School Finance Reform and the Distribution of Student Achievement*

October 2015

PRELIMINARY AND INCOMPLETE

Julien Lafortune
University of California, Berkeley
julien@econ.berkeley.edu

Jesse Rothstein
University of California, Berkeley
and NBER
rothstein@berkeley.edu

Diane Whitmore Schanzenbach
Northwestern University
and NBER
dws@northwestern.edu

ABSTRACT

We study the impact of post-1990 school finance reforms, during the so-called "adequacy" era, on the distribution of school spending and student achievement between high-income and low-income school districts. Using an event study design, we find that reform events – court orders and legislative reforms – lead to sharp, immediate, and sustained increases in mean school spending and in relative spending in low-income school districts. Using test score data from the National Assessment of Educational Progress, we also find that reforms cause gradual increases in the relative achievement of students in low-income school districts, consistent with the goal of improving educational opportunity for these students. The implied effect of school resources on educational achievement is large.

* This research was supported by funding from the Spencer Foundation and the Washington Center for Equitable Growth. We are grateful to Patrick Lapid, Apurba Chakraborty, and Elora Ditton for excellent research assistance. We thank Tom Downes, Rucker Johnson, Bo Jackson, and conference and seminar participants at APPAM, AEFP, and Brookings for helpful comments and discussions.

Introduction

Schools are a key link in the transmission of economic status from generation to generation: Children from low-income families have lower test scores, lower rates of high school and college completion, and eventually lower earnings.[2] The achievement gap between rich and poor children has widened in recent years, even as racial gaps have shrunk (Reardon 2011). One potential contributing factor to gaps in educational outcomes is inequity in school resources. U.S. schools are traditionally funded out of local property taxes, and because wealthier families tend to live in richer communities with larger tax bases, their children have tended to attend schools that spend more than do those attended by the children of low-income families.

There is a longstanding debate in the education policy literature about whether additional school resources are productive (see, e.g., Hanushek 2003; Krueger 2003; Burtless 1996). Time series and cross-district observational comparisons tend to show small or zero effects of school spending on academic achievement (Hanushek 2006; Coleman et al. 1966), though state-level comparisons (Card and Krueger 1992) and randomized experiments (Krueger 1997; Chetty et al. 2011) are more positive.

A major issue in the debate is the extent to which observational designs are confounded by omitted factors, as compensatory funding allocations – additional state aid for disadvantaged school districts – would create a downward bias in the estimated effect of school resources. But it is exactly these allocations that are of

[2] See Barrow and Schanzenbach (2014) for a review of this literature.

interest for policy evaluation, as the state funding formula is the main tool available to policymakers aiming to address inequities in academic outcomes. Indeed, the redesign of state funding formulas has been a locus for reform efforts. Beginning with the 1971 *Serrano v. Priest* decision, in which a federal court found California's finance system unconstitutional, many U.S. states have moved away from local school funding to more centralized systems aimed at increasing opportunity for low-income students.[3]

A long literature examines the implications of these reforms for the distribution of school spending (see, e.g., Ladd and Fiske, 2008; Hanushek and Lindseth 2009). Most relevant for our study, Corcoran and Evans (2008; see also Corcoran et al., 2004) find that plaintiff court victories are associated with reduced inequality of spending across districts, driven by increases in low-spending districts, while Card and Payne (2002) find that these victories lead to increased relative spending in districts with low family incomes (which may or may not be low-spending districts). Fischel (1989) and Hoxby (2001) argue, however, that poorly designed reforms sometimes led to "leveling down" of spending in high-income districts rather than through increased spending in low-income districts.

Leveling down was possible because reforms in the 1970s and 1980s were focused on reducing gaps in funding between rich and poor districts. A new wave of reforms in the 1990s was based on a different legal theory: That state constitutions required not just *equitable* education spending but an *adequate* level of educational

[3] Cascio and Reber (2013) and Cascio, Gordon, and Reber (2013) examine an earlier form of school finance reform, the introduction of federal Title I funding to low-income schools via the 1965 Elementary and Secondary Education Act.

quality. In judging adequacy, courts focused on the *level* of spending in low-income districts, so there was less scope to level down in response to an adverse ruling.

Finance reforms are arguably the most important policy for promoting equality of educational opportunity since the turn away from school desegregation in the 1980s. Although intellectual attention has shifted in recent years to accountability and other process reforms as more important levers for educational opportunity, finance policy changes remain quite important, with at least 20 school finance reform cases decided since 2000. Several authors have examined individual adequacy-based reforms as case studies.[4] But to our knowledge Sims (2011) is the only systematic study of the effects of these reforms, taken as a group, on realized school finance, and Sims' sample ends in 2002. There is thus little known about the effect of adequacy-based reforms on realized school spending.

An even bigger gap in the literature concerns the impact of school finance reforms on student outcomes. As noted above, a long but inconclusive literature attempts to identify the effects of school spending using observational variation. But school finance reforms are the means by which state policymakers can influence spending, so represent highly policy-relevant variation in spending. They are also discrete events, with timing due more to legal processes than to potentially endogenous trends in other determinants of student outcomes, making them attractive candidates for natural experimental analyses of the causal effects of spending on outcomes. The barrier to this has been the absence of nationally comparable student outcome data. A few authors have tried to circumvent this by

[4] See, e.g., Clark (2003) and Flanagan and Murray (2004) on Kentucky, and Hyman (2013), Papke (2005, 2008), Cullen and Loeb (2004), and Chaudhary (2009) on Michigan.

examining particular states (Clark 2003; Hyman 2013); by focusing on the selected subset of students who take the SAT college entrance exam (Card and Payne 2002); or by examining less proximate outcomes like eventual educational attainment, health, and labor market outcomes (Jackson, Johnson, and Persico, forthcoming).

We provide the first evidence from nationally representative data regarding the impact of school finance reforms on student achievement. We rely on data from the National Assessment of Educational Progress (NAEP), also known as "the Nation's Report Card." The NAEP standardized assessments in math and reading have been administered to representative samples of hundreds of thousands of 4th and 8th grade students roughly biannually since 1990.[5] The rarely-used NAEP microdata allow us to examine the distribution of student achievement across high- and low-income school districts in each state over time.

Conveniently, the beginning of the NAEP panel coincides with the beginning of the adequacy era of school finance, which is often dated to the 1989 decision in *Rose v. Council for Better Education*[6] in Kentucky and the resulting Kentucky Education Reform Act (KERA) of 1990. Our results thus pertain to the more modern type of finance reform, which as noted above is likely to have had different impacts – closer to policymakers' intents – than did earlier equity-based reforms.

The first part of our analysis documents the impact of post-1990 reforms on absolute and relative spending levels in low- and high-income school districts. Using an event study framework, we indeed find that reforms lead to sharp, immediate,

[5] The NAEP testing program began in the early 1970s. But until the "state NAEP" was introduced in 1990, with the aim of providing state-level estimates, samples were too small to support the analysis we undertake here.
[6] 790 SW 2d 186.

and sustained increases in state revenues and total spending in low-income districts. There are no signs of negative impacts on spending in high-income districts; rather, these impacts are generally positive as well, though smaller than those in the districts targeted by the reforms. Although there is some evidence of subsequent reductions in local effort in high-income districts, even in these districts reforms have a positive effect on total revenues for at least a dozen years.

We use two measures of the progressivity of a state's school finance system: the slope of per-pupil revenues with respect to a district's log mean household income, and the gap in mean revenues between districts in the first and fifth quintiles of the state's district mean income distribution. Equity-increasing reforms should reduce the slope (making it less positive or more negative) and increase (make more positive or less negative) the Q1-Q5 gap. We find that the average reform does exactly this, through a substantial increase in the progressivity of state aid to local districts that does not seem to be clawed back through local revenue responses: The impact on the progressivity of total revenues is nearly as large as (and statistically indistinguishable from) the impact on the progressivity of state aid. Again, these effects are immediate following the reform event and persist or even grow over at least the next decade.

We next turn to student outcomes, focusing on analogous measures of the relationship between district mean test scores and the log mean household income in the school district. Using a similar event study framework, we find that the "progressivity" of test scores grows significantly – that scores rise in low-income districts relative to high-income districts – in the years following a finance reform.

This is exactly what one would expect if school resources are productive in low-income, poorly resourced settings. Our results imply that the (local) average effect of an extra $1,000 per pupil in spending is to raise student test scores ten years later by 0.18 standard deviations. This is roughly twice as large as the effect implied by the Project STAR class size experiment (which, translated into these terms, corresponded to an approximately 0.085 SD effect per $1,000 per pupil[7]). Chetty et al. (2011) find that each 0.1 SD increase in a student's test score in early elementary school is associated with an additional $5,350 in present discounted value lifetime earnings, suggesting that marginal increases in school resources in low-income, poorly resourced school districts are cost effective even when the only benefits considered are those operating through subsequent earnings.

A final analysis considers the impact of finance reforms on overall educational equity, measured as the gap in achievement between high- and low-income students or between white and minority students. We find no discernable effect of reforms on either gap. The reason is that low-income and minority students are not very highly concentrated in school districts with low mean incomes, so are not very closely targeted by district-based finance reforms. Our estimates indicate that the average reform event raises relative spending in low-income districts by over $500 per pupil per year, but that this raises relative spending on the average low-income student by under $100 (not statistically distinguishable from zero).

[7] STAR raised costs by about 30% in K-3, and raised test scores by 0.17 SDs. Current spending per pupil in Tennessee is around $6,700, so STAR would today cost around $2,000 per pupil. We thus divide the test score effect by two. This comparison implicitly assumes that maintaining the smaller class sizes beyond 3rd grade would yield no additional growth in test scores.

Thus, while our analysis suggests that finance reforms can be quite effective at reducing between-district inequities, other policy tools aimed at *within*-district resource and achievement gaps will be needed to address the overall gap.

I. School finance reforms[8]

American public schools have traditionally been locally managed and financed out of local property tax revenue. As local jurisdictions vary widely in their tax bases and inclinations to fund local schools, this has meant that the resources available to a child's school depended importantly on where he or she lives.

In the groundbreaking *Serrano v Priest*[9] decision in 1971, the California Supreme Court accepted a novel legal theory (propounded in various forms by Wise 1967; Horowitz 1966; Kirp 1968; and Coons, Clune, and Sugarman 1970; among others) that the Equal Protection Clause of the U.S. Constitution created a right of equal access to good schools. California's legislature responded by implementing a highly centralized school finance system where per-pupil resources are nearly perfectly equalized across districts.

The U.S. Supreme Court rejected the Equal Protection theory in *San Antonio Independent School District v. Rodriguez*[10] in 1973. Reform efforts then shifted to state courts. Unlike the U.S. Constitution, many state constitutions address education specifically. Courts in many states found requirements for greater equity in school finance, while other states' legislatures acted without court decisions (perhaps to stave off potential rulings). The new finance regimes created in this

[8] Our discussion here draws heavily on Koski and Hahnel (2015).
[9] 487 P.2d 1241.
[10] 411 US 1.

second wave of reforms took a variety of forms, ranging from California-style centralization of school finance to "power equalization" formulas that aimed merely to provide poor districts with similar tradeoffs between tax rates and spending as are faced by rich districts. These second-wave reforms proceeded through the 1970s and 1980s, and have been much studied (see, e.g., Hanushek and Lindseth, 2009; Corcoran and Evans, 2008; Card and Payne, 2002; Murray, Evans, and Schwab, 1998).

We focus on the much less studied third wave of adequacy-based finance reforms. These began in 1989 when the Kentucky Supreme Court found that the state constitutional requirement for an "efficient system" of public schools required that "[e]ach child, *every child*, … must be provided with an equal opportunity to have an adequate education" (*Rose v. Council for Better Education*[11]; emphasis in original). The decision made clear that adequacy was defined in terms of outcomes, and required more than equal inputs (e.g., "sufficient levels of academic or vocational skills to enable public school students to compete favorably with their counterparts in surrounding states, in academics or in the job market"). To achieve this, spending would need to be increased substantially in low-income districts, which had historically underperformed in Kentucky and elsewhere. Indeed, subsequent courts and legislatures have often interpreted this sort of requirement to demand *higher*

[11] 790 SW 2d 186.

spending in low-income than in high-income districts, to compensate for the disadvantages that low-income students bring to the educational system.[12]

The Kentucky legislature responded quickly with the Kentucky Education Reform Act of 1990 (KERA), which revamped the state's educational finance, governance, and curriculum. KERA led to substantial increases in spending per pupil in low-income districts, and the correlation between district median income and total current expenditures per pupil went from positive to negative (Clark 2003; Flanagan and Murray 2004).

Since 1990, many other state courts have found adequacy requirements in their own constitutions. We identify reform events in 27 states over this period, many of them adequacy based. We discuss our tabulation of post-1990 finance reform events – court orders and major legislative changes – in Section II.

As with earlier equity-based reforms, there has been no single definition of adequacy, and states have varied in the finance systems that they have adopted. Despite this heterogeneity, there is reason to believe that adequacy-based reforms will have different implications for the level and distribution of school funding than did earlier reforms predicated on equity principles. Where an equity-based court order might permit leveling down to a stingy but equal funding formula, a state cannot satisfy an adequacy mandate by leveling down. Many states seem instead to have leveled *all* districts up in order to meet adequacy criteria in low-income districts while still allowing higher-income districts to differentiate themselves.

[12] A long literature has developed concerning the calculation of spending levels needed to satisfy an adequacy standard. See, e.g., Downes and Steifel, 2008, and Duncombe, Nguyen-Hoang, and Yinger, 2008.

Overall, then, one might expect that adequacy-based reforms would lead to higher spending across the board than would equity-based reforms, but perhaps also to smaller reductions in inequality (Baker and Green, 2015; Downes and Stiefel, 2015). This points to the importance of examining both the average impact of reforms and their differential effect on low-income vs. high-income school districts. We develop a framework to assess both in the next section. Later, we apply it to study impacts on both spending levels (Section IV) and student test scores (Section V).

II. Analytic approach

We develop our analytic approach in three parts. First, we describe the reform event database that we use. Second, we discuss our summary measures of the progressivity of school finance and student outcomes in each state in each year. Third, we discuss our event study methodology for relating reform events to subsequent outcomes.

A. Characterizing events

The most clear cut school finance reform events are when state supreme courts find the state school financing systems to be unconstitutional, and order changes in the funding formula. Much of the prior school finance reform literature has focused on court-ordered reforms, and we start by compiling these, drawing on lists in Jackson et al. (forthcoming), Hanushek and Lindseth (2009), and Corcoran and Evans (2015) and supplementing them with our own research into case histories. We focus on events in 1990 and thereafter, corresponding both to the period covered by our NAEP panel (discussed below) and to the adequacy era of school finance reform.[13] (A complete list of our events, along with a comparison to those used in other studies, is presented in Appendix Table A1).

We use an inclusive definition of events, including many court orders that were subsequently reversed or were ignored by the legislature. We date events to the court judgment – typically a Supreme Court or a significant appellate decision – not to actual flows of money (which may never occur). In contrast to some prior work, we do not restrict attention to initial orders, but we also try not to label every single procedural ruling a separate event.[14] In particular, when a lower court decision is stayed pending appeal, we do not count the event until a higher court upholds the initial decision and lifts the stay.

Not all major school finance reform events resulted from court orders. In some important cases (e.g., California, Colorado), legislatures reformed finance systems without prior court decisions, perhaps to forestall adverse judgments in threatened or ongoing lawsuits. As a result, we also include major legislative reforms that change school finance systems in our event list.

