorting. The results indicate average spending effects of \$417-\$463 per pupil through the fifth post-election year and \$415-\$444 per pupil through the seventh post-election year. Similarly, achievement effects have stabilized to 0.056-0.075 of a standard deviation in those years. These results imply that each additional \$1,000 per pupil in operational spending leads to 0.13-0.16 of a standard deviation in student achievement.[19]

It is difficult to benchmark these estimates against those in other quasi-experimental studies of education spending because of the timing issues we describe above. However, it appears that Lafortune et al. (2018) find a relatively steady relationship between spending and achievement 10 years after districts experienced revenue increases. According to their preferred specification, districts had averaged \$424 more in education spending per pupil by that tenth year, which translates to an increase of 0.1 of a standard deviation in test scores. That implies that increasing spending by \$1,000 per pupil leads to an increase in achievement of 0.24 of a standard deviation—almost 0.1 of a standard deviation higher than our most conservative estimates and just below our largest estimates. Using Chetty et al.'s (2011) finding that a 0.1 standard deviation increase in kindergarten test scores translates to increased earnings in adulthood with a present value of \$5,350 per pupil, Lafortune and

[18]Indeed, an analysis of achievement by grade for those post-election years does not indicate differences in achievement effects for students who experienced higher spending for more years (e.g., there is no statistically significant difference between test score effects for grades 7-8 and those for grades 3-4). Although the estimates are imprecise and this evidence is merely suggestive, these results are consistent with the notion that achievement effects are not accumulating in years 5-7.

[19]We also estimated Two Stage Least Squares (2SLS) models that account for the fact that some districts in which referenda initially failed passed a subsequent referendum (see Table D9 in Appendix D). As one would expect, these treatment-on-the-treated (TOT) estimates are similar: the implied returns of \$1,000 in spending per pupil are approximately 0.12-0.16 of a standard deviation in achievement. We also estimated 2SLS models to directly estimate the impact of operational spending on achievement and graduation rates, using equation 2 as the first stage. The instruments are weak and the results are imprecise, but the estimates once again indicate that increasing spending by \$1,000 translates to achievement gains of 0.12-0.16 of a standard deviation (see Table D10 in Appendix D).

colleagues conclude that spending \$424 per year through grade 8 (with a present discounted cost of \$3,500 in kindergarten) yields a benefit-cost ratio of approximately 1.5. Using a similar calculation, our results linking \$450 in per pupil spending to gains in achievement of 0.07 of a standard deviation yield a benefit-cost ratio just above 1 for the average district, and a ratio just over 2 for districts below the median in spending.

Chetty et al.'s (2014) national study on the benefits of effective teachers yields estimates similar to Chetty et al. (2011). They find that teachers whose students show relative achievement gains of 0.13 of a standard deviation have 1.3-1.6 percent higher earnings in their late 20s. These earnings gains imply \$39,000 more in earnings over the course of their lives, which has a present value of \$7,000 when students are 12 years old. Thus, achievement gains of 0.1 standard deviations translate to a present value of approximately \$5,400 in seventh grade. Using data limited to Texas, Dobbie and Fryer (2019) find that students who attended effective charter schools also had higher earnings in their 20s. Specifically, charter schools that posted relative achievement gains of 0.1 of a standard deviation were associated with earnings gains of \$624.47 - \$1,510.04—approximately 2-5 times larger than the effects in Chetty et al. (2014). These studies are based on different assumptions about the timing of benefits, but they suggest achievement-returns to spending outweighed the costs for our sample of districts.

Research also indicates that schooling has complex cognitive and behavioral benefits that scores on state tests fail capture (Kraft, 2019). Graduation rates, however, may capture some of these benefits (Jackson, 2018b). Our results indicate that spending an extra \$1,000 per pupil for 7-8 years leads to an increase in graduation rates of around 8-10 percentage points. Dobbie and Fryer (2019) find that a 10 percentage-point increase in a charter school's graduation rate corresponds to earnings gains of \$1,629.81-\$2,647.80. Based on Chetty et al.'s (2014) calculations, that implies an increase in lifetime earnings with a present value of at least \$35,000 when students are 12 years old. Thus, our estimates suggest a benefit-cost ratio over 4.

These are approximate calculations. The above studies vary somewhat in terms of their inflation adjustments, years in which earnings are observed, and the details of their present-value calculations. They also assign no value to education benefits other than students' future incomes. Nevertheless, they suggest that our most conservative estimates of the achievement returns to spending (gains of approximately 0.15 of a standard deviation in achievement for every $1,000 in per-pupil spending) are about worth the cost. They also suggest that the benefits significantly exceed costs if one considers the returns to graduation or focuses on districts below the median in spending.

6 Conclusion

This study provides evidence that, even after the large increase in K-12 spending associated with state finance reforms in the late 20th Century, additional investments continue to have significant impacts on student achievement and attainment. Indeed, using a different design focused on more recent years, we obtain effect sizes that are very similar to Lafortune et al.'s (2018) and Candelaria and Shores's (2019) difference-in-differences analyses of state finance reforms. That we get results implying similar achievement and attainment returns to district spending puts additional weight behind their important, nationwide findings. Moreover, although our results are based on a relatively small number of districts, the districts in our analyses to a large extent mirror the demographic diversity of the United States and are typical in terms of spending levels and student achievement and attainment. To the extent that our results are indeed generalizable, they suggest that policymakers could realize significant gains in student achievement and attainment through increased district funding.

On the other hand, we find evidence of diminishing returns. Districts that spent above our sample median ($10,893 per pupil) had no achievement or attainment effects to accompany their increases in spending. That does not mean that high-spending districts did not put those extra funds to productive uses. For example, perhaps parents and voters in those districts were concerned with realizing outcomes that are orthogonal to math and ELA achievement. But these findings do suggest that if one's goal is to increase math and ELA achievement in grades 3-8—or to improve graduation rates—then funds are best spent in districts that spend less per pupil.

In addition to contributing to an emerging and rigorous evidence base on education spending and student outcomes, this study also contributes to the debate over the mechanisms that explain student outcomes. Unlike several prominent studies, the effects we detect are based purely on operational expenditures—which districts directed toward employee salaries, as opposed to smaller student-teacher ratios. We are unable to determine whether districts raised salaries of existing personnel or replaced them with higher-paid individuals. It is possible that districts realized achievement gains because they were able to employ higher-quality teachers or increase teacher effort (e.g., see Biasi, 2018). For example, it could be that low-spending districts that passed referenda suddenly were able to hire more effective full-time teachers, as opposed to relying on temporary options or less qualified individuals. Whatever the precise mechanism, we find suggestive evidence that higher teacher pay yields returns, which is consistent with an accumulating body of research (see Britton and Popper, 2016).

Unlike Brunner et al. (2019), we find little difference in spending choices between districts in states with strong and weak collective-bargaining laws. As we note in the introduction, the nature of our sample may be preventing us from finding differences in effects. Our effective analytic sample consists of districts seeking extra revenues (districts holding a referendum for a tax increase), as well as an electorate willing (or nearly willing) to provide those revenues through increased taxes. In other words, the districts in our sample may be unusual in that they had clear needs and the motivation to allocate resources to best meet those needs. Thus, it may be that the returns to district funding we find are greater than they would be for the average district.

References

Baron, E. Jason. (2019). "School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin," Working paper downloaded November 20, 2019 from `https://sites.google.com/view/jasonbaron/home`.

Biasi, Barbara. (2018). "The Labor Market for Teachers Under Different Pay Schemes," Working Paper, Dowloaded October 29, 2019 from `https://www.barbarabiasi.com/uploads/1/0/1/2/101280322/teacherlm.pdf`

Britton, Jack and Carol Propper. (2016). "Teacher pay and school productivity: Exploiting wage regulation," *Journal of Public Economics* 133: 75-89.

Brunner, Eric J., Joshua Hyman, and Andrew Ju. (2019). "School Finance Reforms, Teachers' Unions, and the Allocation of School Resources," *Review of Economics and Statistics*.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295-2326.

Calonico, S., Cattaneo, M. D., Farrell, M., and Titiunik, R. (2018). "Regression Discontinuity Designs Using Covariates," *Review of Economics and Statistics*.

Candelaria, Christopher A. and Kenneth A. Shores. (2019). "Court-Ordered Finance Reforms in The Adequacy Era: Heterogeneous Causal Effects and Sensitivity. Education Finance and Policy," *Education Finance and Policy* 14(1):31-60.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. (2010). "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *Quarterly Journal of Economics* 125(1):215–261.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach,and Danny Yagan. (2011). "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star," *Quarterly Journal of Economics* 126(4): 1593–1660.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. (2014). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review* 104(9): 2633–2679.

Chubb, John and Terry Moe. (1990). *Politics, Markets, and America's Schools* Washington, D.C.: Brookings Institution Press.

Cornman, Stephen Q., Lei Zhou, Malia R. Howell, and Jumaane Young. (2018). *Revenues and Expenditures for Public Elementary and Secondary Education: School Year*

2014–15 (NCES 2018-301) Washington, D.C.: U.S Department of Education.