As shown in Figure 1, we identified a total of 68 events in 27 states between 1990 and 2013. 51% were court orders and 40% were legislative actions; in 9% of cases, we identified one of each in the same year, and count them as a single event. Figure 2 shows the geographic distribution of the events, using shading to represent the date of the first post-1989 event and numerals to indicate the number of events.

[13] Note that the 1990 start date is after the 1989 decision in the *Rose* case in Kentucky but does capture the KERA legislation that it spurred.
[14] The appendix presents analyses that use just initial post-1989 events (Appendix Table A3), with qualitatively similar results.

Reform events are widely spread throughout the country, with the exceptions of the deep South and upper Midwest.

B. Measuring school finance systems and student outcomes

Next we turn to the measurement of the independent variables of interest. Some authors categorize school finance systems by the form of the formula itself (e.g., minimum foundation plan, power equalization, etc. – see Hoxby, 2001 and Card and Payne 2002), while others distinguish based on the legal theory (equity vs. adequacy) that led to the system's implementation. But finance formulas do not always conform to the simple categories that researchers have used, and even two states with formulas of the same type may vary substantially in the extent of intended or actual redistribution. And while the underlying legal theory may guide the choice of a new finance regime, it does not determine it.

A more promising approach measures finance systems by the realized distribution of resources. Here, a challenge is how to summarize the distribution and its movement over time. Corcoran and Evans (2015), for example, examine the standard deviation of spending per pupil and other summaries of the univariate distribution. This approach does not account for the relationship of spending to area economic resources. Since the central issues in school finance litigation and reforms are the equity of resource distribution across rich and poor districts and the adequacy of resources available to the lowest-income districts, we prefer a measure that corresponds more directly to these concepts. We consider both absolute and relative measures of funding in disadvantaged districts, corresponding roughly to the adequacy and equity of the funding system, respectively.

Our primary measure of school district disadvantage is the average family income in the district relative to the state average. The Appendix reports analyses using alternative measures (e.g., mean home values, or the share of families under 185% of poverty), with similar results.[15]

We use two measures of equity. The first is the difference in average per-pupil revenue – either in total or only from state sources – between districts in the bottom and top quintiles of the state family income distribution. But, while the extremes of the distribution are certainly of particular interest in equity discussions, one might also be interested in the distribution of resources for districts in the middle three quintiles. To summarize the relationship between spending and income across the entire income distribution, our second measure follows Card and Payne (2002) in measuring the bivariate relationship between finance and economic disadvantage across districts in the state. In measuring this relationship, we control for district enrollment and the grade span covered by the district to account for potential economies of scale and for differences in production functions across different grades. Thus, we estimate the following regression, separately for each state and each year:

(1) $R_{ist} = \alpha_{st} + \theta_{st} \ln(Y_i) + X_{ist}'\gamma_{st} + u_{ist}$.

[15] Some of the prior literature and litigation has focused on disparities in property tax bases, which are not perfectly correlated with family incomes or even home values. We are not aware of a nationally comparable measure of district property tax bases that takes account of variation in the definition of the tax base or in the presence of taxable non-residential property.

Here, R_{ist} measures revenues per student in district i in state s in year t, $\ln(Y_i)$ is the mean household income in the school district (measured in 1990), and X_{ist} contains the log district enrollment and indicators for elementary-only and secondary-only districts. The regression is weighted by log district enrollment.[16]

The θ_{st} coefficient summarizes the progressivity or regressivity of the finance system. A more positive coefficient means greater gaps in funding between high- and low-income districts, as would generally be expected with local finance, while a negative coefficient (observed in about 40% of the state-year cells in our sample) means that per-pupil revenues are negatively correlated with mean household incomes across districts in the state. Both equity and adequacy theories of school finance reforms can imply higher spending levels in lower-income districts, in recognition of their greater needs.

When we turn to our examination of student outcomes, we use parallel measures to those used in our finance analysis: For adequacy, the mean test scores of students at districts in the bottom quintile of the family income distribution; for equity, the gap between this mean and the mean at districts in the top quintile, and the slope from a regression of mean test scores on district family income.[17] Each is estimated separately for each available state-year-subject-grade combination.

---

[16] Our weight variable is the mean log enrollment in the district across all years in the sample, so does not vary over time. This reduces volatility in the estimates. By contrast, the enrollment measure in the X_{ist} vector is the time-varying log enrollment from year t, to capture sensitivity of funding formulas to district scale.

[17] The regression used to estimate test score slopes differs in minor ways from that used for the finance measures. See the Appendix.

C. Ohio Case Study

To illustrate these measures and their relationships to the school finance reform events, we present Ohio as a case study. Figure 3 shows how the relationship between district income and state revenues in Ohio changed between 1990 and 2010. On the horizontal axis is the log of the average household income in a school district, as reported in the 1990 School District Data Book, a tabulation of 1990 Census data at the district level. On the vertical axis, we show state revenues per pupil, in inflation-adjusted 2013 dollars, in 1990 (left panel) and 2011 (right panel). These data come from the National Center for Education Statistics (NCES) Common Core of Data (CCD) School Finance Survey. (We discuss the data sources at greater length in Section III.) In each panel, we overlay a regression lines with slope θ_{st}.[18] The slope is negative in both years, indicating progressive state funding to districts, but is much more negative in 2011 than in 1990. In 1990, each 10% increase in mean household income was associated with about $144 less in state aid per pupil; the corresponding figure in 2011 is $469. The change in slope is driven by a dramatic increase in state aid to low-income districts. Higher-income districts also saw increases, but their gains were much smaller.

The dashed lines in Figure 3 show mean state revenue per pupil for districts in each quintile of the mean log income distribution. In 1990, the bottom quintile Ohio districts received an average of $1,102 per pupil more than did the top quintile districts, but by 2011 this had grown to $3,387. These gaps are quite similar to what is implied by the log-linear regressions from the two years.

[18] Following (1), we control for enrollment and grade span in estimating θ_{st}; the lines in Figure 3 correspond to a unified district with mean enrollment.

Figure 4a presents the scatterplot of state revenue-income slopes, θ_{st}, in 1990 and 2011 across all states. It shows that Ohio, highlighted in the figure, is not an outlier. Fully 39 states are below the 45 degree line, indicating smaller slopes (more progressive distributions) in 2011 than in 1990.

State transfers represent only about half of total district revenue nationally, and state-level formula changes may have been offset by local effort responses (Hoxby 2001). Moreover, there is some arbitrariness in whether certain tax revenues (e.g., those raised by a property tax that is locally administered but set at the state level) are labeled as local or state revenues. Thus, changes in state revenues may not be reflective of changes in the actual resources available. Figure 4b shows the corresponding scatterplot for the slope of total revenues, inclusive of state revenues, local tax collections, and federal transfers, per pupil with respect to district income. Although total revenue slopes are generally larger and more often positive – while state revenue formulas are often progressive, local tax collections are not – we see a similar pattern of declining gradients over time in most states.

Figure 3 shows that Ohio's finance formula changed substantially between 1990 and 2011, and Figure 4 shows that this is not an isolated case. But to what extent were the changes due to intentional reforms? To answer this, we need to relate the changes in finances to the reform events described earlier. In the clearest cases, a court decision finding the state's finance system to be unconstitutional results in a prompt, discrete change in spending. Often, however, there is a complex interaction between the courts and the legislature, with multiple court decisions and legislative changes over many years, and spending changes gradually.

Ohio is again a useful illustration. The state Supreme Court ruled four times on the *De Rolph v. State* case, in 1997, 2000, 2001, and 2002. The 1997 ruling declared the state's finance system unconstitutional, on adequacy grounds, and specifically rejected the state's reliance on local property taxes. The court ordered the legislature to enact a "complete systematic overhaul" of the school funding system. In 2000, the Ohio Supreme Court determined that the legislature had failed to act and that funding levels remained inadequate. The same year, the legislature revised the system and a subsequent ruling in 2001 determined that the new system would now be constitutional as long as the legislature enacted a few minor, additional changes. In 2002, the court – with new judges since the previous year – reversed itself, determining that the state was not in compliance after all. To our knowledge, there have not been substantial reforms to the finance system since then. We code Ohio as having reform events in 1997, 2000, and 2002. The 2000 event is a joint statutory-judicial event, while the others are judicial.

Figure 5a shows the estimated state revenue-income and total revenue-income slopes over time for Ohio. Vertical lines indicate the reform events: Plaintiff court victories in 1997, 2000, and 2002, with a statutory reform also in 2000. The figure shows a clear effect of the first decision, with gradual declines in each gradient between 1997 and 2002 following a period of stability before 1997. There is less visual evidence of an effect of the 2000 ruling, which does not seem to have interrupted the previous trend, while the 2002 ruling seems to coincide with an *end* to the decline in the gradient. Indeed, there was some backsliding in 2002-2005, though in broad terms the gradients were stable from 2002 to 2011. There is little sign that changes in state aid are offset through changes in local effort, as the two sets of gradients move in parallel throughout the period. Figure 5b presents similar time series evidence for the differences in mean state aid or total revenue between districts in the bottom and top quintiles of the Ohio district mean income distribution. Again, it is not hard to see evidence that the initial 1997 ruling had an eventual effect, though the only sharp break came following the 2002 ruling and has the opposite of the expected sign.

D. Event study methodology

To model the relationship between school finance reform events and measures of school finance progressivity, we adopt an event-study framework. Our strategy is based on the idea that states without events in a particular year form a useful counterfactual for states that do have events in that year, after accounting for fixed differences between the states and for common time effects.

We estimate parametric and non-parametric models for the effect of an event on subsequent outcomes. In the non-parametric model, we specify the outcome for state s in year t as:

$$(2) \qquad \theta_{snt} = \delta_{sn} + \kappa_t + \Sigma_{r=-kmin}^{+kmax}\, 1(t = t_{sn}^* + r)\beta_r + \varepsilon_{snt}.$$

Here, n indexes the potentially several events in a state. We discuss this below; for now, consider the case where each state has only a single event. β_r represents the effect of an event in year t_{sn}^* on outcomes r years later (or previously, for r<0). These effects are measured relative to year r=0, which is excluded. We censor r at kmin=-5, so β_{-5} represents average outcomes five or more years prior to an event, relative to those in the event year. κ_t is a calendar year effect that is assumed to be constant across states, while δ_{sn} represents a fixed effect for each copy of each state's data.

As discussed above, we use several measures for the outcome, θ_{snt}. One is the slope of revenues or test scores with respect to log district income. Another is the mean revenue (or test score) among districts in the bottom quintile of the state district mean income distribution; a third is the difference between this mean and the mean in the top quintile. The event study models for slopes are weighted by the inverse estimated sampling variance of θ_{snt}; models for quintile means or gaps are unweighted. In each case, error terms are allowed to be freely correlated among observations from the same state, but are assumed to be independent across states.

The event study framework yields estimates of the causal effects of events if event timing is random, conditional on state and year effects. This need not be true. The interplay between courts and legislatures may produce changes in finance or outcomes in the years immediately prior to our identified events – for example, when a court responds to an inadequate reform effort from the legislature, as in Ohio in 2000 and 2002. Our inclusion of terms capturing pre-event dynamics – $\{\beta_{-k}, \ldots, \beta_{-1}\}$ in (2) – is designed to capture this. An indication that these coefficients were non-zero would suggest that we are unable to distinguish the causal effects of events from the prior dynamics that led to them. As it happens, in none of the specifications that we examine do we find that the pre-event effects are meaningfully or significantly different from zero. This supports our reliance on an event study framework.

In specification (2), the effect of the event is allowed to be entirely different in each subsequent and prior year. We present estimates from this nonparametric specification, but we focus our attention on a more parametric specification that replaces the relative time effects in (2) with three parametric terms:

(3) $\theta_{snt} = \delta_{sn} + \kappa_t + (t-t_{sn}^*)\beta^{trend} + 1(t > t_{sn}^*)\beta^{jump} + (t-t_{sn}^*)*1(t > t_{sn}^*)\beta^{phasein} + \varepsilon_{snt}$.

Here, β^{jump} captures a discrete change in the outcome following the event, while $\beta^{phasein}$ captures a gradually growing event effect that produces a kink in the linear trend on the date of the event. β^{trend} represents a linear trend that predates the event and continues afterward, and is interpreted as a potential confound, analogous to the pre-event effects in (2), rather than as the effect of the event itself. As before, this coefficient is never practically significant. Comparisons of the parametric and non-parametric estimates indicate that the three-coefficient structure does a good job of capturing dynamics in outcomes surrounding events, though the change captured by the post-event "jump" coefficient is sometimes delayed a year or spread out over two to three years following the event.

A complication we face in implementing the event study framework is that states may have multiple events. In our preferred estimates, we treat each of several events in a state separately.[19] Specifically, suppose that state s has event number n (out of N_s total events) in year t_{sn}^*. We create N_s copies of the state-s panel, labeling them n=1…N_s, and we code copy n as having a single event in t_{sn}^*. (For states without events, we make a single copy and set all relative time variables to zero.)

[19] Results are qualitatively unchanged when we use only the first event in a state, when we reweight so that states with multiple events are not overrepresented, or when we use one panel per state with a running count of events to date as the key variable. See Appendix Table A3.

This yields a panel data set characterized by three dimensions – state, time, and event number, where the first two dimensions are balanced but the number of events varies across states. We use this panel data set to estimate equations (2) and (3), with state-event and year fixed effects.

Our decision to treat each of several events in a state separately affects the interpretation of the post-event coefficients. The coefficient β_r, r>0, estimates the reduced-form effect of an event in year t_{sn}^* on the outcome measure in $t_{sn}^* + r$, not holding constant subsequent events.[20] In some cases it takes many events (e.g., court rulings) before the finance reform is actually implemented. Thus, gradual increases in β_r may not indicate that states are slow to implement new finance formulas, but rather that the true finance formula change did not occur for several years after one of our focal events. As we show below, this is not very important empirically – effects on finance outcomes appear almost immediately following our designated events, and persist without growing thereafter.

We also use equations (2) and (3) to investigate student outcomes, replacing the dependent variable with test score-income slopes or between-quintile gaps in mean scores and replacing the year effects κ_t with subject-grade-year effects. We expect a different time pattern of effects here. Because student outcomes are cumulative and a sudden infusion of resources in 8th grade is not likely to have as large an effect as would a flow of resources every year from Kindergarten onward, we expect the primary effect of reforms on student outcomes to occur through the $\beta^{phasein}$ coefficient or, alternately, through gradual growth in the β_rs.

[20] See Cellini, Ferreira, and Rothstein (2010) on the interpretation of reduced-form event study effects with repeated events.

III. Data

Our analysis draws on data from several sources. We begin with our database of school finance reform events, discussed above. We merge this to district-level school finance data, from the National Center for Education Statistics' (NCES) Common Core of Data (CCD) school district finance files (also known as the "F-33" survey) and the Census of Governments; demographics, from the CCD school universe files; household income distributions, from the 1990 Census; and student achievement outcomes in reading and math in 4^{th} and 8^{th} grade, from the NAEP.

The CCD district finance data, collected by the Census Bureau on behalf of NCES, report enrollment, revenues and expenditures annually for each local education agency (LEA). The census has been collected annually since school year 1994-95, as well as in 1989-90 and 1991-92. We supplement it with data from the Census Bureau's Annual Survey of Government Finances for 1992-93 and 1994-95. We convert all dollar figures to 2013 dollars per pupil.[21] We use the CCD annual census of schools from 1986-87 through 2012-13, aggregated to the district level, for school racial composition, free lunch share, and pupil-teacher ratios.