Dobbie, Will and Roland G. Fryer. (2019). "Charter Schools and Labor Market Outcomes," *Journal of Labor Economics*

Gordon, Nora. (2004). " Do Federal Grants Boost School Spending? Evidence from Title I," *Journal of Public Economics* 88: 1771-1792.

Hanushek, Eric A.(2003). "The Failure of Input-based Schooling Policies," *The Economic Journal* 113(February):F64–F98.

Heckman, James J., and Paul A. LaFontaine. (2010). "The American High School Graduation Rate: Trends and Levels," *Review of Economics and Statistics* 92(2): 244–262.

Hong, Kai and Ron Zimmer. (2016). "Does Investing in School Capital Infrastructure Improve Student Achievement?" *Economics of Education Review* 53(8):143–158.

Hoxby, Carolyn Minter. (1996). "How Teachers' Unions Affect Education Production" *The Quarterly Journal of Economics* 111(3): 671–718.

Jackson, C. Kirabo. (2018a). "Does School Spending Matter? The New Literature on an Old Question," National Bureau of Economic Research, Cambridge, MA, Working Paper No. 25368.

Jackson, C. Kirabo. (2018b). "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes," *Journal of Political Economy* 126(5): 2072-2107.

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. (2016). "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms," *The Quarterly Journal of Economics*, 131(1):157–218.

Jackson, C. Kirabo, Cora Wigger, and Heyu Xiong. (2018). "Do School Spending Cuts Matter? Evidence from the Great Recession," National Bureau of Economic Research, Cambridge, MA, Working Paper No. 24203.

Kogan, Vladimir, Stéphane Lavertu, and Zachary Peskowitz. (2017). "Direct Democracy and Administrative Disruption," *Journal of Public Administration Research and Theory* 27(3): 381-399.

Kraft, Matthew. (2018). "Interpreting Effect Sizes of Education Interventions," Brown University Working Paper. Downloaded Tuesday, April 16, 2019, from `https://scholar.harvard.edu/files/mkraft/files/kraft_2018_interpreting_effect_sizes.pdf`.

Kraft, Matthew. (2019). "Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies," *Journal of Human Resources* 54(1):1-36.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. (2018). "School Finance Reform and the Distribution of Student Achievement," *American Economic Journal: Applied Economics* 10(2):1–26.

Lavertu, Stéphane and Travis St. Clair. (2018). "Beyond Spending Levels: Revenue Uncertainty and the Performance of Local Governments," *Journal of Urban Economics* 106: 59-80.

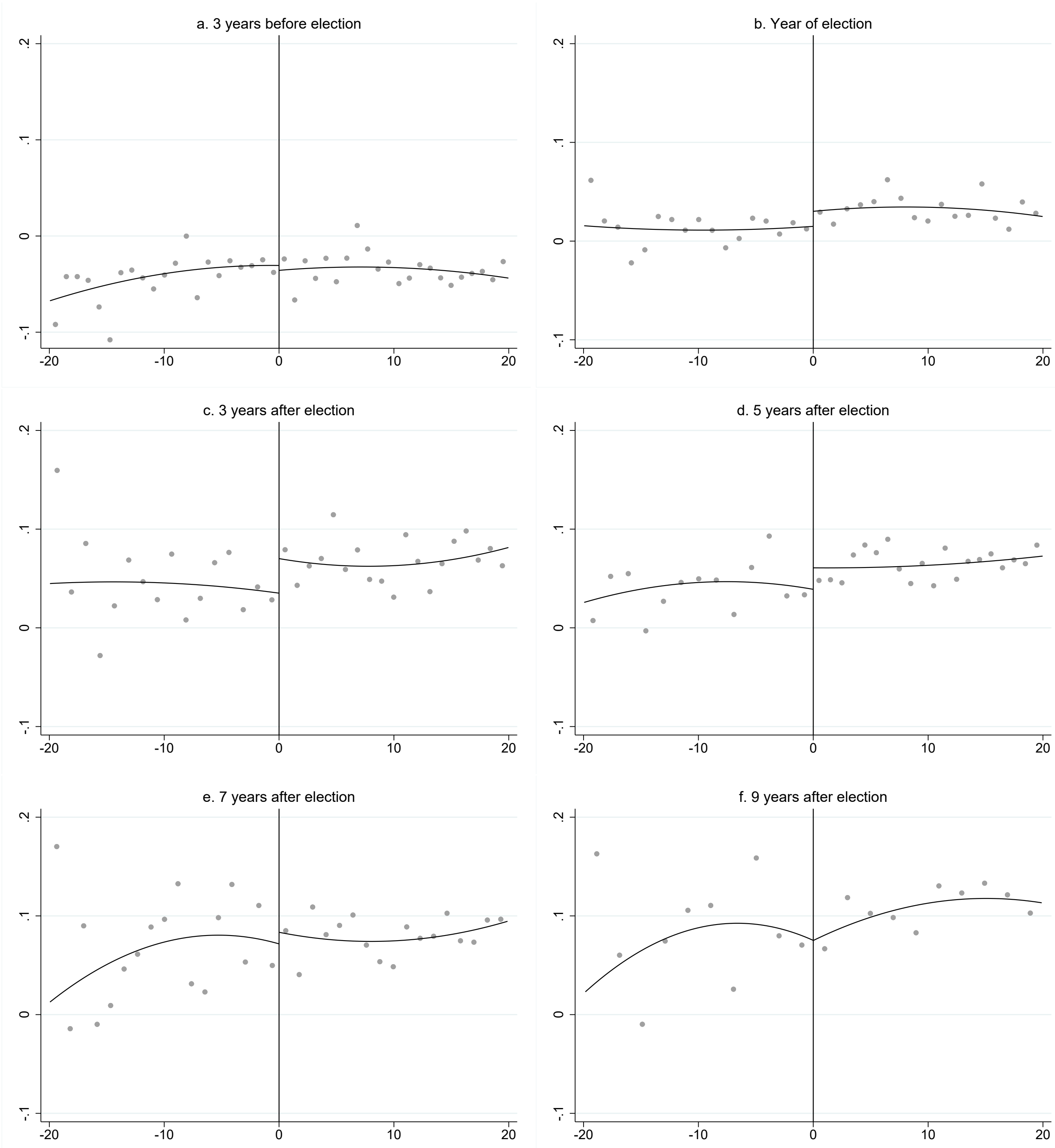
Martorell, Paco, Kevin Stange, and Isaac McFarlin. (2016). "Investing in Schools: Capital Spending, Facility Conditions, and Student Achievement," *Journal of Public Economics* 140: 13–29.

Oklahoma State Department of Education (OSDE). (2017). *Oklahoma School Finance Technical Assistant Document*. Financial Services Division. Downloaded April 2, 2019 from `https://sde.ok.gov/sites/ok.gov.sde/files/documents/files/`.

Rauscher, Emily. (2019). "Delayed Benefits: Effects of California School District Bond Elections on Achievement by Socioeconomic Status," *Sociology of Education*

Reardon, Sean F., Andrew D. Ho, Benjamin R. Shear, Erin M. Fahle, Demetra Kalogrides, and Richard DiSalvo. (2018). *Stanford Education Data Archive (Version 2.1)*. Retrieved from http://purl.stanford.edu/db586ns4974

Figure 1: Changes in logged revenues (2012 dollars)



Note: The figure compares changes in the natural log of revenues per pupil between districts in which tax referenda passed (vote margin > 0) and those in which referenda failed (vote margin < 0) as of a) 3 years before the election, b) the year of the election, c) 3 years after the election, d) 5 years after the election, e) 7 years after the election, and f) 9 years after the election. Specifically, the y axes capture the difference in logged revenues (2012 dollars) between each of these years and the year before the election. The x axes capture the vote margin—the difference between the percent of votes in favor of passage and the threshold needed for passage (typically 50 percent). The dots are local means and the curves are fitted using a quadratic polynomial. The sample is limited to levies within 20 percentage points of passage and those for which we observe post-election achievement in at least one year.

TABLE 1. Descriptive statistics for tax referenda – full vs. effective samples

	Referendum Count	Percent Passed	Mean Pct Yes Vote	Mean Vote Count	Election Years
Overall	3,070	75.08	61.04	3,205	2000 - 2015
	[2,025]	[67.11]	[54.41]	[3,382]	[2000-2015]
By State					
Arkansas	525	85.90	65.89	592	2002 - 2015
	[305]	[78.69]	[57.68]	[695]	[2002-2015]
Louisiana	510	87.06	66.74	6,846	2000 - 2015
	[256]	[78.52]	[57.37]	[5,525]	[2000-2015]
Michigan	659	87.41	66.33	2,448	2002 - 2015
	[375]	[79.47]	[58.49]	[3,193]	[2002-2015]
Missouri	96	63.54	53.01	6,563	2001 - 2015
	[82]	[62.20]	[51.35]	[7,047]	[2002-2015]
Pennsylvania	105	70.48	56.05	3,144	2001 - 2015
	[71]	[71.83]	[55.84]	[2,864]	[2001-2015]
Texas	474	73.63	62.62	1,434	2006 - 2015
	[293]	[63.48]	[53.70]	[2,011]	[2006-2015]
Wisconsin	701	49.93	49.08	3,952	2000 - 2015
	[643]	[51.63]	[49.86]	[4,118]	[2000-2015]

Note: The table provides descriptive statistics for tax referenda. Descriptive statistics in brackets are for the effective sample within 20 percentage points of the vote threshold for passage, as all analyses are based on effective samples within this range.

TABLE 2. Characteristics of districts near threshold and w/ post-election achievement obs.

	District Count	Total Expnd. Per Pupil (2012$)	Oper. Expnd. Per Pupil (2012$)	Pct. Rural	Stdnt Count	Pct. White Stdnts	Pct. F/R Lunch Stdnts	Teach. FTE	SEDA Test Scores (SDs)	Grad. Rate (4yr)
All Refs.	566	11,185	9,700	51.01	3,820	75.95	41.57	258	-0.03	85.61
By Result										
Passed	429	10,938	9,555	50.82	4,118	74.89	43.18	277	-0.05	84.67
Failed	235	11,647	9,972	51.38	3,263	77.93	38.56	223	0.002	87.41
By State										
Arkansas	60	9,686	8,393	42.95	2,644	73.47	57.61	183	-0.18	81.49
Louisiana	36	9,706	8,934	61.80	10,098	58.81	64.52	696	-0.27	72.21
Michigan	98	10,874	9,511	47.13	2,828	79.53	39.03	165	-0.11	79.03
Missouri	37	10,865	8,714	37.50	4,293	85.02	32.80	285	0.15	89.57
Penn.	15	12,431	10,078	40.00	2,775	89.27	24.56	177	0.15	91.31
Texas	171	11,109	9,193	57.38	4,116	54.99	48.94	288	-0.05	92.60
Wisconsin	149	12,673	11,229	53.02	2,875	91.84	24.86	198	0.15	92.17

Note: The table provides statistics on the districts in which referenda were held and for which we observe achievement at least once 0-5 years after the election. Nearly all statistics are based on averages across referenda in the year prior to an election. The test score and graduation rate statistics, however, are are averages from 0-5 years after the election. Districts that placed multiple referenda on the ballot figure more prominently in the averages. The mean standardized test scores in the final column are based on 2009-2015 data from the Stanford Education Data Archive (SEDA). The four-year, cohort-based graduation rates are from state education department websites. All other data are from the NCES Common Core of Data Finance and Universe files. Expenditures are in 2012 dollars.

TABLE 3. Descriptive stats through 5 post-election yrs. (elections near pass. thresh.)

	N	Mean	SD	Min	Max
Expenditures Per Pupil (2012$)					
Operations	946	10,061	1,588	6,946	19,636
Instruction	946	6,025	926	3,716	11,331
Support services	946	3,555	770	1,911	7,615
Other operations	946	480	110	0	1,009
Capital outlays	946	941	889	24	6,882
Salaries	946	5,816	759	4,085	9,049
Instructor Salaries	946	4,013	543	2,562	6,016
Benefits	946	2,207	929	659	4,811
Instructor Benefits	946	1,474	632	457	3,244
Salaries per FTE (2012$)					
Salaries per staff	914	43,003	6,864	27,676	67,419
Instructor salaries per teacher	913	57,109	8,560	35,603	81,201
Staffing					
Student-teacher ratio	914	14.51	2.61	9.22	51.51
Student-aide ratio	905	102.38	144.71	22.67	2,369
Student-administrator ratio	916	76.62	19.68	29.37	252.71
Student-counselor ratio	858	475	264	157	3,071
Student-librarian ratio	745	1,039	2,890	251	74,022
Achievement					
SEDA math/ELA scores (SDs)	946	-0.03	0.29	-1.19	0.96
Attainment					
Graduation rate (4-year, cohort)	849	86.66	11.39	34.38	100

Note: These are descriptive statistics for select outcome variables, which capture averages between 0-5 post-election years for districts that came within 20 percentage points of passing or failing a referendum. The mean standardized test scores in the final column are based on 2009-2015 data from the Stanford Education Data Archive (SEDA). The four-year, cohort-based graduation rates are from state education department websites. All other data are from the NCES Common Core of Data Finance and Universe files. Expenditures in 2012 dollars.

TABLE 4. Impact of tax elections on spending, achievement, and graduation rates

	Election Year	Post-El. Year 1	Post-El. Year 2	Post-El. Year 3	Post-El. Year 4	Post-El. Year 5	Post-El. Year 6	Post-El. Year 7	Poly./ Band.
Op. Expnd.									
	234***	383***	410**	520***	614***	611**	440**	167	linear
	(79.4)	(98.4)	(168)	(169)	(216)	(247)	(212)	(183)	/ 10
	277***	358***	360**	495***	598***	667***	590**	363*	quad.
	(89.2)	(103)	(162)	(184)	(218)	(241)	(254)	(206)	/ 20
Achievement									
	0.052	0.13*	0.15**	0.12*	0.14**	0.17***	0.18***	0.19***	linear
	(0.047)	(0.067)	(0.071)	(0.068)	(0.069)	(0.063)	(0.057)	(0.068)	/ 10
	0.036	0.092	0.13	0.082	0.099	0.14**	0.17***	0.16**	quad.
	(0.057)	(0.077)	(0.077)	(0.081)	(0.080)	(0.063)	(0.059)	(0.072)	/ 20
w/ demog.	0.049	0.069	0.081	0.054	0.094	0.082*	0.09***	0.054	quad.
controls	(0.052)	(0.064)	(0.067)	(0.080)	(0.090)	(0.043)	(0.030)	(0.038)	/ 20
w/ precis.	0.020	0.059	0.072	0.043	0.056	0.11**	0.13***	0.098**	quad.
weights	(0.045)	(0.054)	(0.057)	(0.056)	(0.059)	(0.050)	(0.047)	(0.049)	/ 20
Grad. Rate									
	3.64	1.79	2.50	0.57	0.74	1.30	3.42*	3.21	linear
	(3.05)	(2.75)	(2.75)	(1.93)	(1.93)	(1.61)	(1.80)	(2.00)	/ 10
	2.40	0.55	1.01	0.16	0.21	1.21	3.24*	2.59	quad.
	(3.31)	(2.89)	(2.89)	(2.00)	(2.05)	(1.77)	(1.66)	(1.78)	/ 20
w/ demog.	1.13	1.95	1.57	1.15	2.29	2.02	3.93**	3.85**	quad.
controls	(2.19)	(2.99)	(3.12)	(2.72)	(3.13)	(2.15)	(1.84)	(1.63)	/ 20

Note: Each coefficient is from a separate OLS regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed, from the year of the election through the seventh post-election year. The dependent variables are operational ("current") expenditures, SEDA's measure of standardized student achievement in math and reading, and 4-year cohort-based graduation rates. All models include school-year fixed effects and baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts unless "precision weights" are included, which capture the inverse of the squared standard error in district-level test score estimates. Specifications labeled "w/ demog. controls" include post-election student composition (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) as controls to account for possible changes in student composition. The final column indicates whether the model features a linear or quadratic specification and whether the sample is restricted to a bandwidth of 10 or 20 percentage points in the percent of votes in favor of a tax levy. Standard errors clustered at the district level appear in parentheses below coefficient estimates. *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 5. Impact of tax elections on district student characteristics

	Election Year	Post-El. Year 1	Post-El. Year 2	Post-El. Year 3	Post-El. Year 4	Post-El. Year 5	Post-El. Year 6	Post-El. Year 7	Poly./ Band.
F/R Lunch									
	-0.0028	-0.014	-0.021	-0.00040	-0.018	-0.025	-0.042	-0.056**	linear
	(0.011)	(0.021)	(0.017)	(0.015)	(0.019)	(0.026)	(0.026)	(0.025)	/ 10
	-0.0074	-0.013	-0.029	-0.0021	-0.018	-0.011	-0.037	-0.055**	quad.
	(0.013)	(0.025)	(0.023)	(0.019)	(0.022)	(0.030)	(0.029)	(0.027)	/ 20
Hispanic									
	-0.0031	0.00021	0.0023	0.0044	0.0028	0.0010	0.00029	-0.0053	linear
	(0.0025)	(0.0022)	(0.0030)	(0.0043)	(0.0054)	(0.0069)	(0.0086)	(0.011)	/ 10
	-0.0048	-0.00057	0.00085	0.0031	0.00012	-0.00041	-0.0047	-0.012	quad.
	(0.0032)	(0.0025)	(0.0035)	(0.0053)	(0.0059)	(0.0068)	(0.0089)	(0.011)	/ 20
Black									
	0.0019	0.0042*	0.0019	0.0022	0.0060	0.0031	0.0045	0.0094	linear
	(0.0015)	(0.0023)	(0.0029)	(0.0042)	(0.0047)	(0.0057)	(0.0063)	(0.0074)	/ 10
	0.00054	0.0032	-0.0016	-0.0021	0.00014	-0.0035	-0.0011	0.0077	quad.
	(0.0017)	(0.0026)	(0.0038)	(0.0051)	(0.0057)	(0.0068)	(0.0076)	(0.0084)	/ 20
IEP									
	0.00063	-0.0042	-0.014**	-0.011	-0.0057	-0.0037	-0.0070	-0.0065	linear
	(0.0039)	(0.0041)	(0.0068)	(0.0065)	(0.0050)	(0.0054)	(0.0056)	(0.0058)	/ 10
	0.0055	0.0017	-0.0095	-0.0050	-0.0001	0.0012	-0.0012	-0.0020	quad.
	(0.0046)	(0.0045)	(0.0066)	(0.0062)	(0.0050)	(0.0053)	(0.0054)	(0.0059)	/ 20
LEP									
	0.019*	0.025**	0.027***	0.023**	0.023**	0.019	0.023*	0.019	linear
	(0.011)	(0.010)	(0.0096)	(0.010)	(0.011)	(0.012)	(0.013)	(0.014)	/ 10
	0.022*	0.026**	0.028***	0.025**	0.028**	0.023*	0.026*	0.019	quad.
	(0.012)	(0.011)	(0.010)	(0.012)	(0.012)	(0.014)	(0.015)	(0.016)	/ 20
Enrollment									
	-1,276	-1,315	-2,101	-2,088	-1,483	-1,763	-1,468	-1,774	linear
	(2,543)	(2,563)	(2,226)	(2,341)	(2,343)	(2,471)	(2,551)	(2,697)	/ 10
	-1,320	-1,387	-1,851	-1,846	-1,060	-1,282	-948	-1,643	quad.
	(2,404)	(2,415)	(2,084)	(2,110)	(2,281)	(2,356)	(2,462)	(2,516)	/ 20