We draw information on district-level mean household income from the 1990 School District Data Book, a tabulation of data from the 1990 Census to the school district level. We focus on the log of mean household income, and drop districts below the 2^{nd} or above the 98^{th} percentile of their state's (unweighted) distribution.

[21] We exclude districts with highly volatile enrollment (year-over-year changes of 15% or more in any year, or where more than one-third of years have enrollment more than 10% off of a log-linear trendline) and those with total revenue per pupil less than 20% or more than 500% of the (unweighted) state-year mean.

Finally, our measures of student outcomes come from the restricted-use NAEP microdata. Estimation of the achievement-income gradient in a state-year cell requires a dense sample. Accordingly, we limit attention to NAEP administrations designed to produce representative samples for each participating state. These began as the "State NAEP" in 1990, with 8th grade math and 42 states participating, and have been administered roughly every two years since (with subjects and grades staggered in the early years). Since 2003, there have been 4th and 8th grade assessments in both math and reading in every odd-numbered year, with all states participating.[22] Table 1 shows the schedule of assessments, the number of participating states, and the number of students assessed. We generally have over 100,000 students per subject-grade-year, with a representative sample of about 2,500 students in 100 schools per state.

The NAEP uses a consistent scoring scale across years for each subject and grade. We standardize scores to have mean zero and standard deviation one in 1990, 1992, or 1994, depending on the grade and subject, but allow both the mean and variance to evolve since then. We then aggregate to the district-year-grade-subject level and merge to the CCD and SDDB.[23] We estimate separate quintile mean test scores and income gradients for each state-year-subject-grade in our sample. We stack these data, so that our student outcome event-study sample consists of state-subject-grade-event number-year cells.

---

[22] The NAEP also collects 12th grade achievement data, but high school dropout makes the samples nonrepresentative. We do not use data from subjects other than math or reading, which are administered less frequently.

[23] The pre-2000 data do not use the same district codes as the CCD. We crosswalk using a link file produced for NCES by Westat (and obtained from the Educational Testing Service), using district names to check and supplement the crosswalk.

Table 2a presents summary statistics for our district-level analysis file, pooling data from 1990-2011. As discussed in Section III, we construct several summaries of finance and achievement outcomes in each state-year cell. Table 2b presents summary statistics for the state-year panel.

IV. Results: School Finance

We begin by investigating the effects of finance reform events on transfers from states to school districts, measured in real 2013 dollars per pupil. The solid line in Figure 6 presents estimates of the non-parametric event study specification (2), taking the income gradient of state revenues per pupil as the dependent variable. The state revenue gradient is roughly stable in the years leading up to a finance reform event, but declines by roughly $500 (scaled as dollars per pupil per one-unit change in log mean income) in the three years following the event. The gradient continues to decline over the next eight years, reaching a minimum total effect of -$937 in the 11^{th} year after the event before rebounding somewhat, but in rough terms is stable from about year seven onward. Dotted lines in the graph show pointwise 95% confidence intervals for the β_r effects. These are wide, but exclude zero in years 2-15. A test of the joint significant of all the post-event effects has a p-value less than 0.001, while the test that all pre-event effects equal zero has p=0.22.

Figure 6 also shows the parametric specification (3), as a dashed line. Not surprisingly, given the nonparametric results, this shows a small pre-event trend, a sharp downward jump following the event, and a slow continued decline in the state revenue gradient in subsequent years. This three-parameter model fits the non-parametric pattern quite well.

Columns 1-3 of Table 3 present estimates from various versions of the parametric specification. In column 1, we include only state and year effects and the post-event indicator (i.e., we constrain $\beta^{trend} = \beta^{phasein} = 0$ in equation (3)). Column 2 adds the phase-in effect, while column 3 also adds the trend term. (This third specification is what is shown in Figure 6.) The table also reports tests of the joint hypothesis that the post-event jump and the phase-in effect are both zero. These have p-values of 0.03 in columns 2 and 3. In column 3, both the trend and phase-in effects are small, and neither approaches statistical significance. Only the post-event effect is statistically significant or economically meaningful. We thus focus on the simpler specification in Column 1. Here, the post-event jump coefficient indicates that reform events lead to an immediate decline in the gradient of state aid per pupil with respect to district income of about \$500 per pupil, or about 5% of mean total revenues per pupil in our sample.

Figure 8 shows event study analyses for mean state revenues in the first and fifth quintiles of the district mean income distribution in the state (panels A and B) and for the difference between these (Panel C). The figure indicates sharp post-event increases in state revenues in the first quintile districts, with smaller but still substantial increases in the fifth quintile. The former effects grow in the years following the event, while the latter erode. As a result, the effect on the between-quintile gap (Panel C) is relatively small at first but grows substantially over time.

Closer inspection indicates that revenues are trending up in first quintile districts before the events and that there is little change in the trend following an event.

Estimates from the parametric model, in Table 4a, confirm this. None of the trend or post-event trend change coefficients are significant in either quintile, so we focus on the models without these terms in Columns 1, 3, and 5. They imply that state revenues rise by $1023 per pupil in first quintile districts after an event. The increase in fifth quintile districts is smaller, $510 (not significantly different from zero); the differential effect on first quintile districts is thus $518. The gap in mean log incomes between the first and fifth quintile districts is only about 0.6, so this is a larger increase in progressivity than is implied by the slope coefficients in Table 3.

Our reform event list intentionally includes many events that do not – because of subsequent judicial reversals or legislative foot-dragging – ever lead to implemented changes in school finance. We thus view our estimates of the effect of an event as intention-to-treat (ITT) effects, representing the average of the effects of actually implemented finance reforms with the null effects of events that did not lead to changes in funding formulas. The effects of implemented finance reforms are almost certainly larger than those that we estimate.

Districts may respond to changes in state transfers by changing their local tax rates. It is also possible for changes in the state aid formula to induce property value changes that affect local revenues with fixed rates (Hoxby 2001). We thus turn next to models for the progressivity of total revenues per pupil, inclusive of state and local components. Models for the district income slopes are presented in Figure 7 and in Columns 4-6 of Table 3. The figure shows that events are associated with a discrete downward jump in the total revenue gradient. Though no individual coefficient is statistically significant in the non-parametric model, we decisively reject the hypothesis that all post-event effects are zero. The parametric model shows a fall in the gradient of about \$320 per pupil following an event, about one-third smaller than in the state revenue models, but this is statistically insignificant both individually and jointly with the post-event trend (Table 3).

We next repeat the quintile mean analyses for total revenues, in Figure 9, panels A-C, and in Table 4B. These are very similar to the state revenue results, and are much more precise than the total revenue slope analyses. Here, we find statistically significant increases of \$500 per pupil in relative total revenues in first quintile districts, with point estimates slightly larger than for state revenues. Again, this is about twice as is implied by the (insignificant) revenue-income slope results.

Our analyses thus far have focused on relative spending in high- and low-income districts. As discussed in Section I, a central concern in the school finance reform literature is whether reforms lead to voter revolts and ultimately to reductions in total educational spending. To assess this, we examine average state revenue and total revenue per pupil across all districts in the state, in Table 5. Average state revenues per pupil rise by about \$760 immediately following an event, with no sign of practically or statistically significant pre-event trends or phase-in effects. The increase in total revenues is somewhat smaller, around \$550, but equally sharp and also highly significant. (These results are shown graphically in Panel D of Figures 8 and 9.)

Taken together, our event study models indicate large increases in the progressivity of state and total revenues following finance reform events, driven by increases in low-income districts and with no sign of declines in high-income districts or in overall means. The income gradient and quintile mean analyses are broadly similar, though the quintile mean results suggest larger increases in progressivity. Average total revenues per pupil in first quintile districts are around $11,500, so the approximately $1000 average absolute increase that they see following an event represents a bit under 10% of their total revenues; the *relative* increase compared to higher income districts is about half as large.

Our estimated revenue impacts are notably larger than in the comparable specifications in Card and Payne's (2002) study of finance reforms in the 1980s, likely reflecting the extra "bite" of adequacy reforms in our sample. Card and Payne also estimate the impact of state aid on total revenues, using finance reforms as instruments for the former, and find that about $0.50 of each dollar of state aid "sticks." While our slope estimates are roughly consistent with this, our quintile analyses imply that a much larger share of the state aid increase persists in total revenues, perhaps in part because at least some adequacy reforms have involved state or judicial oversight of local property tax rates in addition to changes in the distribution of state aid.

V. Results: Student Outcomes

The above results establish that reform events are associated with sharp, immediate increases in the progressivity of school finance, with absolute and relative increases in revenues in low-income school districts. If additional funding is productive, we might expect to see impacts on student outcomes.

Figure 10 presents parametric and non-parametric event study estimates of the effect of reforms on the gradient of mean student test scores with respect to log mean income in the school district. The pattern of effects is notably different than in the finance analyses. There is no sign of an immediate effect here, but there is a clear change in the trend following reform events. The nonparametric estimates indicate a very smooth, very nearly linear decline in the test score gradient following an event, implying that relative scores in low-income districts rise gradually over many years. This is exactly the pattern one would expect, as 4th and 8th grade test scores are cumulative outcomes that presumably reflect inputs not only in the current grade but also inputs in earlier grades.

The pattern deviates from expectations in one respect, however: There is no indication that the phase-in of the effect slows five or nine years after the event, when the 4th and 8th graders, respectively, will have attended school solely in the post-event period. Our estimates of the out-year effects are imprecise, however, so we cannot rule out this sort of slowing.[24]

Estimates of the parametric model are presented in Table 6. As discussed in Section II.D, we treat each state-subject-grade-event combination as a separate panel (though standard errors are clustered at the state level to capture correlations among them). Columns 1-3 include state-event and subject-grade-year effects, while

---

[24] We observe outcomes *r* years after the event only for events in 2011-*r* and earlier. The resulting imbalance is partly offset by the increasing frequency of NAEP assessments over time (Table 1). Figure A1 in the Appendix shows the distribution of relative event time in our analytical sample. Samples are quite large for effects up to ten years out, but start to drop off thereafter.

columns 4-6 include state-subject-grade-event and year effects. This choice has little import for the results. There is no evidence of a pre-reform trend or a jump following events in any specification, so we focus on the models with just a phase-in effect, in Columns 1 and 4. These indicate that the test score-income gradient falls by about 0.009 per year after a reform event, for a total decline over ten years of 0.09.

Figure 11 and Table 7 repeat the test score analysis, this time using the gap in scores between first and fifth quintile districts. Results are quite similar: There is no immediate effect, but relative mean scores in first quintile districts begin to rise linearly following the event, accumulating to 0.07 standard deviations over ten years. Effects are driven by increases in low-income districts, with essentially no change in mean scores in high-income districts. Recall that the between-quintile gap in log mean incomes is about 0.6, so the 0.007 coefficient in Table 7 is quite consistent with the 0.009 coefficient in the test score slope model in Table 6.

The divergent time patterns of impacts on resources and on student outcomes, combined with the cumulative nature of the latter, prevents a simple instrumental variables interpretation of the reduced-form coefficients in terms of the achievement effect per dollar spent – it is not clear which years' revenues are relevant to the accumulated achievement of students tested *r* years after an event. In Section VIII we present estimates that divide the impact on student achievement ten years following an event by the impact on total discounted revenues over those ten years. The ten-year effect can be interpreted as the impact of a change in school resources for every year of a student's career (through 8th grade), an interpretation that is facilitated by the apparent lack of dynamics in the revenue effects.

Nevertheless, the calculation is somewhat arbitrary, as there is no sign that effects on grade g achievement stop growing g years after the events (perhaps because it takes time for districts to implement successful new programs after funding increases). We would obtain larger estimates of the achievement effect per dollar if we used estimates for more than ten years after events, or smaller effects if we used a shorter window.[25]

Table 8 presents estimates of the key coefficients from separate models by grade and subject, using the same specifications as Column 1 in Table 6 and Column 5 of Table 7. Effects are somewhat larger for math than for reading scores and for 4th than for 8th grade scores, but neither of these differences is statistically significant.

VI. Mechanisms

Our results thus far show that school finance reforms lead to substantial increases in relative revenues in low-income school districts, achieved through absolute increases in both high- and low-income districts that are larger in the former than the latter. Over time, they also lead to increases in the relative and absolute achievement of students in low-income districts. In an effort to understand the mechanisms through which increased revenues are translated into improved student outcomes, we analyze intermediate factors such as pupil-teacher ratios, teacher and student characteristics, and subcategories of spending.

First, we investigate student characteristics to determine whether changes to enrollment or the composition of the student body are likely to contribute to

[25] Another source of error in this calculation is the allocation of funding across grades. If it is evenly divided, roughly one-quarter of the funds are spent in grades 9-12, where they could not possibly impact 8th grade scores.

improvements in test scores. We estimate the same type of event-study analysis shown in Tables 3-4, but focusing on district demographic composition. Results are shown in Table 9. We find no evidence of effects of finance reform events on the share of students who are minority or low-income, either when examining gradients with respect to district income (first panel) or first-fifth quintile gaps (second panel). This suggests that compositional changes in the student body are not likely to be the mechanism for the rise in achievement.[26]

Other rows of Table 9 show proxies for classroom quality: The average pupil-teacher ratio and teacher salary. There are no significant effects on the distribution of these outcomes, either. Point estimates indicate reductions in the relative number of pupils per teacher in low-income districts, but these are quite imprecisely estimated.

Table 10 shows parallel results for components of spending. Total expenditures per pupil become discretely more progressive after a school finance reform event, though as with total revenues this is statistically significant only in the quintile analysis. When we divide spending into instructional and non-instructional components, only the non-instructional effect is robustly significant, and appears to account for about two-thirds of the total. Within this category, there is evidence of impacts on capital outlays and, less robustly, on student support services.[27] Neither of these is obvious as the most efficient route to increased learning, but neither is it

[26] Appendix Table 3 shows no relationship between the change in district income between 1990 and 2011 and the school finance reform events.

[27] Many of the court cases in our event database specifically concern inadequacy of school facilities in poor school districts, so it is not surprising that plaintiff victories lead to increases in this category.

implausible that either could be productive (see, e.g., Cellini et al., 2010; Martorell, Stange, and McFarlin, 2015, and Neilson and Zimmerman 2014).

Our research design is poorly suited to identifying the optimal allocation of school resources across expenditure categories, or to testing whether actual allocations are close to optimal. It is possible that the achievement effects would have been much larger had districts spent their extra revenues in some other way. The most that we can say is that the average finance reform – which we interpret to involve roughly unconstrained increases in resources, though in some cases the additional funds were earmarked for particular programs or tied to other reforms – led to a productive, though perhaps not maximally productive, use of the funds.[28]

VII. Effects on Achievement Gaps

The final question that we investigate is whether finance reforms closed overall test score gaps between high- and low-achieving, minority and white, or low-income and non-low-income students in a state. These are perhaps better measures of the overall effectiveness of a state's educational system at delivering equitable, adequate services to disadvantaged students. However, because only a small portion of income or other inequality is across rather than within districts, school finance reforms that target the distribution of resources across districts may not be well enough targeted to meaningfully close these gaps.

Table 11 presents estimates of effects on mean test scores across different subgroups of interest. The first panel shows small and insignificant effects on mean

[28] Stronger school accountability may provide incentives to schools to allocate their resources more efficiently (Hanushek 2006). We investigated specifications that allowed for interactions between finance reform events and the state's accountability policy, but found no evidence for this.