Note: Each coefficient is from a separate OLS regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed, from the year of the election through the seventh post-election year. The dependent variables capture the fraction of students who qualify for free- or reduced-price lunches, who are Hispanic, who are Black, who have an independent education plan (IEP), and who are designated "limited English proficient," as well as total district enrollments. All models include school-year fixed effects and baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. The final column indicates whether the model features a linear or quadratic specification and whether the sample is restricted to a bandwidth of 10 or 20 percentage points in the percent of votes in favor of a tax levy. Standard errors clustered at the district level appear in parentheses below coefficient estimates. $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$

TABLE 6. Inputs through 5 and 7 post-election years (observations with achievement data)

		$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Expend. P.P.	Operations	425**	444**	428**	473**
(2012 dollars)		(175)	(179)	(190)	(188)
	Instruction	211**	215**	164	170*
		(86.1)	(87.6)	(102)	(92.4)
	Support services	206**	236**	255**	286***
		(95.7)	(98.1)	(105)	(104)
	Capital outlays	23.1	39.4	89.5	135
		(177)	(175)	(190)	(206)
	Salaries	405***	426***	393***	457***
		(110)	(116)	(129)	(130)
	Instructor Salaries	272***	247***	258***	244***
		(69.9)	(71.5)	(82.8)	(82.3)
	Benefits	12.9	33.1	11.8	34.2
		(46.3)	(52.0)	(53.7)	(58.9)
	Instructor Benefits	-21.9	-3.15	-27.9	-14.6
		(29.9)	(32.9)	(41.0)	(43.3)
Staffing	Student-teacher ratio	0.038	-0.038	0.16	0.0079
		(0.31)	(0.28)	(0.41)	(0.38)
	Student-aide ratio	-5.92	-21.5	27.9	23.5
		(15.8)	(22.8)	(27.2)	(33.1)
	Student-counselor ratio	-26.7	-29.0	0.50	1.34
		(28.9)	(31.2)	(37.0)	(40.5)
Math/ELA (SDs)		0.15***	0.12**	0.16***	0.15**
		(0.058)	(0.058)	(0.058)	(0.059)
Grad. rate		2.87	2.32	7.29***	7.04***
		(2.18)	(2.30)	(2.35)	(1.94)
Bandwidth		+/- 10	+/- 20	+/- 10	+/- 20
Specification		linear	quadratic	linear	quadratic
Baseline Covariates		yes	yes	yes	yes
DV Lags		yes	yes	yes	yes
Demographic Controls		no	no	no	no
School Year FEs		yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across up to five or seven post-election years, and test scores and graduation rates in the highest year of the relevant range (year 5 or 7). All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 7. Impact of tax referenda on salaries per teacher and staff FTEs

	Election Year	Post-El. Year 1	Post-El. Year 2	Post-El. Year 3	Post-El. Year 4	Post-El. Year 5	Post-El. Year 6	Post-El. Year 7	Poly./ Band.
Salaries/	390	87.2	1,176*	1,439	2,331***	2,477**	3,456***	2,805***	linear
staff FTE	(474)	(608)	(676)	(947)	(783)	(981)	(1,006)	(748)	/ 10
	536	-5.37	1,058	1,176	1,899**	2,249**	3,383***	2,460***	quad.
	(483)	(572)	(798)	(1,054)	(820)	(1,050)	(1,057)	(856)	/20
Inst.Sal./	1,233***	1,696***	2,630***	3,048***	3,410***	3,967***	4,249***	3,471***	linear
Inst.FTE	(475)	(593)	(931)	(908)	(930)	(1,007)	(823)	(1,049)	/10
	1,279**	1,157*	2,022*	2,297**	2,747***	3,899***	4,925***	3,595***	quad.
	(535)	(616)	(1,096)	(966)	(966)	(1,155)	(983)	(1,263)	/20
Staff	45.1**	27.6	48.3*	60.7**	37.7	5.69	-14.2	12.9	linear
FTE	(17.5)	(22.4)	(26.1)	(30.3)	(29.1)	(40.0)	(31.9)	(38.3)	/10
	38.8*	32.8	46.1	57.6	9.13	-26.0	-35.0	-5.45	quad.
	(21.4)	(24.7)	(30.2)	(37.4)	(34.2)	(51.2)	(39.2)	(45.4)	/20
Teacher	-0.87	-2.99	9.28	0.11	-2.70	-23.3	-22.7	-1.34	linear
FTE	(5.88)	(9.04)	(11.3)	(11.4)	(17.2)	(20.3)	(17.9)	(19.7)	/10
	1.65	2.01	11.7	2.28	-12.7	-33.8	-26.0	-12.4	quad.
	(6.93)	(8.89)	(13.2)	(12.9)	(18.6)	(27.7)	(19.8)	(22.2)	/20

Note: Each coefficient is from a separate OLS regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed from the year of the election through the seventh post-election year. The dependent variables are salary expenditures divided by staff FTEs, instructor salary expenditures divided by teacher FTEs, staff FTEs, and teacher FTEs. All models include school-year fixed effects and baseline covariates capturing the dependent variable one and three years prior to the election, logged enrollments and operational expenditures one and three years prior to the election, and student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. The final column indicates whether the model features a linear or quadratic specification and whether the sample is restricted to a bandwidth of 10 or 20 percentage points in the percent of votes in favor of a tax levy. Standard errors clustered at the district level appear in parentheses below coefficient estimates. *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 8. Inputs through 5 and 7 post-election years (post-election student demog. controls)

		$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Expend. P.P.	Operations	417**	463**	444**	415**
(2012 Dollars)		(181)	(180)	(187)	(180)
	Instruction	197**	216**	159	114
		(84.8)	(88.3)	(102)	(94.1)
	Support services	203**	249**	260***	272***
		(100)	(97.3)	(100)	(98.4)
	Capital outlays	-92.8	-13.4	-81.7	-45.4
		(161)	(173)	(161)	(199)
	Salaries	369***	419***	362***	400***
		(117)	(117)	(125)	(125)
	Instructor Salaries	242***	237***	217***	173**
		(68.8)	(69.6)	(78.3)	(73.3)
	Benefits	31.1	50.6	17.6	10.7
		(45.5)	(47.0)	(59.2)	(58.2)
	Instructor Benefits	-11.3	6.98	-20.4	-26.5
		(29.9)	(31.1)	(44.5)	(45.1)
Staffing	Student-teacher ratio	0.32	-0.029	0.13	0.015
		(0.46)	(0.30)	(0.42)	(0.37)
	Student-aide ratio	-8.06	-23.0	-6.88	9.88
		(15.9)	(23.7)	(21.0)	(24.8)
	Student-counselor ratio	-7.82	-16.4	-4.78	-3.14
		(29.3)	(34.4)	(31.5)	(38.5)
Math/ELA (SDs)		0.066*	0.075*	0.067**	0.056*
		(0.038)	(0.043)	(0.030)	(0.032)
Grad. rate		1.74	1.69	3.99**	3.80**
		(1.80)	(2.09)	(1.99)	(1.61)
Bandwidth		+/- 10	+/- 20	+/- 10	+/- 20
Specification		linear	quadratic	linear	quadratic
Baseline Covariates		yes	yes	yes	yes
DV Lags		yes	yes	yes	yes
Demographic Controls		yes	yes	yes	yes
School Year FEs		yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across up to five or seven post-election years, and test scores and graduation rates in the highest year of the relevant range (year 5 or 7). All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions also include controls for post-election student demographics and are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 9. Outcomes through 5 post-election years – by child poverty and baseline spending