(pooled) test score and on the 25th and 75th percentiles of the state distributions. The absence of a mean score effect is somewhat of a puzzle, given the increases in mean revenues documented earlier. It must be noted, however, that our research design is more credible for *disparities* in outcomes than for the *level* of outcomes, as the latter would be confounded by unobserved shocks to average outcomes in a state that are correlated with the timing of school finance reforms.

The second and third panels present results for mean scores by race and free lunch status, respectively. There is no discernible effect on mean scores for any group, or on achievement gaps defined by race or free-lunch status. Point estimates are roughly one full order of magnitude smaller than the earlier estimates for first-quintile district mean scores.

Appendix Tables A5 and A6 resolve the discrepancy. While non-white and free lunch students are more likely than their white and non-free-lunch peers to attend school in low-income school districts, the differences are not very large. Roughly one-quarter of non-white students, and 30% of free lunch students, live in first quintile districts, while the shares in fifth quintile districts are about half as large. This suggests that finance reforms may not have much effect on the relative resources to which the typical minority or low-income student is exposed.

To assess this more carefully, we assigned each student the mean revenues for the district that he/she attends, and estimated event study models for the black-white or free lunch/no free lunch gap in these revenues. Results, reported in Appendix Table A6, indicate that finance events raise relative per-pupil revenues in the average black student's school district by only \$220 (S.E. 166) and in the average free lunch student's district by only $79 (S.E. 166). Even if this funding was *more* productive than the average effect implied by our pooled analysis, it would still not be enough to yield detectable effects on black or free lunch students' average test scores. Thus, while reforms aimed at ameliorating resource shortages in low-income districts appear to have been successful at raising resources and outcomes in these districts, we conclude that within-district changes would be necessary to have a meaningful impact on the average low-income or minority *student*.

VIII. Conclusion

After school desegregation, school finance reform is perhaps the most important education policy change in the United States in the last half century. But while the effects of the first- and second-wave reforms on school finance have been well studied, there is little evidence about the finance effects of third-wave, "adequacy" reforms or about the effects of any of these reforms on student achievement. Our study presents new evidence on each of these questions.

We find that state-level school finance reforms enacted during the adequacy reform era markedly increased the progressivity of school spending. They did not accomplish this by "leveling down" school funding, but rather by increasing spending across the board, with larger increases in low-income districts. Although we cannot rule out the possibility that a portion of this funding was offset through local decisions, much or all of it "stuck," leading to appreciable increases in spending in low-income school districts. Using nationally representative data on student achievement, we find that this spending was productive: The reforms also led to increases in the absolute and relative achievement of students in low-income districts. Our estimates thus complement the analysis of Jackson et al. (forthcoming), who examine the long-run impacts of earlier school finance reforms and find substantial positive impacts on a variety of long-run outcomes.

To put our results into context, consider the implied effect of an average-sized reform on a district with log average income one point below the state mean, relative to a district at the mean. According to our estimates, the reform raised relative state revenue per pupil in the former district by $500 immediately, an effect that persisted for many years. Relative total revenues rose by about $320, again immediately and persistently. Over the following years, relative test scores rose as well, cumulating to a 0.09 standard deviation impact in the tenth year after the reform event that if anything continued to grow thereafter.

The cost-effectiveness of these reforms can be assessed by comparing the finance effects to the achievement effects. To do so, we assume that finance effects are uniform over time. $320 per pupil in spending each year of a student's career, discounted to the student's kindergarten year using a 3% rate, corresponds to a present discounted cost of $3505. Chetty et al. (2011) estimate that a 0.1 standard deviation increase in kindergarten test scores translates into increased earnings in adulthood with present value of $5,350 per pupil. Our ten-year reform effect estimates thus imply that the additional spending yields increased earnings of $4,815 per pupil, implying a benefit-to-cost ratio of nearly 1.4.

This ratio is not wholly robust. Our quintile analysis shows larger revenue effects, implying a benefit-cost ratio below one. Note, however, that these comparisons count only 4^{th} and 8^{th} grade test score increases as benefits, while counting as costs expenditures in all grades (including 9-12). This biases the benefit-cost ratio downward. Another downward bias comes from our use of earnings effects of kindergarten test scores to value increases in 8^{th} grade test scores, which are presumably better proxies for adult earnings. Jackson et al.'s (forthcoming) analysis of the effects of earlier finance reforms on students' adult outcomes implies much larger benefits per dollar than does our calculation. Thus, although these sorts of calculations are quite imprecise, the evidence appears to indicate that the spending enabled by finance reforms was cost-effective, even without accounting for beneficial distributional effects.

Our results thus show that money can and does matter in education, and complement similar results for the long-run impacts of school finance reforms from Jackson et al. (forthcoming). School finance reforms are blunt tools, and some critics (Hanushek, 2006; Hoxby, 2001) have argued that the funds that they provide will be offset by changes in district or voter choices over tax rates or will be spent so inefficiently as to be wasted. Our results do not support these claims. Evidently, it is possible for courts to force improvements in school quality for students in low-income districts.

But there is an important caveat to this conclusion. As we discuss in Section VII, the average low-income student does not live in a particularly low-income district, so is not very well targeted by a transfer of resources to the latter. Thus, we find that finance reforms reduced achievement gaps between high- and low-income school districts but did not have detectable effects on the resource or achievement gaps between average high- and low-income (or white and black) students. Attacking these gaps via school finance policies would require changing the allocation of resources *within* school districts, something that was not attempted by the reforms that we study.

References

Burtless, G. (1996). *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: Brookings Institution Press.

Card, D., & Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, *100*(1), 1–40.

Card, D., & Krueger, A.B. (1992). School quality and black-white relative earnings: A direct assessment. *Quarterly Journal of Economics*, 107(1), 151-200.

Card, D., & Payne, A.A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, *83*(1), 49-82.

Cascio, E. U., Gordon, N., & Reber, S. (2013). Local responses to federal grants: evidence from the introduction of title I in the South. *American Economic Journal: Economic Policy*, *5*(3), 126-159.

Cascio, E. U., & Reber, S. (2013). The Poverty Gap in School Spending Following the Introduction of Title I. *American Economic Review*, *103*(3), 423-427.

Cellini, S., F. Ferreira, & J. Rothstein (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *Quarterly Journal of Economics 125*(1), 215-261.

Chaudhary, L. (2009). Education inputs, student performance and school finance reform in Michigan. *Economics of Education Review*, *28*(1), 90-98.

Clark, M. A. (2003). Education reform, redistribution, and student achievement: Evidence from the Kentucky Education Reform Act. Unpublished working paper, Mathematica Policy Research, Princeton, NJ.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. (1966). *Equality of educational opportunity.* Washington, DC, 1066-5684.

Coons, J.E., Clune, W.H., & Sugarman, S. (1970). *Private Wealth and Public Education*. Cambridge, MA: Belknap Press.

Corcoran, S. P., & Evans, W. N. (2008). Equity, Adequacy, and the Evolving State Role in Education Finance. In H.F. Ladd and E.B. Fiske, eds., *Handbook of Research in Education Finance and Policy*. New York: Routledge.

Corcoran, S., Evans, W. N., Godwin, J., Murray, S. E., & Schwab, R. M. (2004). The changing distribution of education finance, 1972–1997. *Social inequality*, 433-465.

Cullen, J. B. and Loeb, S. (2004). School finance reform in Michigan: evaluating Proposal A. In J. Yinger, ed., *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, pp. 215-50, Cambridge, MA: MIT Press.

Downes, T. and L. Stiefel (2008). Measuring equity and adequacy in school finance. In Ladd, H. F., & Fiske, E. B., eds., *Handbook of Research in Education Finance and Policy*. New York, NY: Routledge.

Duncombe, W.D., P Nguyen-Hoang, and J. Yinger (2008). Measurement of cost differentials. In Ladd, H. F., & Fiske, E. B., eds., *Handbook of Research in Education Finance and Policy*. New York, NY: Routledge.

Fischel, W.A. (1989). Did *Serrano* cause Proposition 13? *National Tax Journal* 42(4): 465-73.

Flanagan, A. E., and Murray, S. E. (2004). A Decade of Reform: The Impact of School Reform in Kentucky. In John Yinger, (Ed.), *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity* (pp. 165-213). Cambridge, MA: MIT Press.

Guryan, J. (2001). Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts. National Bureau of Economic Research Working Paper No. w8269.

Hanushek, E.A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113, F64-F98.

Hanushek, E.A. (2006). School resources. In Hanushek, E.A. and F. Welch, eds., *Handbook of the Economics of Education*, vol. 2. Elsevier.

Hanushek, E.A. & Lindseth, A.A. (2009). *Schoolhouses, Courthouses and Statehouses: Solving the Funding-Achievement Puzzle in America's Public Schools*. Princeton: Princeton University Press.

Hanushek, E. A., Rivkin, S. G., & Taylor, L. L. (1996a). Aggregation and the estimated effects of school resources. *The Review of Economics and Statistics* 78(4), 611-627.

Hanushek, E. A., Rivkin, S. G., & Taylor, L. L. (1996b). The identification of school resource effects. *Education Economics*, *4*(2), 105-125.

Horowitz, H. (1966). Unseparate but unequal: The emerging Fourteenth Amendment issue in public school education. *UCLA Law Review*, 13, 1147-1172.

Hoxby, C. M. (2001). All school finance equalizations are not created equal. *The Quarterly Journal of Economics*, *116*(4), 1189-1231.

Hyman, J. (2013). Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment. Unpublished manuscript.

Jackson, C.K., Johnson, R.C., & Persico, C. (forthcoming). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. Forthcoming, *Quarterly Journal of Economics*.

Kirp, D.L. (1968). The poor, the schools, and equal protection. *Harvard Educational Review* 38, 635-668.

Koski, W.S., & Hahnel, J. (2008). The past, present and future of educational finance reform litigation. In H.F. Ladd and E.B. Fiske, eds., *Handbook of Research in Education Finance and Policy*. New York: Routledge.

Krueger, A.B. (2003). Economic considerations and class size. *The Economic Journal* 113, F34-F63.

Krueger, A.B., & Whitmore, D.M. (2002). Would smaller classes help close the black-white achievement gap? In John E. Chubb and Tom Loveless, eds., *Bridging the Achievement Gap*. Washington: Brookings Institution Press.

Ladd, H. F., & Fiske, E. B. (Eds.). (2008). *Handbook of Research in Education Finance and Policy*. New York, NY: Routledge.

Martorell, P., Stange, K. M., & McFarlin, I. (2015). Investing in schools: Capital spending, facility conditions, and student achievement. NBER Working Paper 21515, September.

Murray, S. E., Evans, W. N., & Schwab, R. M. (1998). Education-finance reform and the distribution of education resources. *American Economic Review, 88*(4), 789-812.

Nielson, C., & Zimmerman, S. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics 120*.

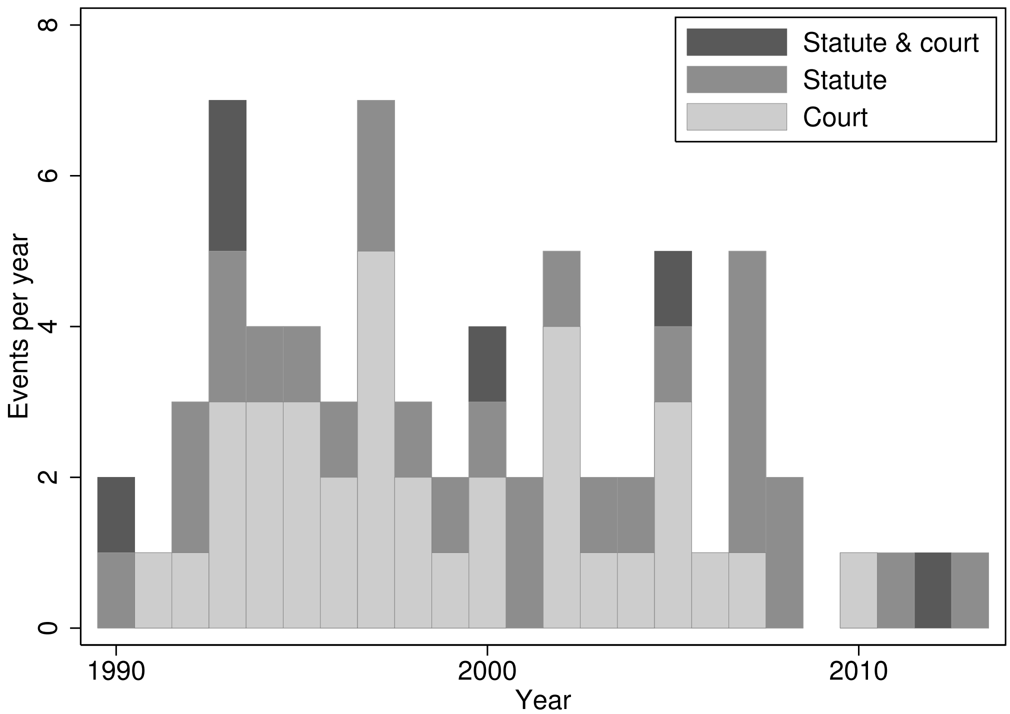
Papke, L. (2005). The Effects of Spending on Test Pass Rates: Evidence from Michigan. *Journal of Public Economics* 89(5), 821-839.

Papke, L. (2008). The Effects of Changes in Michigan's School Finance System. *Public Finance Review* 36(4), 456-474

Wise, A. (1967). *Rich Schools, Poor Schools: The Promise of Equal Educational Opportunity*. Chicago, IL: University of Chicago Press.