		pass	pass*(high pov)	pass	pass*(high spend)
Expend. P.P.	Operations	256	301	383	50.6
		(202)	(281)	(251)	(278)
	Instruction	73.9	251	271*	-133
		(94.6)	(154)	(147)	(177)
	Support	137	112	185	36.3
		(107)	(149)	(135)	(142)
	Capital	-121	38.4	35.7	-230
		(298)	(364)	(198)	(371)
	Salaries	351**	18.9	387***	-40.8
		(170)	(200)	(137)	(198)
	Instructor Sal.	211**	48.1	292***	-97.7
		(90.2)	(122)	(101)	(132)
	Benefits	2.48	63.2	-38.1	122
		(58.9)	(80.3)	(68.3)	(98.6)
	Instructor Ben.	-52.6	83.7	-31.6	34.1
		(46.3)	(60.2)	(45.7)	(63.9)
Staffing	Student-teacher	0.73	-0.91	-0.29	1.09
		(0.63)	(0.67)	(0.47)	(1.24)
	Student-aide	-39.9	61.5	8.05	-27.7
		(31.5)	(39.7)	(32.8)	(34.2)
	Student-counsel.	-8.84	-6.96	-19.6	22.9
		(40.5)	(55.7)	(41.0)	(63.8)
Math/ELA (SDs)		0.090**	-0.055	0.15***	-0.15*
		(0.045)	(0.068)	(0.057)	(0.089)
Grad. rate		1.76	-0.95	3.68	-3.76
		(2.14)	(3.23)	(3.27)	(5.16)
Bandwidth/Spec.		+/-10		+/-10	
Specification		linear		linear	
Baseline Covariates		yes		yes	
DV Lags		yes		yes	
Demog. Controls		yes		yes	
School Year FEs		yes		yes	

Note: Each pair of coefficients (a passage indicator and that passage indicator interacted with an indicator for high district poverty or high district spending) is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across five post-election years, and test scores and graduation rates in year 5. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election, as well as post-election student demographics. All regressions are weighted by baseline student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 10. Outcomes through up to 5 post-election years – by CB and election timing

		pass	pass*CB	pass	pass*(late elect.)
Expend. P.P.	Operations	267	251	708***	-722**
		(227)	(275)	(264)	(344)
	Instruction	213	-46.5	366***	24.3
		(130)	(165)	(124)	(167)
	Support	103	183	337**	-328*
		(125)	(150)	(147)	(188)
	Capital	-21.3	-175	-39.2	-153
		(193)	(419)	(230)	(306)
	Salaries	208*	170	542***	-433**
		(113)	(182)	(152)	(178)
	Instructor Sal.	220**	2.22	350***	-264**
		(87.1)	(127)	(94.8)	(120)
	Benefits	-34.3	117	80.5	-111
		(62.4)	(99.8)	(58.7)	(89.7)
	Instructor Ben.	-19.7	8.68	21.9	-81.7
		(43.3)	(65.0)	(40.2)	(61.4)
Staffing	Student-teacher	-0.41	1.26	0.29	-0.073
		(0.46)	(1.37)	(0.64)	(0.70)
	Student-aide	14.8	-37.7	-3.86	-16.0
		(18.0)	(25.9)	(16.5)	(34.1)
	Student-counsel	52.0	-91.4	7.00	-43.6
		(48.4)	(77.6)	(34.0)	(50.0)
Math/ELA (SDs)		0.069	-0.049	0.12***	-0.13
		(0.055)	(0.079)	(0.045)	(0.084)
Grad. rate		0.75	1.97	2.19	0.98
		(2.31)	(2.77)	(2.23)	(4.21)
Bandwidth		+/-10		+/-10	
Specification		linear		linear	
Baseline Covariates		yes		yes	
DV Lags		yes		yes	
Demog. Controls		yes		yes	
School Year FEs		yes		yes	

Note: Each pair of coefficients (a passage indicator and that passage indicator interacted with an indicator for strong collective bargaining or an election late in our sample) is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across five post-election years, and test scores and graduation rates in year 5. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election, as well as post-election student demographics. All regressions are weighted by baseline student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Tax Levy Information by State

Table A1. Local Tax Levy Information by State

Arkansas	Local revenues come primarily from property taxes. There are no local sales or income taxes. There are no limits on local property tax rates that school districts can levy, but there is a minimum of 25 mills for maintenance and operations. A majority of voters must approve increases to property tax rates beyond this minimum. The constitution also mandates that millage rates that are unchanged appear on the ballot every year, but the vote outcome does not affect a district's millage rate.
Louisiana	Local revenues come primarily from sales and property taxes (slightly more from sales taxes). Sales taxes are capped at 3 percent for all local entities combined and a majority of voters must approve them. Parish school boards also have the authority to levy a "constitutional" property tax of up to 5 mills (13 mills in New Orleans). Districts can supplement this by obtaining voter approval to levy additional property taxes for a specific purpose relating to operations, maintenance, or capital expenses. There is a maximum of 70 mills above the "constitutional" tax and it is limited in duration.
Michigan	Local revenues come primarily from property taxes. School districts must get approval from a majority of voters if they wish to exceed caps on local property taxes that the state set in 1994. In general, there is a cap of 18 mills on non-homestead property taxes. A majority of school district voters must approve millage increases for non-homestead properties and must renew these mills over time. Although law generally prohibits property taxes on homesteads (including noncommercial and agricultural property), districts for which the 18 mill cap does not allow them to get to pre-1994 levels may levy additional taxes to get up to that level. There were 52 (of 555) traditional school districts that met this "hold harmless" condition after the law's passage.
Missouri	Local revenues come primarily from property taxes. Local school districts can levy taxes of up to \$2.75 per \$100 of assessed valuable without voter approval and must do so to receive increases in state aid. Tax rates above \$2.75 require approval of 50 percent of voters and rates that exceed \$6.00 per \$100 of valuation must be approved by two thirds of voters. There is no limit on tax rates.
Pennsylvania	Local revenues come primarily from property taxes. A 2006 law requires voter approval for any proposed tax increase that exceeds an index capturing increases in wages and employment costs for schools. The 2006 law also sought to provide homeowners with property tax relief and gave districts the option of shifting some of the tax burden from local property taxes to local income taxes to offset losses. The vast majority of districts placed such a measure on the ballot in May of 2007, which voters overwhelmingly defeated throughout the state.
Texas	Local revenues come primarily from property taxes. Beginning in 2007, districts could put for a vote whether to impose taxes to maintain revenues equal to revenues prior to the 2006 tax reform. Districts adopt maintenance and operations tax rates each year. School districts must hold a tax referendum if school boards adopt a tax rate that exceed the "rollback rate." This rate effectively allows minimal increases in district revenues without voter approval.
Wisconsin	Local revenues come primarily from property taxes. District can raise property tax revenues up to a state-mandated revenue limit. Districts must obtain approval from a majority of district voters to exceed the state revenue limit. Voters can elect to exceed a revenue limit permanently.

B Covariate Balance Tests

TABLE B1. Pre-Election Balance Tests (various bandwidths and specifications)

	(1) Level t_1	(2) Level t_1	(3) Level t_1	(4) Trend (t_1-t_3)	(5) Trend (t_1-t_3)	(6) Trend (t_1-t_3)
totalrev12pp	-548	-595	-407	156	207	202
	(431)	(670)	(441)	(135)	(197)	(145)
tcurelsc12pp	-412	-721	-248	78.3	-74.3	40.9
	(405)	(718)	(417)	(157)	(229)	(158)
tcurinst12pp	-248	-470	-165	8.99	-95.0	-18.1
	(255)	(460)	(262)	(94.3)	(142)	(98.5)
tcurssvc12pp	-173	-256	-72.2	49.0	12.3	46.3
	(176)	(294)	(181)	(72.3)	(102)	(74.7)
tcuroth12pp	8.95	4.24	-10.7	20.4	8.39	12.6
	(28.0)	(40.2)	(33.8)	(17.4)	(22.7)	(20.6)
tnonelse12pp	-8.78	-0.88	11.5	53.6	52.3	57.3
	(45.1)	(69.9)	(48.6)	(37.7)	(68.4)	(44.1)
tcapout12pp	-193	-489	-303	-172	-582*	-271
	(266)	(451)	(308)	(266)	(335)	(264)
salaries12pp	-108	-371	22.5	85.3	30.3	76.6
	(235)	(402)	(250)	(85.5)	(114)	(88.0)
benefits12pp	-327*	-205	-266	-53.7	-46.0	-70.7
	(175)	(315)	(189)	(61.6)	(106)	(66.9)
stufees12pp	19.3	6.72	24.7	8.45*	2.17	9.89
	(31.7)	(45.4)	(33.3)	(5.13)	(8.33)	(6.61)
stucount	567	7,724	395	-152	236	297
	(7,029)	(9,166)	(7,461)	(228)	(249)	(402)
fracwhite	-0.080	-0.056	-0.073	-0.0046	-0.0025	-0.000067
	(0.080)	(0.094)	(0.090)	(0.0036)	(0.0063)	(0.0043)
fracblack	0.075	0.041	0.047	0.0049	0.0049	0.0016
	(0.052)	(0.088)	(0.066)	(0.0038)	(0.0050)	(0.0047)
frachisp	0.0054	0.0091	0.030	0.00049	-0.0020	0.00050
	(0.088)	(0.085)	(0.092)	(0.0027)	(0.0033)	(0.0029)
fracfrl	0.068	0.017	0.062	0.010	0.041**	0.017
	(0.050)	(0.080)	(0.064)	(0.016)	(0.017)	(0.014)
fraciep	0.0043	-0.0031	0.0088	0.0042	-0.010	-0.00099
	(0.0071)	(0.013)	(0.0078)	(0.0064)	(0.0090)	(0.0064)
fraclep	0.017	0.0069	0.024	-0.013	-0.037**	-0.016
	(0.017)	(0.023)	(0.021)	(0.016)	(0.017)	(0.017)
Polynomial	linear	quad.	quad.	linear	quad.	quad.
Bandwidth	+/-10	+/-10	+/-20	+/-10	+/-10	+/-20

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. Differences in levels are estimated based on the year prior to the election, and differences in pre-treatment trends do the same based on within-district differences in levels between one year prior and three year prior. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

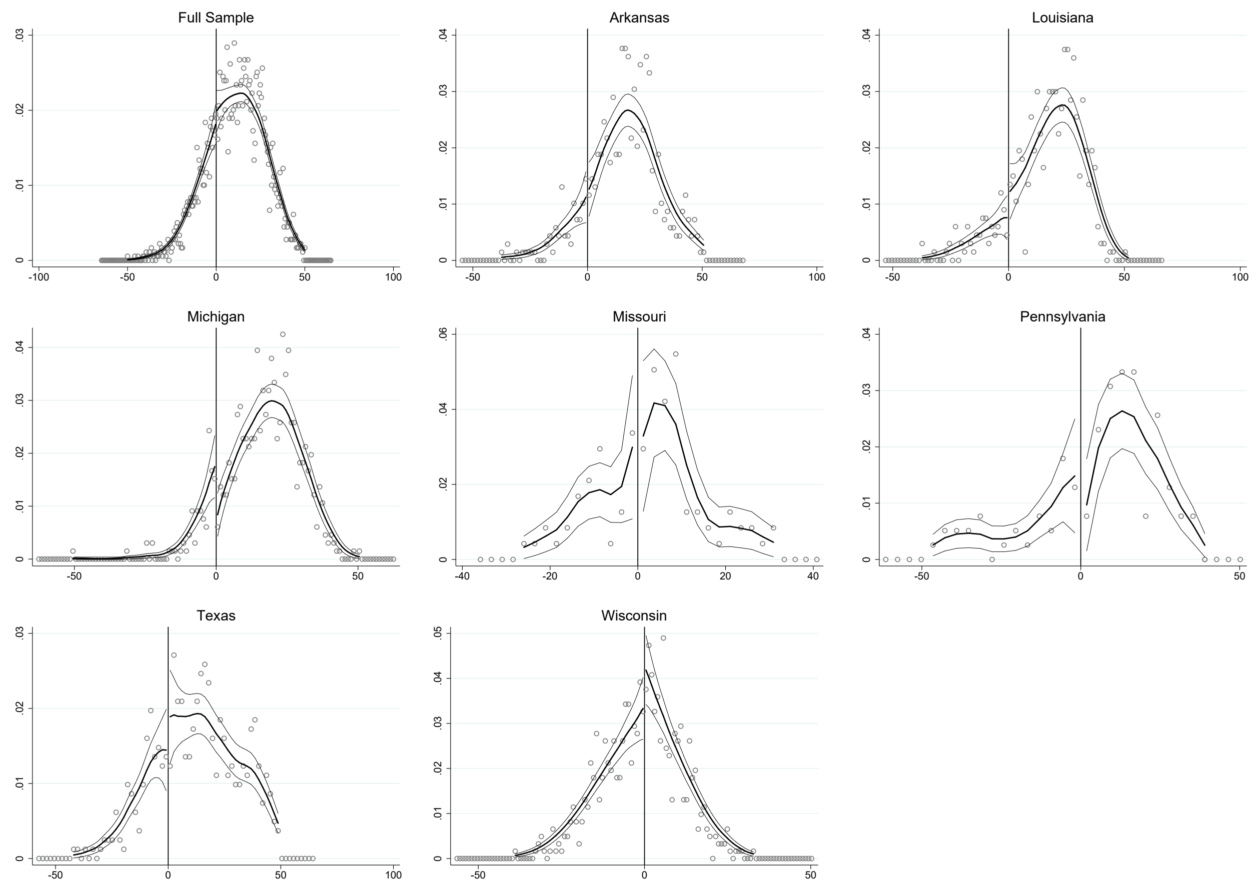
TABLE B1, cont'd. Pre-Election Balance Tests (various bandwidths and specifications)

	(1)	(2)	(3)	(4)	(5)	(6)
	Level t_1	Level t_1	Level t_1	Trend (t_1-t_3)	Trend (t_1-t_3)	Trend (t_1-t_3)
teachfte	61.6	565	52.2	-3.28	8.39	47.1
	(473)	(606)	(500)	(14.0)	(17.8)	(31.8)
instaides	15.3	102	45.2	-18.3**	-15.8	2.39
	(114)	(185)	(138)	(8.68)	(11.8)	(13.4)
counselors	-44.2	18.5	-8.32	-39.5	-14.4	-8.67
	(54.7)	(85.6)	(58.8)	(26.6)	(33.1)	(31.1)
librarians	-2.32	3.77	-4.80	-0.76	1.14	1.96*
	(9.57)	(11.3)	(10.1)	(0.62)	(0.74)	(1.14)
schladmin	-6.11	38.1	5.49	5.71	-16.2	1.39
	(30.4)	(34.6)	(29.2)	(19.7)	(16.0)	(17.1)
schladminstaff	8.33	61.1	11.0	17.0	3.98	16.6
	(40.2)	(53.3)	(44.6)	(21.3)	(12.0)	(20.7)
distadmin	-4.58	-0.59	-0.11	-1.52	-2.73	-0.76
	(6.58)	(4.16)	(7.08)	(2.46)	(3.82)	(2.89)
distadminstaff	18.6	-82.2	30.0	23.4	-74.8	43.6
	(35.2)	(134)	(40.5)	(34.7)	(134)	(37.5)
stuschl	37.5	39.5	45.5	-10.1	-1.92	-12.4
	(39.0)	(52.1)	(42.2)	(9.17)	(12.5)	(9.54)
stuteafte	-0.66**	-0.66	-0.60*	-0.013	-0.11	0.24
	(0.32)	(0.48)	(0.33)	(0.15)	(0.23)	(0.40)
stustaff	-0.24	0.19	-0.33	0.054	0.37	0.17
	(0.23)	(0.40)	(0.25)	(0.12)	(0.23)	(0.29)
stuaide	139	-15.7	104*	-4.08	11.7	-13.8
	(98.7)	(52.8)	(61.9)	(25.2)	(24.3)	(30.7)
stuadmin	0.97	2.54	-2.62	-4.17	7.36	-7.38
	(4.67)	(7.38)	(4.57)	(4.96)	(8.44)	(5.24)
stucounsl	-44.2	18.5	-8.32	-39.5	-14.4	-8.67
	(54.7)	(85.6)	(58.8)	(26.6)	(33.1)	(31.1)
stulib	-48.3	-139	-25.5	-75.5	-189*	-399
	(138)	(172)	(147)	(87.1)	(113)	(288)
dpr01	-0.013	-0.0088	-0.039	-0.00023	0.011	-0.014
	(0.051)	(0.072)	(0.065)	(0.014)	(0.021)	(0.021)
Polynomial	linear	quad.	quad.	linear	quad.	quad.
Bandwidth	+/-10	+/-10	+/-20	+/-10	+/-10	+/-20
Joint Hyp. Test						
Chi2 (32)	39.57	50.04	38.10	27.58	30.59	21.97
Prob>Chi2	0.1680	0.0221	0.2114	0.6898	0.5377	0.9081

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. Differences in levels are estimated based on the year prior to the election, and differences in pre-treatment trends do the same based on within-district differences in levels between one year prior and three year prior. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Density Tests

Figure 2: Density Tests



Note: The figure presents the results of McCrary (2008) density tests.