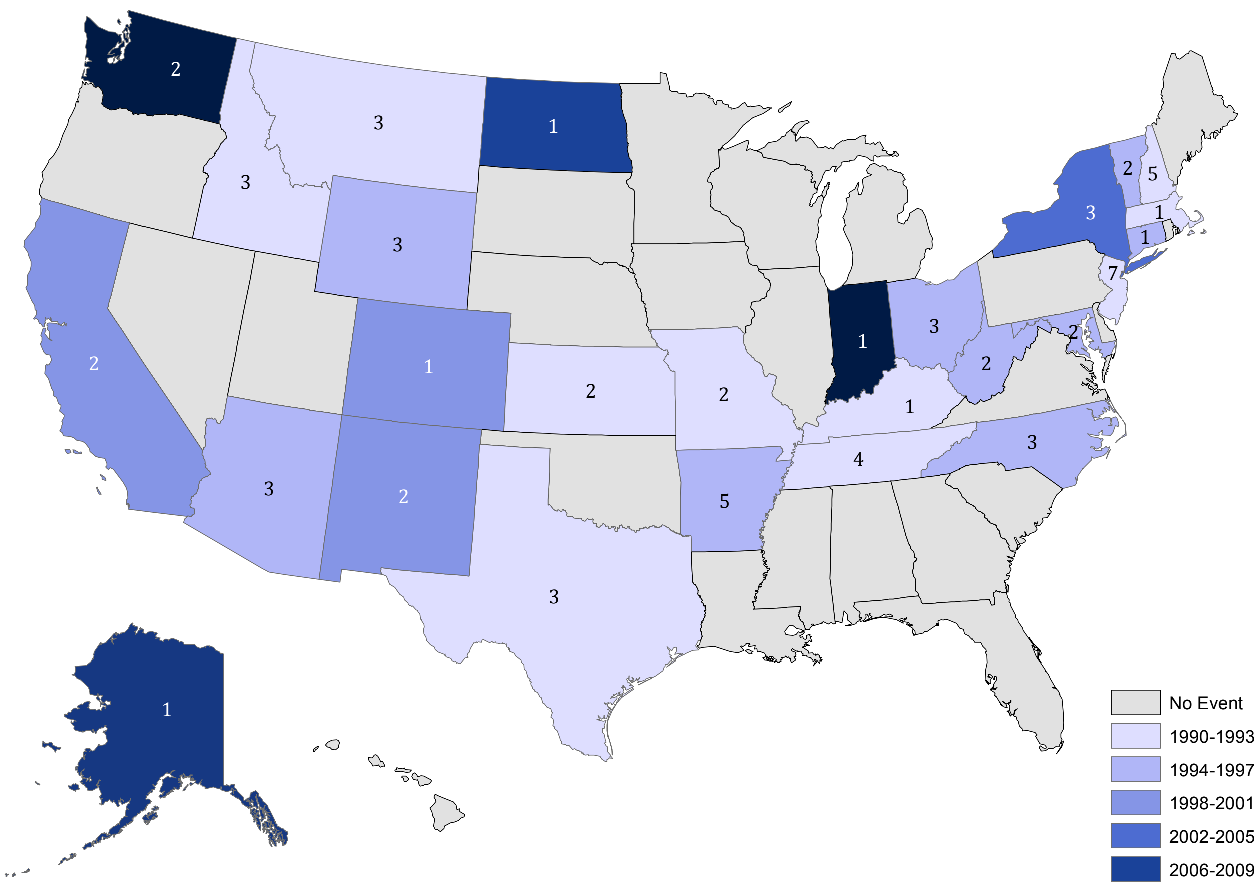
1 Figures

Figure 1: Timing of school finance events



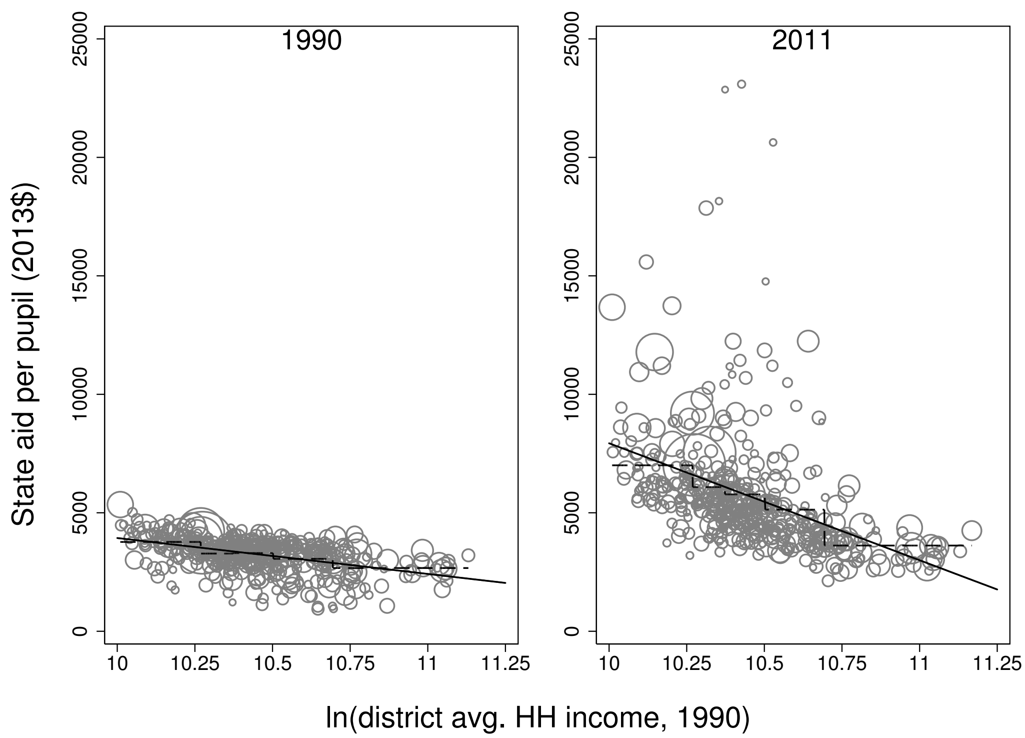
Notes: When multiple events occur in a state in a given year, they are combined into a single event for this chart.

Figure 2: Geographic distribution of post-1989 school finance events



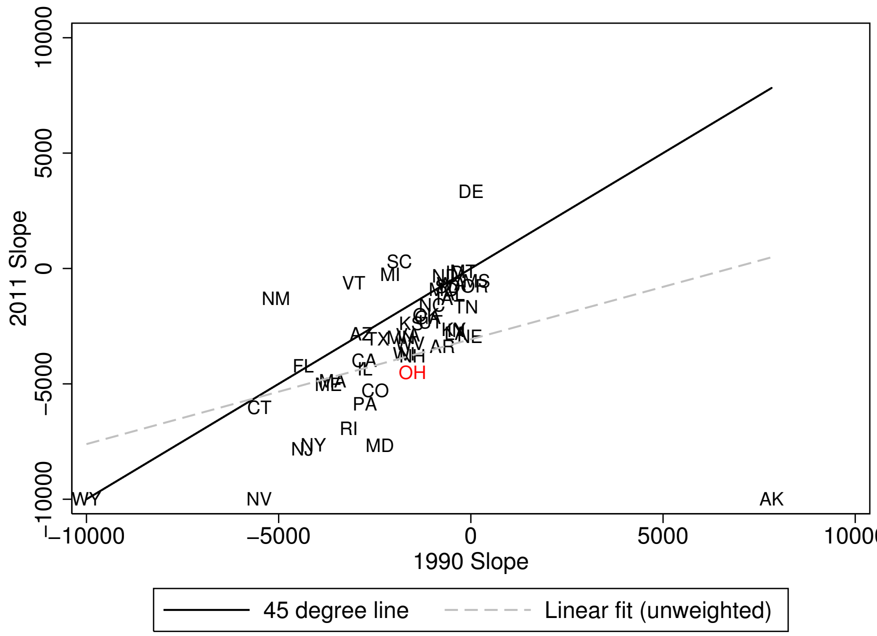
Notes: Colors correspond to the date of the first post-1989 school finance event. Numbers indicate the number of events in that period.

Figure 3: State aid vs. district income, Ohio, 1990 and 2011

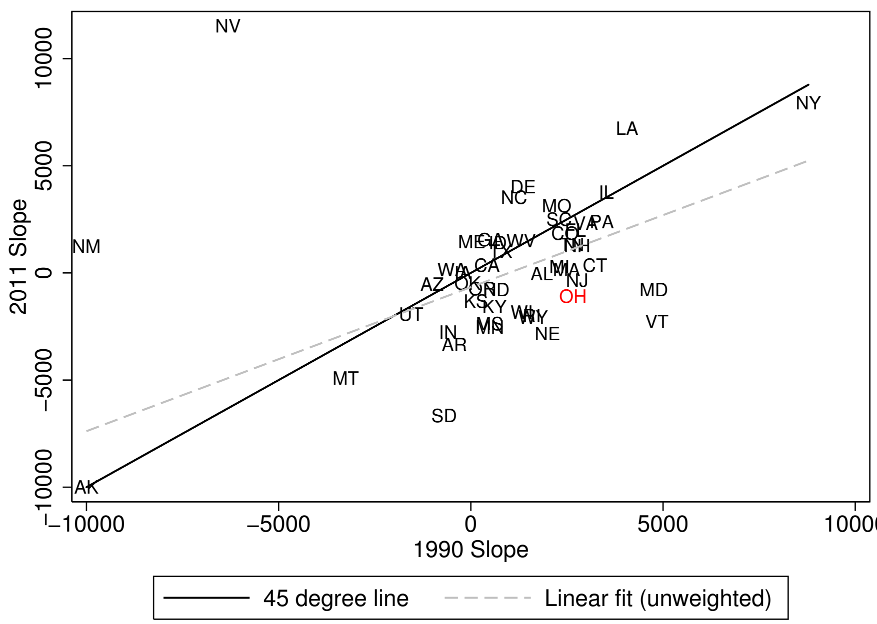


Notes: Each point represents one district. Circle sizes are proportional to average district enrollment over 1990-2011. Solid lines have slope equal to θ_{st} from equation (1) and correspond to predicted values for a unified district of average log enrollment. Dashed lines represent means among districts in each quintile of the district mean income distribution.

Figure 4: State-level slopes of school finance with respect to ln(district income), 1990 and 2011



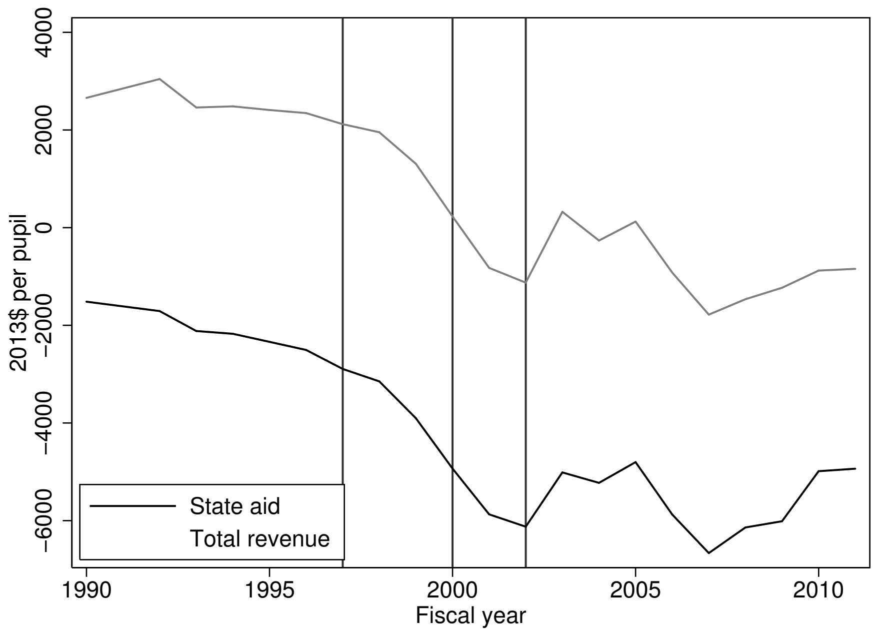
(a) State revenue per pupil



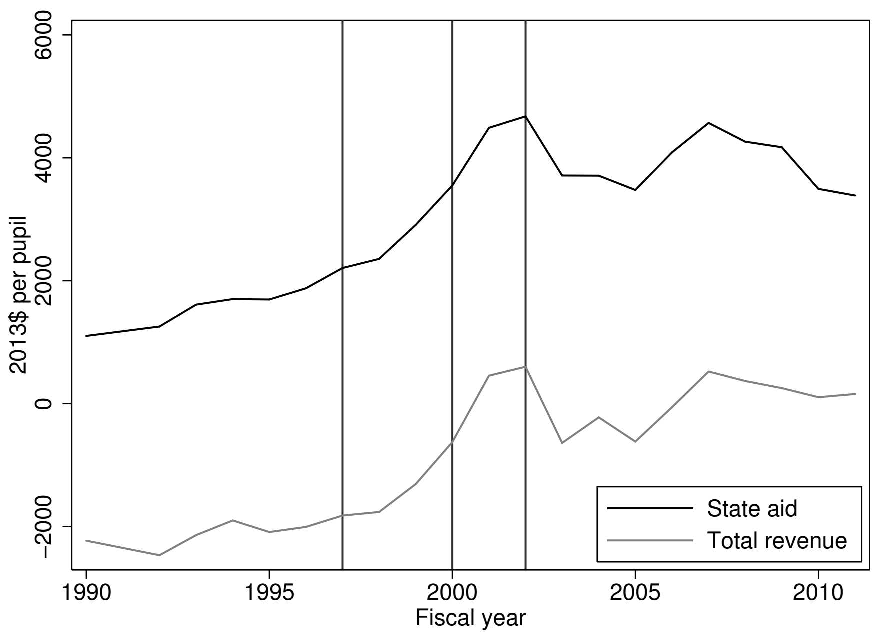
(b) Total revenue per pupil

Notes: Points indicate θ_{st} for t = 1990, 2011. Slopes are censored below at -10,000 for graphical display, but uncensored values are used in computing the (unweighted) linear fit.

Figure 5: Summaries of school finance in Ohio, 1990-2011



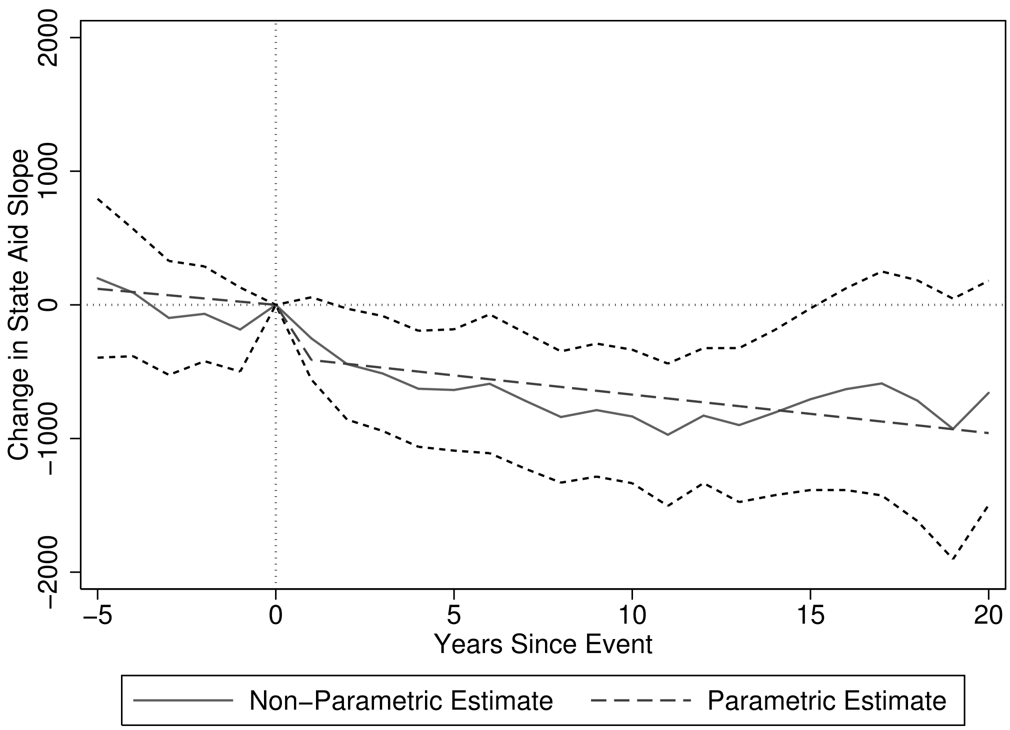
(a) Log income gradients



(b) Mean difference between 1st and 5th quintile of district mean log income

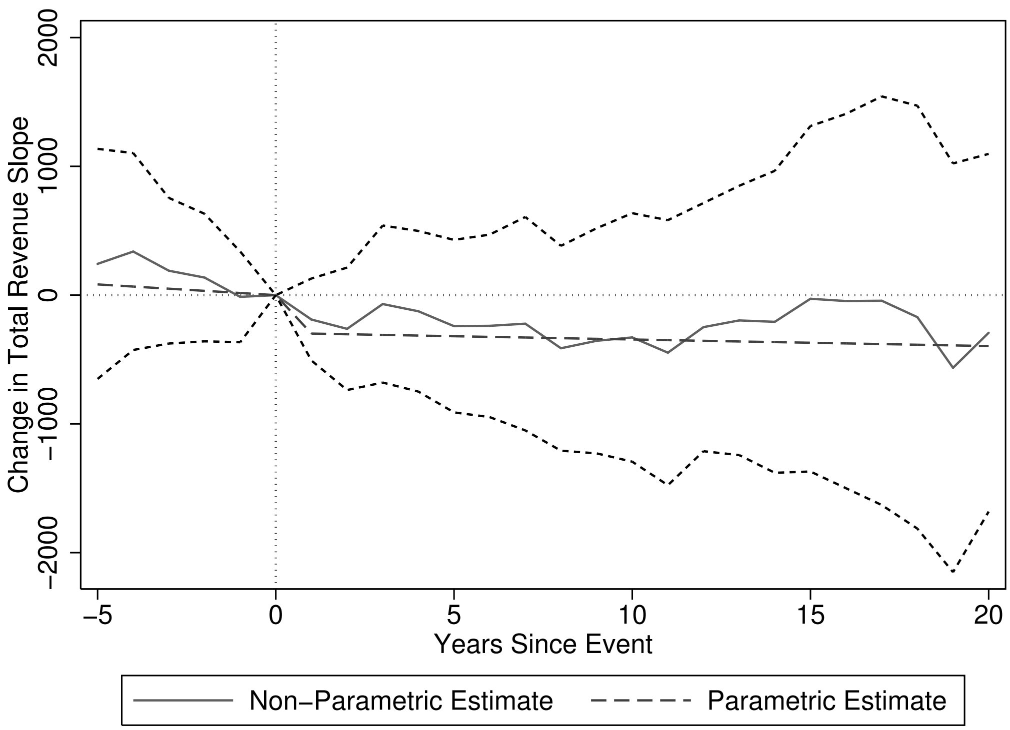
Notes: In panel (a), series represent θ_{st} from equation (1), varying the dependent variable, with 95% confidence intervals. In panel (b), series are the difference in the mean of the relevant revenue variable between districts in the first and fifth quintiles of the district mean income distribution. Solid vertical lines represent plaintiff victories in the Ohio Supreme Court in *De Rolph v state I, II, and IV* in 1997, 2000, and 2002. In 2000 there was also a statutory reform.

Figure 6: Event study estimates of effects of reform events on state revenue slope



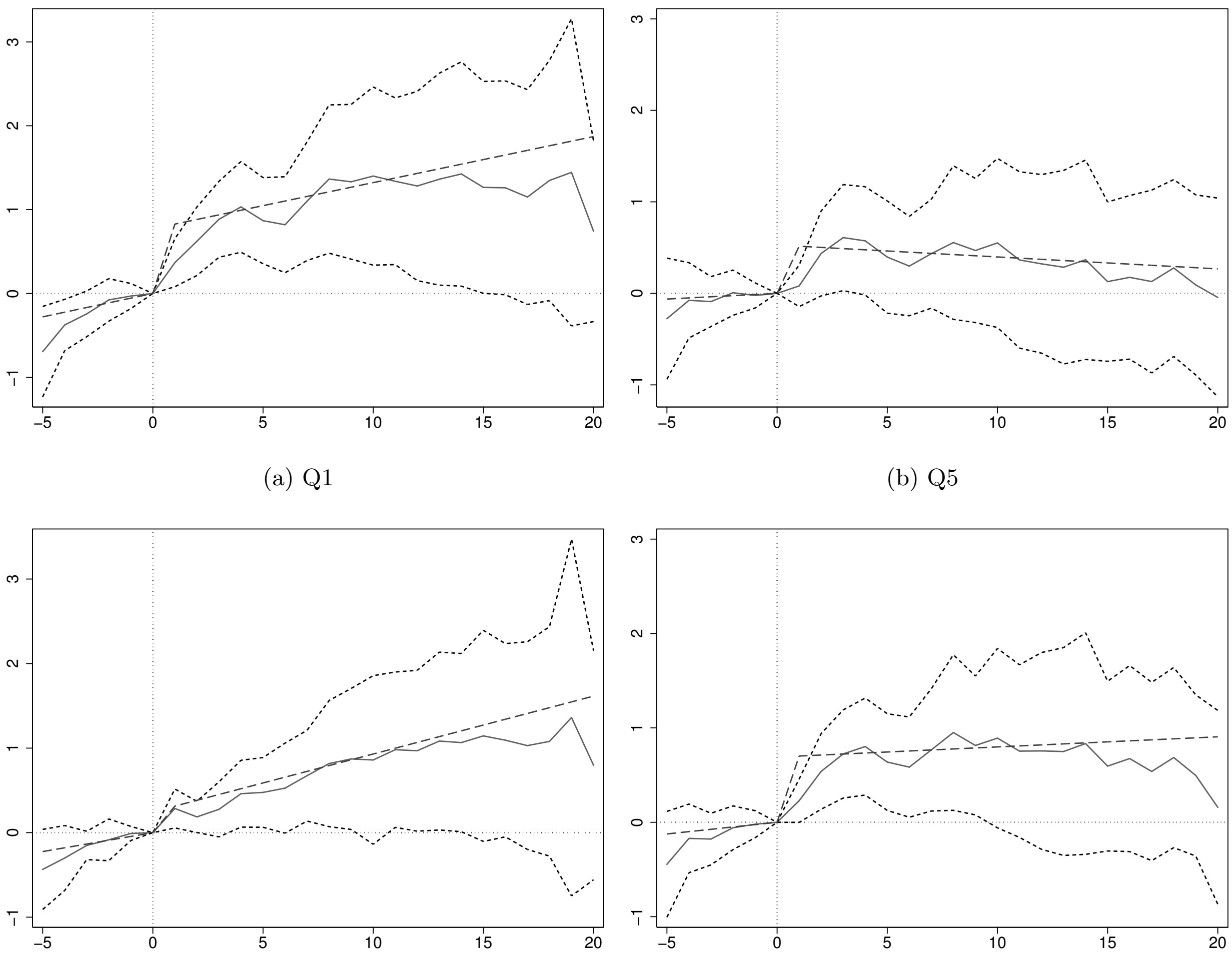
Notes: Dependent variable is the slope of state revenue per pupil with respect to ln(district income) in the state-year cell. Figure shows parametric and non-parametric estimates of the effect of a finance event on this slope, by years since (or prior to) the event, along with 95% confidence intervals for the non-parametric model. See text for specification. The null hypothesis that all post-event coefficients in the non-parametric model are zero is rejected (p<0.001). Estimates for the parametric model are reported in Table 3, Column 3.

Figure 7: Event study estimates of effects of reform events on total revenue slope



Notes: Dependent variable is the slope of total revenue per pupil with respect to ln(district income) in the state-year cell. Figure shows parametric and non-parametric estimates of the effect of a finance event on this slope, by years since (or prior to) the event, along with 95% confidence intervals for the non-parametric model. See text for specification. The null hypothesis that all post-event coefficients in the non-parametric model are zero is rejected (p<0.001). Estimates for the parametric model are reported in Table 3, Column 6.

Figure 8: Event study estimates of effects of reform events on mean state revenues per pupil by district income quintile



(a) Q1

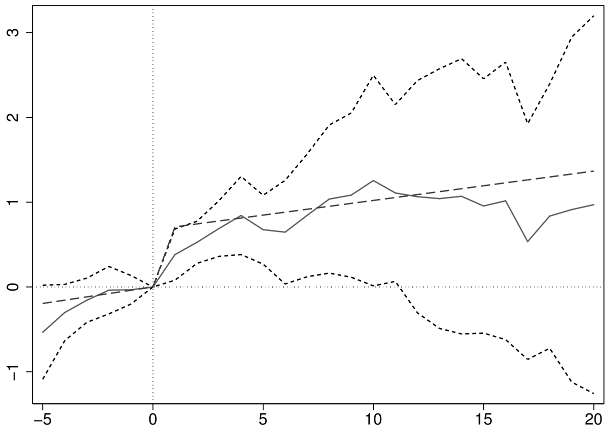
(b) Q5

(c) Q1 - Q5

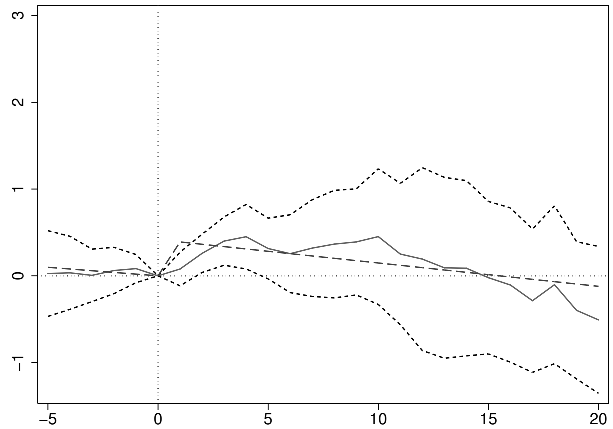
(d) Overall

Notes: Dependent variable is mean state aid per pupil in the relevant subgroup of districts. In Panels A and B, the mean is for districts in the bottom fifth and top fifth, respectively, of the district mean income distribution (unweighted). In Panel C, the dependent variable is the difference between these. All districts are included in the mean in panel D. See text for event study specifications. In the non-parametric specifications, the null hypothesis that all post-event effects equal zero is rejected in each panel. In the parametric specifications, the post-event jump coefficient is significantly different from zero in each panel (though the null hypothesis that the jump and the change in trend are jointly zero is not rejected in panels C and D). Estimates for parametric models are reported in panel a of Table 4.

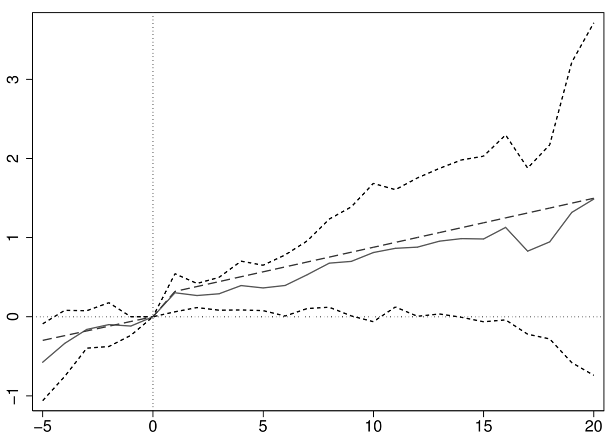
Figure 9: Event study estimates of effects of reform events on mean total revenues per pupil by district income quintile



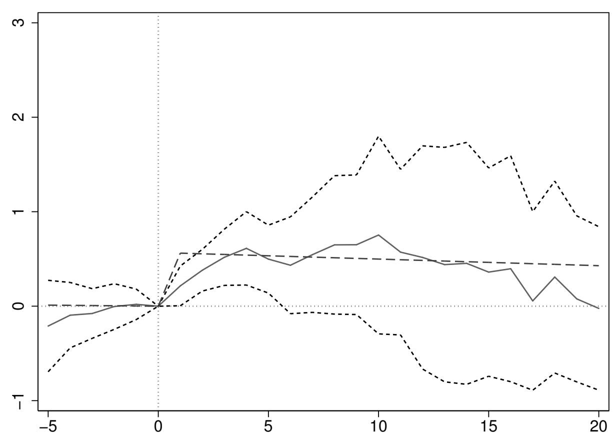
(a) Q1



(b) Q5



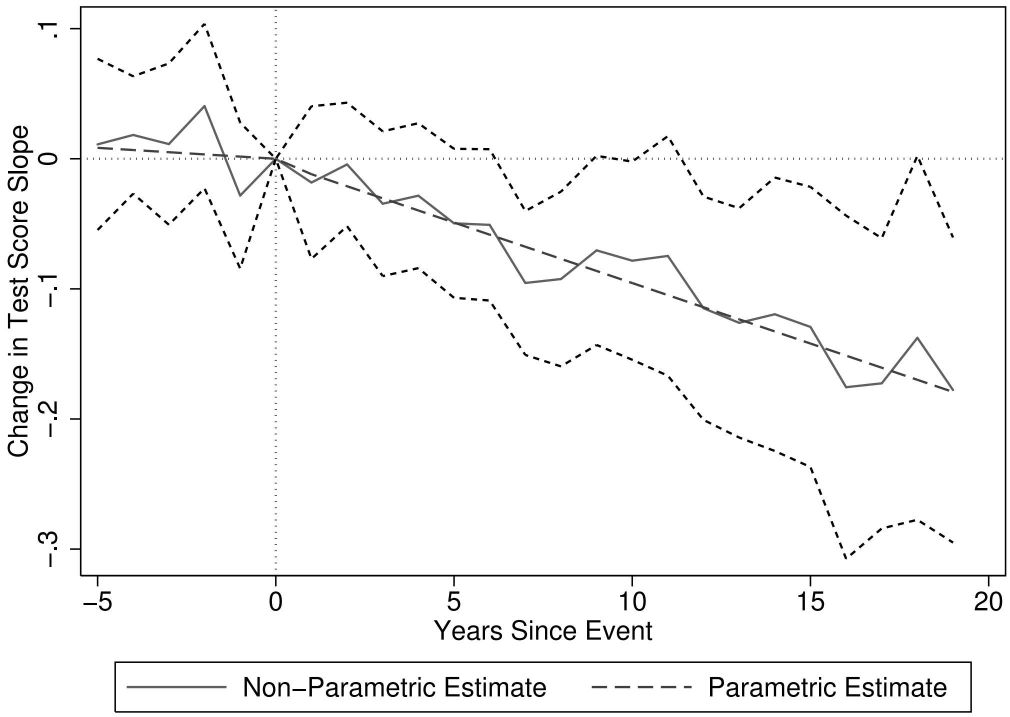
(c) Q1 - Q5



(d) Overall

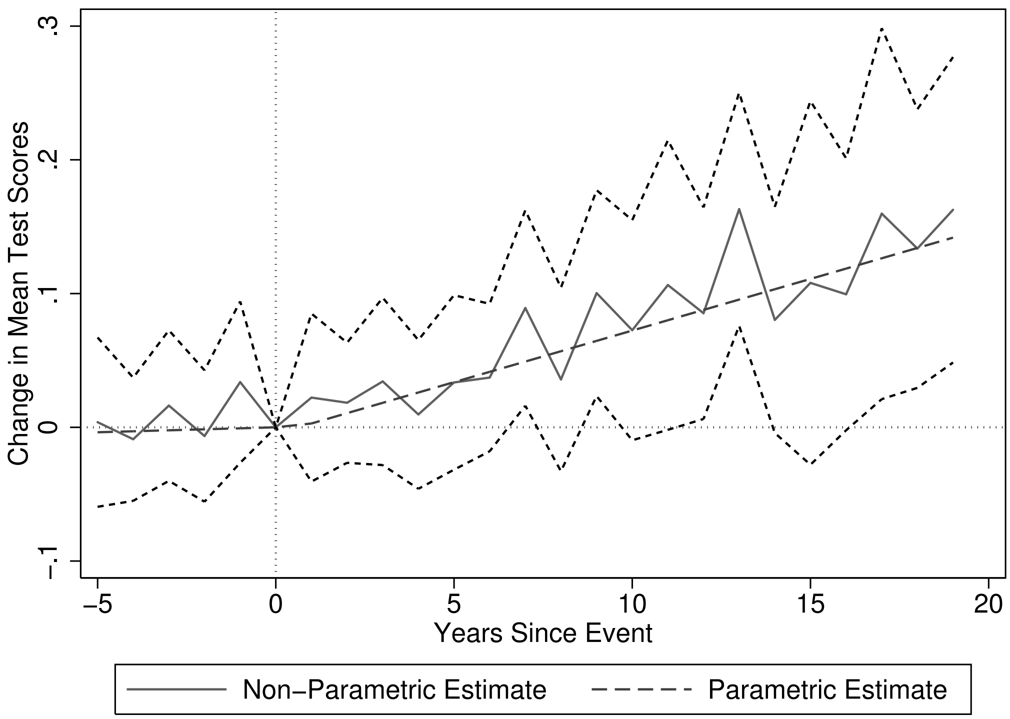
Notes: Dependent variable is mean total revenues per pupil in the relevant subgroup of districts. In Panels A and B, the mean is for districts in the bottom fifth and top fifth, respectively, of the district mean income distribution (unweighted). In Panel C, the dependent variable is the difference between these. All districts are included in the mean in panel D. See text for event study specifications. In the non-parametric specifications, the null hypothesis that all post-event effects equal zero is rejected in each panel. In the parametric specifications, the post-event jump coefficient is significantly different from zero in each panel. Estimates for parametric models are reported in panel b of Table 4.