D Additional Tables

TABLE D1. Revenues and Future Passage Prob. through 5 and 7 post-election years

	$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Future Ref Passage	-0.47***	-0.49***	-0.20**	-0.27**
	(0.10)	(0.12)	(0.089)	(0.11)
Total Revenue Per Pupil	510**	572***	423**	517**
	(200)	(213)	(212)	(220)
Local Revenue Per Pupil	210	268*	229	307*
	(147)	(145)	(153)	(163)
State Revenue Per Pupil	220*	243*	160	148
	(124)	(136)	(137)	(152)
Federal Revenue Per Pupil	83.5	114	241	313
	(68.8)	(75.5)	(167)	(210)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quadratic	linear	quadratic
Baseline Covariates	yes	yes	yes	yes
DV Lags	yes	yes	yes	yes
Demographic Controls	no	no	no	no
School Year FEs	yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture whether or not a district passes another referendum in the first five or 7 years after the election, revenues per pupil, local revenues per pupil, state revenues per pupil, and federal revenues per pupil. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D2. Panel models w/ election fixed effects as in Cellini et al. (2010), with weights

	ln(rev. pp.)	ln(rev. pp.	ln(op. exp.)	ln(op. exp.)
3 yrs prior	-0.015	-0.014	0.016	0.017
	(0.014)	(0.017)	(0.015)	(0.017)
2 yrs prior	-0.0083	-0.000063	0.012	0.011
	(0.013)	(0.016)	(0.012)	(0.012)
Election year	0.010	0.00099	0.00024	0.0063
	(0.016)	(0.016)	(0.013)	(0.016)
1 yr after	0.026*	0.034**	0.033**	0.036**
	(0.015)	(0.015)	(0.014)	(0.016)
2 yrs after	0.11*	0.11*	0.055*	0.042
	(0.060)	(0.054)	(0.030)	(0.029)
3 yrs after	0.099***	0.086**	0.080**	0.063*
	(0.033)	(0.035)	(0.031)	(0.034)
4 yrs after	0.087**	0.080**	0.090**	0.074*
	(0.036)	(0.036)	(0.040)	(0.041)
5 yrs after	0.062**	0.056**	0.094**	0.073*
	(0.024)	(0.026)	(0.044)	(0.044)
6 yrs after	0.052**	0.044*	0.072**	0.053
	(0.022)	(0.025)	(0.033)	(0.036)
7 yrs after	0.061**	0.059**	0.061**	0.053*
	(0.028)	(0.028)	(0.027)	(0.028)
8 yrs after	0.079	0.076	0.12**	0.12**
	(0.048)	(0.048)	(0.058)	(0.060)
9 yrs after	0.042	0.036	0.088*	0.092**
	(0.027)	(0.030)	(0.046)	(0.047)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quadratic	linear	quadratic
Election FEs	yes	yes	yes	yes
Schl Year FEs	yes	yes	yes	yes

Note: Each column presents a set of coefficients from the same regression. Each coefficient captures the difference in outcomes between districts in which referenda passed and those in which referenda failed, from three years prior to the election to nine years after the election. The year prior to the election is the baseline. The dependent variable is the natural log of revenues or operating expenditures. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D3. Logged expenditures through 5 and 7 post-election years

	$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
ln(operations)	0.045***	0.050***	0.037**	0.047**
	(0.016)	(0.016)	(0.017)	(0.019)
ln(instruction)	0.041***	0.042***	0.035**	0.040**
	(0.015)	(0.014)	(0.016)	(0.017)
ln(support services)	0.053*	0.066**	0.045	0.065*
	(0.029)	(0.031)	(0.033)	(0.035)
ln(capital outlays)	0.015	0.11	-0.014	0.093
	(0.15)	(0.18)	(0.16)	(0.18)
ln(salaries)	0.053***	0.058***	0.056***	0.069***
	(0.017)	(0.017)	(0.019)	(0.019)
ln(instructor salaries)	0.065***	0.056***	0.072***	0.071***
	(0.016)	(0.016)	(0.017)	(0.017)
ln(benefits)	-0.047	0.00045	-0.090**	-0.066
	(0.043)	(0.052)	(0.044)	(0.050)
ln(instructor benefits)	-0.056	-0.012	-0.099**	-0.082
	(0.046)	(0.051)	(0.048)	(0.053)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quadratic	linear	quadratic
Baseline Covariates	yes	yes	yes	yes
DV Lags	no	no	no	no
Demographic Controls	no	no	no	no
School Year FEs	yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture the natural log of various expenditures. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions also include controls for post-election student demographics and are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D4. No baseline covariates

	$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Op. expend. p.p.	122	156	403	211
	(456)	(526)	(637)	(696)
Math/ELA (SDs)	-0.000041	0.075	0.013	0.044
	(0.11)	(0.13)	(0.11)	(0.14)
Graduation rate	-3.79	-1.64	0.23	1.14
	(4.75)	(5.31)	(4.99)	(6.89)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quadratic	linear	quadratic
Baseline Covariates	no	no	no	no
DV Lags	no	no	no	no
Demographic Controls	no	no	no	no
School Year FEs	no	no	no	no

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across up to five or seven post-election years, and test scores and graduation rates in the highest year of the relevant range (year 5 or 7). All regressions are weighted by baseline student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D5. No weights

	$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Without Demog. Controls				
Op. expend. p.p.	277**	215*	258*	213
	(114)	(125)	(131)	(150)
Math/ELA (SDs)	0.053	0.050	0.072**	0.083**
	(0.036)	(0.041)	(0.036)	(0.041)
Graduation rate	4.51***	3.53**	3.82**	4.13**
	(1.60)	(1.71)	(1.69)	(1.80)
With Demog. Controls				
Op. expend. p.p.	268**	223*	298**	235
	(114)	(124)	(129)	(149)
Math/ELA (SDs)	0.027	0.032	0.041	0.047
	(0.032)	(0.035)	(0.031)	(0.034)
Graduation rate	4.03***	3.32**	2.77*	3.10*
	(1.49)	(1.58)	(1.66)	(1.67)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quadratic	linear	quadratic
Baseline Covariates	yes	yes	yes	yes
School Year FEs	yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across up to five or seven post-election years, and test scores and graduation rates in the highest year of the relevant range (year 5 or 7). All regressions are weighted by baseline student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D6. First observed referendum for each district (with weights)

	$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Op. expend. p.p.	643***	836***	641***	778***
	(199)	(284)	(214)	(289)
Math/ELA (SDs)	0.096**	0.14**	0.18***	0.16***
	(0.046)	(0.055)	(0.052)	(0.056)
Graduation rate	0.89	1.78	3.72*	4.17**
	(2.36)	(2.39)	(2.15)	(2.09)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quadratic	linear	quadratic
Baseline Covariates	yes	yes	yes	yes
DV Lags	no	no	no	no
Demographic Controls	yes	yes	yes	yes
School Year FEs	yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across up to five or seven post-election years, and test scores and graduation rates in the highest year of the relevant range (year 5 or 7). All regressions are weighted by baseline student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D7. No lagged DVs

		$\leq 5yrs$	$\leq 5yrs$	$\leq 7yrs$	$\leq 7yrs$
Expend. P.P.	Operations	470***	520***	465**	492***
(2012 dollars)		(172)	(186)	(190)	(186)
	Instruction	257***	228**	191*	174*
		(96.4)	(99.1)	(110)	(104)
	Support services	211*	298**	286**	347***
		(118)	(125)	(132)	(131)
	Capital outlays	60.1	76.2	83.1	122
		(188)	(185)	(189)	(201)
	Salaries	377***	473***	406**	541***
		(131)	(126)	(166)	(153)
	Instructor Salaries	339***	297***	320***	318***
		(84.2)	(83.0)	(88.6)	(87.3)
	Benefits	-114	-116	-142	-188*
		(88.8)	(99.6)	(93.2)	(98.6)
	Instructor Benefits	-115	-133*	-156*	-200**
		(72.4)	(76.0)	(86.6)	(81.9)
Staffing	Stu-teacher ratio	-1.33***	-1.04***	-1.21***	-1.06**
		(0.41)	(0.38)	(0.38)	(0.44)
	Stu-aide ratio	21.1	8.85	288	165
		(29.9)	(33.8)	(219)	(114)
	Stu-counselor ratio	-71.3	-22.4	-63.8	-38.5
		(54.8)	(52.0)	(53.4)	(54.9)
Bandwidth		+/- 10	+/- 20	+/- 10	+/- 20
Specification		linear	quadratic	linear	quadratic
Baseline Covariates		yes	yes	yes	yes
DV Lags		no	no	no	no
Demographic Controls		no	no	no	no
School Year FEs		yes	yes	yes	yes

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures, staffing, test scores, and graduation rates averaged across up to five or seven post-election years. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D8. Impact on spending and ed. outcomes (omit one state at a time)