Figure 10: Event study estimates of effects of reform events on test score slope



Notes: Dependent variable is the slope of district-level mean NAEP test scores (in student-level standard deviation units) with respect to ln(district income) in the state-year cell. Figure shows parametric and non-parametric estimates of the effect of a finance event on this slope, by years since (or prior to) the event, along with 95% confidence intervals for the non-parametric model. See text for specification. The null hypothesis that all post-event coefficients in the non-parametric model are zero is rejected (p<0.001). Estimates for the parametric model are reported in Table 6, Column 3.

Figure 11: Event study estimates of effects of Q1-Q5 difference in mean scores



Notes: Dependent variable is difference in mean NAEP test scores (in student-level standard deviation units) between the first and fifth quintiles with respect to ln(district income) in the state-year cell. Figure shows parametric and non-parametric estimates of the effect of a finance event on this slope, by years since (or prior to) the event, along with 95% confidence intervals for the non-parametric model. See text for specification. The null hypothesis that all post-event coefficients in the non-parametric model are zero is rejected ($p<0.001$). Estimates for the parametric model are reported in Table 7, Column 6.

2 Tables

Table 1: NAEP Testing Years

Year	Subject(s)	Grade(s)	Number of States	Number of Students Tested
1990	Math	G8	38	97900
1992	Math, Reading	G4, G8	42	321118
1994	Reading	G4	41	104890
1996	Math	G4, G8	45	228978
1998	Reading	G4, G8	41	206812
2000	Math	G4, G8	42	201106
2002	Reading	G4, G8	51	270230
2003	Math, Reading	G4, G8	51	691359
2005	Math, Reading	G4, G8	51	674416
2007	Math, Reading	G4, G8	51	711360
2009	Math, Reading	G4, G8	51	775062
2011	Math, Reading	G4, G8	51	749246

Table 2: Summary statistics

(a) District-Year Panel

	mean	sd	N
Total revenue per pupil	\$10,979	(3,376)	208207
State revenue per pupil	\$5,155	(2,234)	208207
Local revenue per pupil	\$4,971	(3,184)	208207
Federal revenue per pupil	\$853	(625)	208207
Log(Mean income) - 1990	10.51	(0.27)	208207
Unfied district	0.93	(0.25)	208207
Elementary district	0.05	(0.21)	208207
Secondary district	0.02	(0.14)	208207
Total expenditure per pupil	\$11,149	(3,582)	208212
Total instructional expenditure per pupil	\$5,804	(1,915)	208212
Total non-instructional expenditure per pupil	\$5,346	(2,151)	208212
Enrollment (student weighted)	70,973	(188,868)	208207
Enrollment (unweighted)	4,006	(16,378.2)	208207

(b) State-Year Panel

	mean	sd	N
State revenue slope	-3163.72	(3512.17)	4116
Total revenue slope	326.34	(3665.60)	4116
Test score slope	0.95	(0.36)	1498
Dist income Q1 mean: state revenue	\$6,430	(2,856)	4264
Dist income Q1 mean: total revenue	\$11,462	(3,798)	4264
Dist income Q5 mean: state revenue	\$4,410	(2,278)	4256
Dist income Q5 mean: total revenue	\$11,554	(3,358)	4256
Dist income Q1-Q5 mean: state revenue	\$2,012	(2,094)	4256
Dist income Q1-Q5 mean: total revenue	\$-103	(2,028)	4256
Dist income Q1 mean: test scores	0.08	(0.37)	1573
Dist income Q5 mean: test scores	0.48	(0.41)	1571
Dist income Q1-Q5 mean: test scores	-0.40	(0.30)	1568
Num events to Date	0.77	(1.29)	5100

Table 3: Event study estimates for slopes of state revenue and total revenue with respect to ln(district income)

	St. Rev.	St. Rev.	St. Rev.	Tot. Rev.	Tot. Rev.	Tot. Rev.
Post Event	-501.4**	-441.5**	-383.9**	-321.2	-327.4	-293.7
	(187.6)	(180.0)	(153.8)	(285.1)	(270.4)	(228.1)
Post Event * Yrs Elapsed		-17.97	-4.760		2.178	11.54
		(16.93)	(19.25)		(36.23)	(40.18)
Trend			-24.00			-16.63
			(27.72)			(39.90)
Observations	1890	1890	1890	1890	1890	1890
p, total event effect=0	0.010	0.032	0.034	0.266	0.486	0.438
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: In columns 1-3, the dependent variable is the slope of state revenue per pupil with respect to ln(district income) in the state-year cell. In columns 4-6 the dependent variable is the slope of total revenue per pupil with respect to ln(district income). Regressions are weighted by the inverse of the estimated sampling variance of the dependent variable. See text for further specification details. P-values from the joint hypothesis test that all after-event coefficients equal zero are shown. Standard errors are clustered at the state level.

Table 4: Event study estimates for mean state revenue and total revenues per pupil by district income quintile

(a) State Revenue

	Q1	Q1	Q5	Q5	Q1-Q5	Q1-Q5
Post Event	1022.7***	772.8***	510.2	528.1**	517.6**	245.7**
	(279.9)	(249.4)	(328.6)	(255.9)	(210.5)	(119.3)
Post Event * Yrs Elapsed		-0.817		-25.48		23.73
		(46.53)		(23.81)		(34.61)
Trend		55.74		12.44		44.76
		(36.27)		(32.51)		(28.04)
Observations	1927	1927	1924	1924	1924	1924
p, total event effect=0	0.001	0.008	0.127	0.109	0.017	0.091

(b) Total Revenue

	Q1	Q1	Q5	Q5	Q1-Q5	Q1-Q5
Post Event	838.2***	674.7***	307.3	417.5**	534.9***	258.0**
	(236.8)	(209.8)	(221.0)	(193.1)	(179.5)	(123.0)
Post Event * Yrs Elapsed		-4.259		-7.270		2.308
		(58.69)		(31.20)		(38.73)
Trend		38.88		-19.67		59.65*
		(39.71)		(25.70)		(30.92)
Observations	1927	1927	1924	1924	1924	1924
p, total event effect=0	0.001	0.005	0.170	0.099	0.004	0.118

Notes: The dependent variables are mean state revenue and total revenues per pupil in the in the relevant district income quintile. All specifications include state-event and year fixed effects. Regressions are unweighted. See text for further specification details. P-values from the joint hypothesis test that all after-event coefficients equal zero are shown. Standard errors are clustered at the state level.

Table 5: Event study estimates for mean state aid per pupil and mean total revenues per pupil

	St. Rev.	St. Rev.	St. Rev.	Tot. Rev.	Tot. Rev.	Tot. Rev.
Post Event	762.3**	760.1***	691.1***	544.6**	562.4**	568.6***
	(297.7)	(277.1)	(240.1)	(221.5)	(212.6)	(189.4)
Post Event * Yrs Elapsed		0.749	-14.04		-6.079	-4.732
		(28.99)	(31.69)		(38.73)	(42.26)
Trend			24.77			-2.256
			(31.33)			(29.72)
Observations	1927	1927	1927	1927	1927	1927
p, total event effect=0	0.014	0.029	0.021	0.017	0.036	0.014
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: In columns 1-3, the dependent variable is mean state aid per pupil in the state-year cell. In columns 4-6 the dependent variable is mean total revenues per pupil. See text for further specification details. P-values from the joint hypothesis test that all after-event coefficients equal zero are shown. Standard errors are clustered at the state level.

Table 6: Event study estimates for test score slopes

Post Event * Yrs Elapsed	-0.00882***	-0.00863**	-0.00762**	-0.00875**	-0.00864**	-0.00711*
	(0.00313)	(0.00324)	(0.00369)	(0.00357)	(0.00367)	(0.00419)
Post Event		-0.00707	-0.00253		-0.00410	0.00255
		(0.0187)	(0.0143)		(0.0211)	(0.0168)
Trend			-0.00168			-0.00253
			(0.00365)			(0.00388)
Observations	2743	2743	2743	2743	2743	2743
p, total event effect=0	0.00700	0.0210	0.0555	0.0180	0.0546	0.205
State-Event FEs	X	X	X			
St-Ev-Gr-Sub FEs				X	X	X
Year FEs				X	X	X
Sub-Gr-Yr FEs	X	X	X			

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: Dependent variable is the slope of district-level mean NAEP test scores (in student-level standard deviation units) with respect to ln(district income) in the state-year cell. Columns 1-3 show estimates with state-copy fixed effects and NAEP exam fixed effects (i.e. subject-grade-year). Columns 4-6 show estimates with joint state-copy-grade-subject fixed effects, and separate year effects. Regressions are weighted by the inverse of the estimated sampling variance of the dependent variable. See text for further specification details. P-values from the joint hypothesis test that all after-event coefficients equal zero are shown. Standard errors are clustered at the state level.

Table 7: Event studies for mean subgroup scores

	Q1	Q1	Q5	Q5	Q1-Q5	Q1-Q5
Post Event * Yrs Elapsed	0.00761***	0.00472	0.000708	-0.00169	0.00734***	0.00698***
	(0.00264)	(0.00375)	(0.00176)	(0.00191)	(0.00256)	(0.00253)
Post Event		-0.00538		-0.00400		-0.00485
		(0.0151)		(0.0151)		(0.0121)
Trend		0.00417		0.00382*		0.000740
		(0.00459)		(0.00228)		(0.00295)
Observations	2832	2832	2828	2828	2819	2819
p, total event effect=0	0.00585	0.374	0.689	0.644	0.00600	0.0263
State-Event FEs	X	X	X	X	X	X
Sub-Gr-Yr FEs	X	X	X	X	X	X

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: The dependent variables are district-level mean NAEP test scores (in student-level standard deviation units) in the state-year cell for the relevant district income quintiles. State-copy fixed effects and NAEP exam fixed effects (i.e. subject-grade-year) are included. Regressions are weighted by the sample size in relevant subcategory. See text for further specification details. Standard errors are clustered at the state level.

Table 8: Event studies for test score slopes by subject and grade.

	Test Score Slope	Q1-Q5 Mean
Pooled	-0.00882***	0.00734***
	(0.00313)	(0.00256)
By Subject:		
Math	-0.0106***	0.00803***
	(0.00340)	(0.00304)
Reading	-0.00653*	0.00577**
	(0.00383)	(0.00244)
By Grade:		
4th	-0.0106***	0.00780***
	(0.00396)	(0.00286)
8th	-0.00724**	0.00728**
	(0.00341)	(0.00295)

Notes: Each coefficient represents a separate regression. In column 1, the dependent variables are the slopes of district-level mean NAEP test scores (in student-level standard deviation units) with respect to ln(district income) in the state-year cell for the relevant subject and/or grade subgroups. In column 2, the dependent variables are the difference in mean test scores between quintile 1 and quintile 5 districts. Pooled estimates correspond to column 1 of table 6; prior trends and post event indicators are not included. None of the differences in the above coefficients are statistically significant. State-copy fixed effects and NAEP exam fixed effects (i.e. subject-grade-year) are included. Regressions are weighted by the inverse of the estimated sampling variance of the dependent variable. See text for further specification details. Standard errors are clustered at the state level.

Table 9: Mechanisms: Teacher and student variables

	Post Event (1 para)	Post Event (3 para)	p (1 para)	p (3 para)
Slopes:				
Share black/hispanic	-0.00175	-0.00164	0.377	0.728
	(0.00197)	(0.00225)		
Share free/reduced price lunch	-0.0204	-0.0287	0.282	0.480
	(0.0187)	(0.0239)		
Mean teacher salary	-235.2	-22.09	0.800	0.990
	(921.0)	(748.1)		
Pupil teacher ratio	0.170	0.177	0.221	0.415
	(0.137)	(0.134)		
Q1-Q5 Means:				
Share black/hispanic	-0.00455	-0.00352	0.606	0.718
	(0.00878)	(0.00636)		
Share free/reduced price lunch	0.00510	0.00473	0.628	0.796
	(0.0105)	(0.0104)		
Mean teacher salary	309.2	-46.02	0.651	0.588
	(678.5)	(429.2)		
Pupil teacher ratio	-0.105	0.0614	0.444	0.832
	(0.137)	(0.103)		

Notes: In column 1, estimates of the post-event coefficient are shown for parametric event study models which include only parameter (only the post event variable). P values corresponding to these models are shown in column 3. In column 2, estimates of the post-event coefficent are shown for parametric event study models with 3 parameters (includes a post-event variable, an event-time trend variable, and a post-event time trend). Corresponding p-values for the joint test of both post event coefficients are shown in column 4. The columns in table 10 (following page) are analogously defined.

Table 10: Mechanisms: Revenue and expenditure variables

	Post Event (1 para)	Post Event (3 para)	p (1 para)	p (3 para)
Slopes:				
Total revenue per pupil	-321.2	-293.7	0.266	0.438
	(285.1)	(228.1)		
State revenue per pupil	-501.4***	-383.9**	0.0103	0.0339
	(187.6)	(153.8)		
Local revenue pp	44.34	-31.08	0.834	0.896
	(209.9)	(165.2)		
Federal revenue per pupil	35.43*	28.47*	0.106	0.0325
	(21.51)	(16.41)		
Total expenditures per pupil	-374.4	-333.2	0.194	0.397
	(284.5)	(252.7)		
Current instructional expenditure per pupil	-49.73	-22.53	0.721	0.909
	(138.3)	(108.8)		
Teacher salaries + benefits per pupil	-36.52	-24.14	0.797	0.975
	(141.0)	(125.3)		
Non-instructional expenditure per pupil	-236.0	-282.7*	0.198	0.264
	(181.0)	(170.8)		
Student support per pupil	-69.77	-49.08	0.306	0.465
	(67.41)	(54.17)		
Other current expenditures	-0.862	-7.769	0.943	0.517
	(11.96)	(9.320)		
Total capital outlays	-78.37	-94.84	0.447	0.584
	(102.2)	(92.24)		
Q1-Q5 Means:				
Total revenue per pupil	534.9***	258.0**	0.00445	0.118
	(179.5)	(123.0)		
State revenue per pupil	517.6**	245.7**	0.0175	0.0910
	(210.5)	(119.3)		
Local revenue per pupil	-46.04	2.484	0.794	0.548
	(175.5)	(134.9)		
Federal revenue per pupil	62.95**	9.498	0.0542	0.795
	(31.92)	(24.24)		
Total expenditures per pupil	548.0***	268.6**	0.00131	0.107
	(161.0)	(127.0)		
Current instructional expenditure per pupil	163.4*	-0.574	0.106	0.726
	(99.26)	(66.67)		
Teacher salaries + benefits per pupil	103.4	-10.10	0.202	0.673
	(80.03)	(60.03)		
Non-instructional expenditure per pupil	384.6***	269.2***	0.000102	0.00699
	(91.05)	(82.11)		
Student support per pupil	114.6*	43.62	0.0659	0.524
	(60.93)	(38.21)		
Other current expenditures	-1.977	1.428	0.761	0.0666
	(6.467)	(5.556)		
Total capital outlays	207.4***	167.9***	0.00588	0.0292
	(72.07)	(61.63)		

Notes: See notes to table 9

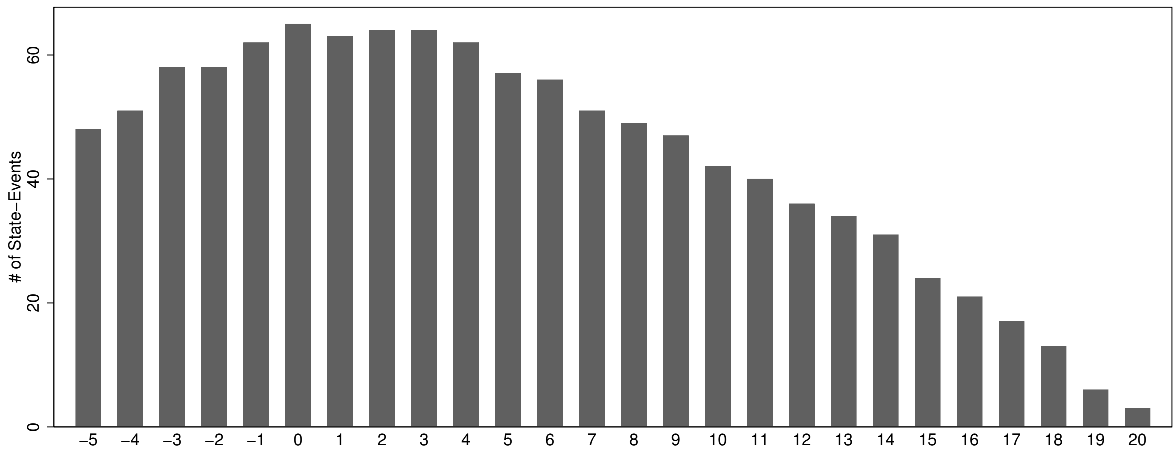
Table 11: Event studies for mean subgroup scores

	Post Event * Yrs Elapsed	
Pooled	0.00123	(0.00210)
25th percentile	0.00153	(0.00257)
75th percentile	0.000425	(0.00167)
By Race:		
White	0.00159	(0.00180)
Black	0.000990	(0.00266)
Black-white gap	-0.00103	(0.00205)
By Free Lunch Status:		
No Free Lunch	0.00123	(0.00184)
Free Lunch	0.000604	(0.00274)
Free lunch-no free lunch gap	-0.00192	(0.00177)

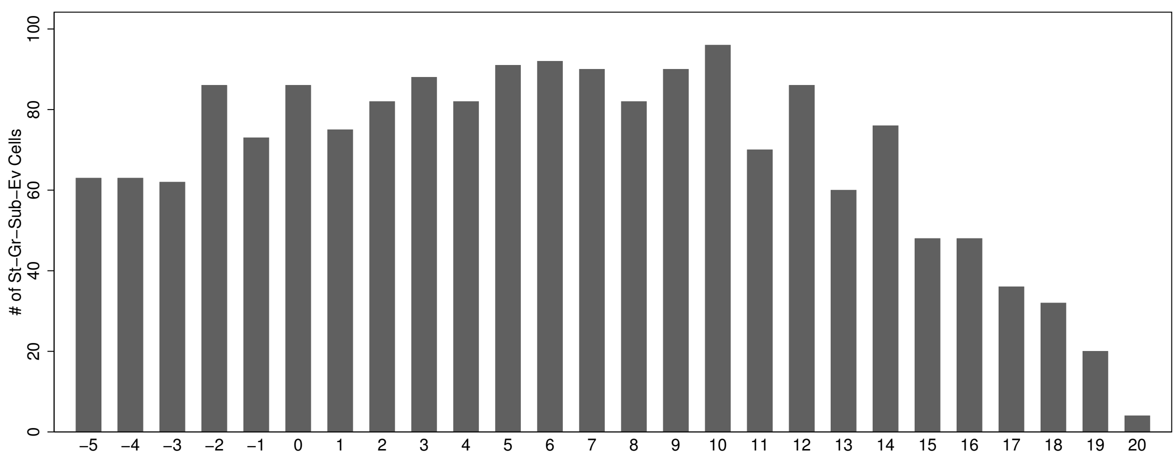
Notes: Black-white gap corresponds to the mean black score minus mean white score in each state-subject-grade-year cell. NAEP sample weights are used in the contruction of these state-subject-grade-year means. The free lunch-no free lunch gap is analogously defined. Regressions of mean score effects are weighted by the inverse of the subgroup sample size used to compute the subgroup sample mean in each state. Regressions with test score gaps as the dependent variable are weighted by the square root of the sum of the inverse subgroup sample sizes (e.g. $\sqrt{\frac{1}{N_a}+\frac{1}{N_b}}$, for subgroups a and b). For this reason, the estimated test score gaps do not necessarily correspond to the difference between the estimated coefficients for each subgroup.

3 Appendix Tables and Figures

Figure A1: Number of states/state-events at each "Event Year"



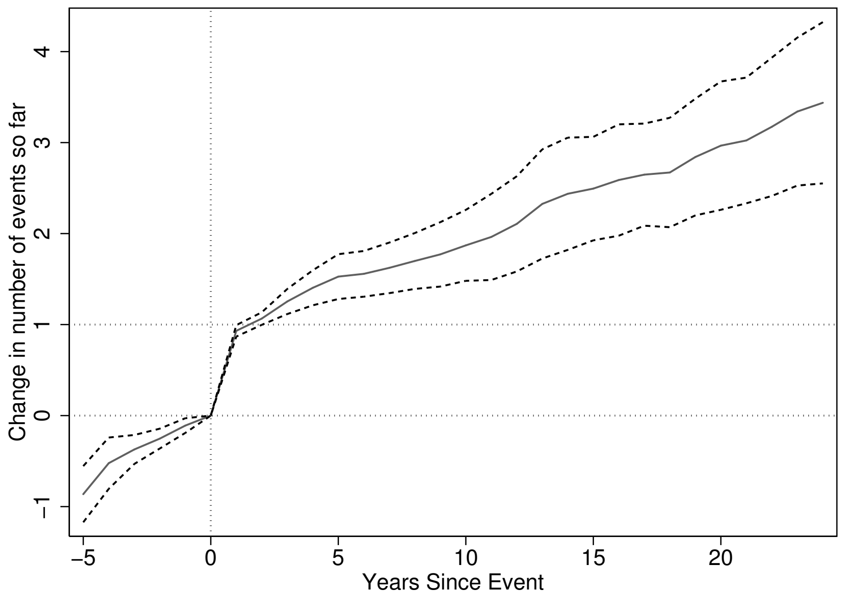
(a) State-year-event sample for finance analysis



(b) State-year-event-subject-grade sample for test score analysis

Notes: X-axis corresponds to "event-time" used in event study figures. States without events (there are 24) are not included in this figure. Panel A shows the number of state-event observations in the finance analysis. Panel B shows the number of state-subject-grade-event observations in the test score analysis.

Figure A2: Event study of number of events to date



Notes: Dependent variable is the number of events to date. Parametric estimates are reported in appendix table 5.

Table A1: List of Events

State	Year of Event	Type of Event
Alaska	2007	Court
Arizona	1994	Court
Arizona	1997	Court
Arizona	1998	Court
Arkansas	1994	Court
Arkansas	1995	Statute
Arkansas	2002	Court
Arkansas	2005	Court
Arkansas	2007	Statute
California	1998	Statute
California	2004	Statute
Colorado	2000	Statute
Connecticut	1996	Court
Idaho	1993	Court
Idaho	1994	Statute
Idaho	2005	Court
Indiana	2011	Statute
Kansas	1992	Statute
Kansas	2005	Court
Kansas	2005	Statute
Kentucky	1990	Court
Kentucky	1990	Statute
Maryland	1996	Court
Maryland	2002	Statute
Massachusetts	1993	Court
Massachusetts	1993	Statute
Missouri	1993	Court
Missouri	1993	Statute
Missouri	2005	Statute
Montana	1993	Statute
Montana	2005	Court
Montana	2007	Statute
New Hampshire	1993	Court
New Hampshire	1997	Court
New Hampshire	1999	Statute
New Hampshire	2002	Court
New Hampshire	2008	Statute

State	Year of Event	Type of Event
New Jersey	1990	Court
New Jersey	1990	Statute
New Jersey	1994	Court
New Jersey	1996	Statute
New Jersey	1997	Statute
New Jersey	1998	Court
New Jersey	2000	Court
New Jersey	2008	Statute
New Mexico	1999	Court
New Mexico	2001	Statute
New York	2003	Court
New York	2006	Court
New York	2007	Statute
North Carolina	1997	Court
North Carolina	2004	Court
North Carolina	2012	Court
North Carolina	2012	Statute
North Dakota	2007	Statute
Ohio	1997	Court
Ohio	2000	Court
Ohio	2000	Statute
Ohio	2002	Court
Tennessee	1992	Statute
Tennessee	1993	Court
Tennessee	1995	Court
Tennessee	2002	Court
Texas	1990	Statute
Texas	1991	Court
Texas	1992	Court
Texas	1993	Statute
Vermont	1997	Court
Vermont	2003	Statute
Washington	2010	Court
Washington	2013	Statute
West Virginia	1995	Court
West Virginia	2000	Court
Wyoming	1995	Court
Wyoming	1997	Statute
Wyoming	2001	Statute

Table A2: Difference in district income, 1990-2011

	(1)	(2)	(3)
Years Since Event (In 2011)	0.00237*	0.0313	-0.0121
	(0.00120)	(0.0353)	(0.0299)
log(Dist avg. HH income)	-0.0863***	-0.0519	-0.114***
	(0.0263)	(0.0397)	(0.0320)
Change in Dist Inc * Years Since Event		-0.00277	0.00130
		(0.00328)	(0.00280)
Observations	12527	12527	15576
Event States	X	X	
All States			X
Post 1990 Events	X	X	X

Standard errors clustered at the state level

Table A3: Alternative ways of handling event sample (C1-C3 slopes; C4-C6 Q1-Q5 means)

(a) First event in each state

	(1) St. Rev.	(2) Tot Rev.	(3) NAEP	(4) St. Rev.	(5) Tot Rev.	(6) NAEP
Post Event	-460.8*	-194.9		427.1	610.3**	
	(254.6)	(395.0)		(308.6)	(238.8)	
Post Event * Yrs Elapsed			-0.00810**			0.00461*
			(0.00332)			(0.00269)
Observations	4116	4116	1498	4256	4256	1568
p, total event effect=0	0.077	0.624	0.019	0.173	0.014	0.093
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X		X	X	
Sub-Gr-Yr FEs			X			X

(b) Reweight states w/ multiple events by 1/n

	(1) St. Rev.	(2) Tot Rev.	(3) NAEP	(4) St. Rev.	(5) Tot Rev.	(6) NAEP
Post Event	-563.9**	-317.7		559.1**	695.0***	
	(244.4)	(345.1)		(263.1)	(202.9)	
Post Event * Yrs Elapsed			-0.00771**			0.00487*
			(0.00362)			(0.00266)
Observations	7560	7560	2743	7696	7696	2819
p, total event effect=0	0.025	0.362	0.039	0.039	0.001	0.073
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X		X	X	
Sub-Gr-Yr FEs			X			X

(c) Number of events to date

	(1) St. Rev.	(2) Tot Rev.	(3) NAEP	(4) St. Rev.	(5) Tot Rev.	(6) NAEP
Num Events to Date	-289.6***	-180.7	-0.0252**	363.1**	337.3**	0.0199
	(101.6)	(164.7)	(0.0111)	(140.6)	(126.2)	(0.0121)
Observations	4116	4116	1498	4256	4256	1568
p, total event effect=0						
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X		X	X	
Sub-Gr-Yr FEs			X			X

Table A4: Alternative income measures (c1-c3 slopes; c4-c6 Q1-Q5 means)

(a) 2010: district income

	(1) St. Rev.	(2) Tot Rev.	(3) NAEP	(4) St. Rev.	(5) Tot Rev.	(6) NAEP
Post Event	-384.4**	-222.1		410.0*	394.9***	
	(168.3)	(220.5)		(209.5)	(146.1)	
Post Event * Yrs Elapsed			-0.00977***			0.00844***
			(0.00286)			(0.00207)
Observations	7560	7560	2743	7696	7696	2827
p, total event effect=0	0.027	0.319	0.001	0.056	0.009	0.000
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X		X	X	
Sub-Gr-Yr FEs			X			X

(b) 1990: housing values

	(1) St. Rev.	(2) Tot Rev.	(3) NAEP	(4) NAEP
Post Event	-331.8**	-214.1		
	(138.6)	(209.4)		
Post Event * Yrs Elapsed			-0.00717***	0.00871***
			(0.00201)	(0.00248)
Observations	7560	7560	2743	2826
p, total event effect=0	0.021	0.312	0.001	0.001
State-Event FEs	X	X	X	X
Year FEs	X	X		
Sub-Gr-Yr FEs			X	X

(c) 1990: share < 185% poverty line

	(1) St. Rev.	(2) Tot Rev.	(3) NAEP	(4) St. Rev.	(5) Tot Rev.	(6) NAEP
Post Event	0.000648	0.000123		-160.6	-208.6	
	(0.000498)	(0.000808)		(343.1)	(304.2)	
Post Event * Yrs Elapsed			-8.40e-09			-0.00201
			(1.20e-08)			(0.00481)
Observations	7560	7560	2743	7644	7644	2787
p, total event effect=0	0.200	0.880	0.489	0.642	0.496	0.678
State-Event FEs	X	X	X	X	X	X
Year FEs	X	X		X	X	
Sub-Gr-Yr FEs			X			X

Table A5: Fraction in each district income quintile

	Q1	Q2	Q3	Q4	Q5
Black	0.243	0.236	0.224	0.171	0.125
Black/Hispanic	0.240	0.228	0.243	0.171	0.117
White	0.198	0.190	0.182	0.202	0.228
Free/reduced-price lunch	0.315	0.216	0.201	0.158	0.110

Table A6: Event studies for per pupil revenue gaps (St. Rev. (C1, C2) and Tot. Rev. (C3, C4))

	(1) Black/White	(2) Free Lunch	(3) Black/White	(4) Free Lunch
Post Event	278.4*	51.99	220.1	79.41
	(147.4)	(211.1)	(166.4)	(165.6)
Observations	1810	1624	1810	1624
State-Event FEs	X	X	X	X
Year FEs	X	X	X	X

Standard errors clustered at the state level