	Omit AR	Omit LA	Omit MI	Omit MO	Omit PA	Omit TX	Omit WI	Poly./ Band.
Through Year 5								
Op. Exp.								
(pp, 2012$)	537***	260**	279	572***	466***	578***	409**	linear
	(174)	(127)	(188)	(170)	(171)	(214)	(182)	/ 10
	594***	297*	363*	632***	522***	576***	461**	quad.
	(183)	(153)	(205)	(181)	(183)	(204)	(202)	/ 20
Math/ELA								
(SDs)	0.13***	0.16***	0.12***	0.11***	0.12***	0.044	0.14**	linear
	(0.042)	(0.045)	(0.045)	(0.043)	(0.043)	(0.028)	(0.068)	/ 10
	0.13**	0.18**	0.15**	0.13**	0.14**	0.059	0.12	quad.
	(0.063)	(0.072)	(0.060)	(0.063)	(0.063)	(0.043)	(0.099)	/ 20
Graduation								
Rate	3.04*	3.69**	1.30	2.73*	2.65*	2.50*	1.97	linear
	(1.60)	(1.64)	(1.40)	(1.59)	(1.54)	(1.44)	(2.05)	/ 10
	2.48	4.79*	2.22	2.98	2.78	0.77	1.28	quad.
	(2.43)	(2.49)	(2.16)	(2.42)	(2.35)	(1.67)	(3.05)	/ 20
Through Year 7 with controls								
Op. Exp.								
(pp, 2012$)	436**	294**	290	546***	440**	440**	293	quad.
	(182)	(134)	(197)	(181)	(179)	(190)	(216)	/ 20
Math/ELA								
(SDs)	0.077**	0.070**	0.031	0.072**	0.064**	0.059*	0.066	quad.
	(0.032)	(0.034)	(0.030)	(0.030)	(0.031)	(0.035)	(0.043)	/ 20
Graduation								
Rate	4.53**	3.38*	0.98	4.48**	4.19**	3.98*	-0.45	quad.
	(2.02)	(2.03)	(2.02)	(2.07)	(1.99)	(2.05)	(2.43)	/ 20

Note: Each coefficient is from a separate regression and captures the difference in outcomes between districts in which referenda passed and those in which referenda failed. The dependent variables capture expenditures and staffing averaged across five or seven post-election years, and test scores and graduation rates in year 5 or 7. All models include school-year fixed effects and baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. The final column indicates whether the models feature a linear or quadratic specification and whether the sample is restricted to a bandwidth of 10 or 20 percentage points in the percent of votes in favor of a tax levy. Standard errors clustered at the district level appear in parentheses below coefficient estimates. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

TABLE D9. 2SLS Treatment on the Treated (TOT) Estimates

	5th Year	5th Year	7th Year	7th Year
First-stage DVs:				
Fraction of years treated	0.41***	0.47***	0.41***	0.46***
	(0.075)	(0.10)	(0.087)	(0.11)
F-test of excluded instrument	*30.59****	*22.13****	*22.01****	*16.69****
Second-stage DVs				
Operational Spend. Per Pupil (avg. through year 5 or 7)	1,031**	1,154**	1,165*	1,064**
	(512)	(506)	(604)	(520)
Math/ELA in year 5 or 7 (SDs)	0.16	0.16	0.16*	0.12*
	(0.10)	(0.10)	(0.084)	(0.072)
Grad. Rate in year 5 or 7	3.54	2.98	11.2	8.96*
	(3.56)	(3.68)	(6.96)	(4.81)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quad.	linear	quad.
Baseline Covariates	yes	yes	yes	yes
DV Lags	yes	yes	yes	yes
Demographic Controls	yes	yes	yes	yes
School Year FEs	yes	yes	yes	yes

Note: Each coefficient is from a separate 2SLS regression. The first stage regresses the fraction of the post-election years (including the election year) that follow a passing levy (e.g., coded 6/6 for a passing referendum, 5/6 for a failing referendum in a district that passes a tax in the subsequent year, 4/6 for a failing referendum in a district that passes a tax two years later, etc.) on the variables specified in equation 2. The first-stage coefficients and F-test results are those for the spending and achievement models. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

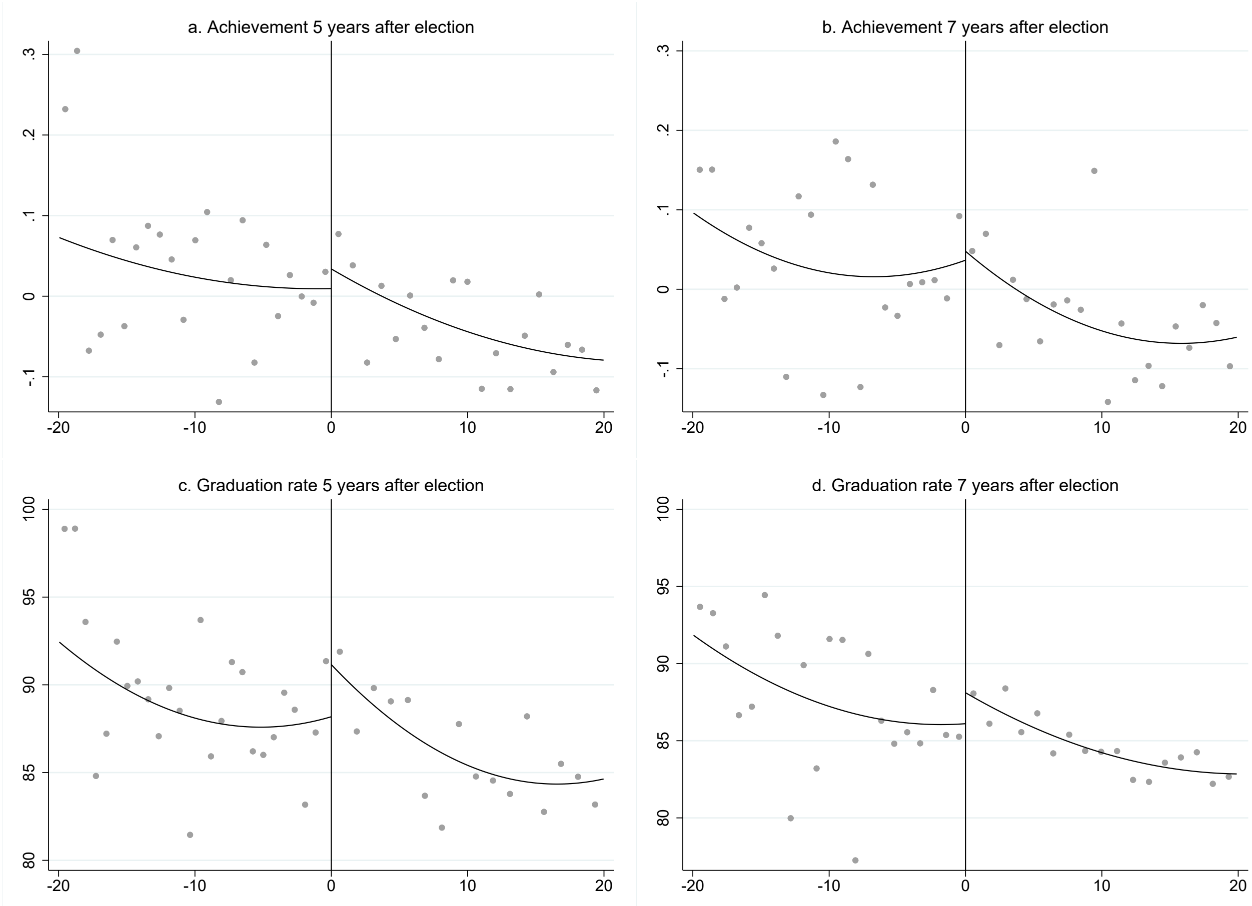
TABLE D10. 2SLS Estimates of the Impact of Spending ($1,000s)

	5th Year	5th Year	7th Year	7th Year
First-stage DVs for Math/ELA models:				
Avg. spending through yr 5 ($1,000s)	0.43**	0.54***		
	(0.018)	(0.019)		
Avg. spending through yr 7 ($1,000s)			0.48**	0.49***
			(0.20)	(0.19)
F-test of excluded instrument	*5.54****	*8.14*****	*5.63****	*6.66****
Second-stage DVs				
Math/ELA in year 5 or 7 (SDs)	0.16	0.14	0.14**	0.12
	(0.10)	(0.088)	(0.07)	(0.07)
Graduation Rate in year 5 or 7	5.03	3.67	7.32**	7.08**
	(5.81)	(4.78)	(3.41)	(3.21)
Bandwidth	+/- 10	+/- 20	+/- 10	+/- 20
Specification	linear	quad.	linear	quad.
Baseline Covariates	yes	yes	yes	yes
DV Lags	yes	yes	yes	yes
Demographic Controls	yes	yes	yes	yes
School Year FEs	yes	yes	yes	yes

Note: Each coefficient is from a separate 2SLS regression. The first stage regresses spending (in $1,000s) on the variables specified in equation 2. The first-stage F-test results are those for the achievement models. All models include baseline covariates capturing logged enrollments and expenditures per pupil one and three years prior to the election, as well as student demographics (fraction Black, fraction Hispanic, and fraction free/reduced-price lunch) one year prior to the election. All regressions are weighted by student counts. Standard errors clustered at the district level appear in parentheses below coefficient estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Additional Figures

Figure E1. Reduced-form RD plots for achievement and graduation rate 5 and 7 years after election



Note: The figure compares achievement (a and b) and graduation rates (c and d) between districts in which tax referenda passed (vote margin > 0) and those in which referenda failed (vote margin < 0) as of 5 years after the election and 7 years after the election. The x axes capture the vote margin—the difference between the percent of votes in favor of passage and the threshold needed for passage (typically 50 percent). The dots are local means and the curves are fitted using a quadratic polynomial. The sample is limited to levies within 20 percentage points of passage and those for which we observe post-election achievement in at least one year.