Washington Center for Equitable Growth

1500 K Street NW, Suite 850
Washington, DC 20005

*Working paper series*

**School Finance Reform and the Distribution of Student Achievement**

Julien Lafortune
Jesse Rothstein
Diane Whitmore Schanzenbach

March 2016

http://equitablegrowth.org/school-finance-reform-and-the-distribution-of-student-achievement

**School Finance Reform and the Distribution of Student Achievement***

February 2016

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 gaps in spending and 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 absolute and relative spending in low-income school districts. Using representative samples 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 Apurba Chakraborty, Elora Ditton, and Patrick Lapid for excellent research assistance. We thank Julie Cullen, Tom Downes, Kirabo Jackson, Rucker Johnson, Richard Rothstein, and conference and seminar participants at APPAM, AEFP, Brookings, and the 2015 Stavanger-Bergen-Berkeley workshop for helpful comments and discussions.

## Introduction

Schools are a key link in the transmission of economic status across generations: 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. Schools in the United States 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.

The productivity of additional school resources is the subject of longstanding debate (see, e.g., Hanushek 2003; Krueger 2003; Burtless 1996). Time series and cross-district observational comparisons tend to show small or zero effects of spending on academic achievement (Hanushek 2006; Coleman et al. 1966), though state-level comparisons (Card and Krueger 1992a) and randomized experiments (Krueger 1999; Chetty et al. 2011) are more positive.

Compensatory funding – additional state aid for disadvantaged school districts – would create a downward bias in the estimated effect of school resources from observational designs. But it is exactly this type of program that is of interest for policy evaluation, as the state funding formula is the main policy tool available to address inequities in academic outcomes. Indeed, state funding formulas have been

[2] See Barrow and Schanzenbach (2012) for a review of this literature.

a locus for reform efforts. Beginning with the 1971 *Serrano v. Priest* decision, in which the California Supreme Court found the state's school finance system unconstitutional, many U.S. states have moved away from local funding to more centralized systems aimed at increasing opportunity for low-income students.[3]

Finance reforms are arguably the most important policy for promoting equality of educational opportunity since the turn away from school desegregation in the 1980s. A long literature examines the implications of these reforms for the distribution of school spending (see, e.g., Ladd and Fiske, 2015; Hanushek and Lindseth, 2009; Corcoran and Evans, 2015). Most relevant for our study, Corcoran and Evans (2015; see also Corcoran et al., 2004) find that plaintiff court victories reduce inequality of spending across districts. Fischel (1989) and Hoxby (2001) argue that poorly designed reforms sometimes led to "leveling down" of the top of the distribution rather than to absolute increases in spending in low-income districts. Nevertheless, Corcoran and Evans (2015) find that plaintiff victories lead to increases at the bottom of the spending distribution. Most relevant for our own study, Card and Payne (2002) find that reforms lead to increased relative spending in districts with low family incomes (which may or may not be low-spending districts).

Leveling down was possible because reforms in the 1970s and 1980s were focused on reducing gaps in funding between districts. A new wave of reforms in the 1990s was based on a different legal theory: That state constitutions required not

[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.

just *equitable* education spending but an *adequate* level of educational 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.

Although 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) and Corcoran and Evans (2015) are the only systematic studies of the effects of these reforms, taken as a group, on realized school finance, and both samples end 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 examining particular states (Clark 2003; Hyman 2013; Guryan 2001); by focusing

[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.

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, 2016; Candelaria and Shores 2015).

We provide the first evidence from nationally representative data regarding the impact of school finance reforms on student achievement. We rely on rarely used microdata from the National Assessment of Educational Progress (NAEP), also known as "the Nation's Report Card," to construct a state-by-year panel of average student achievement and of disparities between high- and low-income school districts. Conveniently, the beginning of our NAEP panel coincides with the onset of the adequacy era of school finance, which dates to the Kentucky Education Reform Act (KERA) of 1990.[5] We thus focus on identifying the effects of adequacy reforms.

The first part of our analysis documents impacts on absolute and relative spending levels in low- and high-income school districts. Using an event study framework, we find that finance reforms lead to sharp, immediate, and sustained increases in both state aid and total revenues in low-income districts. There are no signs of negative impacts on high-income districts; rather, these impacts are generally positive as well, though smaller. Although there is some evidence of subsequent reductions in local effort in high-income districts, even in these districts reforms have positive effects on total revenues for at least a dozen years.

[5] KERA was prompted by a 1989 court ruling in *Rose v. Council for Better Education* (790 SW 2d 186). 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.

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. Each becomes more progressive (via a reduction in the slope and an increase in the Q1-Q5 gap) following a reform event. 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 our 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, indicating that the extra school resources received by the former districts are used productively. The (local) average effect of an extra $1,000 in per-pupil annual 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 annual additional spending in the Project STAR class size experiment (which, translated into these terms, corresponds to an approximately 0.085 SD effect per $1,000 per pupil[6]). It implies

[6] 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 per year. We thus divide the STAR test score effect by two. This comparison implicitly assumes that maintaining the smaller STAR class sizes beyond 3rd grade would yield no additional growth in test scores.

that marginal increases in school resources in low-income, poorly resourced school districts are cost effective from a social perspective, even when the only benefits considered are those operating through subsequent earnings.

In a final analysis, we consider 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 in a state. 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 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 raises relative spending on the average low-income student by under $100 (not statistically distinguishable from zero). 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[7]

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

In the *Serrano v Priest* (1971),[8] the California Supreme Court accepted a novel legal theory (propounded in various forms by Wise 1967; Horowitz 1966;

[7] Our discussion here draws heavily on Koski and Hahnel (2015).

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 with a highly centralized school finance system that nearly perfectly equalizes per-pupil resources across districts.

After the U.S. Supreme Court rejected this legal theory in *San Antonio Independent School District v. Rodriguez*[9] in 1973, reform efforts 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 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, 2015; 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

---

[8] 487 P.2d 1241.
[9] 411 US 1.

an adequate education" (*Rose v. Council for Better Education*[10]; emphasis in original). The decision made clear that adequacy 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. Indeed, subsequent reforms have often aimed at *higher* spending in low-income than in high-income districts, to compensate for the out-of-school disadvantages of low-income students.[11]

The Kentucky legislature responded 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 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

[10] 790 SW 2d 186.

[11] A long literature studies the calculation of spending levels needed to satisfy an adequacy standard. See, e.g., Downes and Steifel, 2015, and Duncombe, Nguyen-Hoang, and Yinger, 2008.

did earlier reforms predicated on equity principles. One important distinction is that equity reforms often focused on inequities in property tax bases, where adequacy reforms focused on student disadvantage. Another is that, 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 to meet adequacy criteria in low-income districts while still allowing higher-income districts to differentiate themselves. 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 introduce our new post-1990 reform event database. Second, we discuss our summary measures of school finance and student outcomes in each state in each year. Third, we discuss our methodology for relating reform events to subsequent outcomes.

### A. Characterizing events

The most clear-cut school finance reform events are when a state's supreme court finds the state school financing system to be unconstitutional, and orders changes in the funding formula. Much of the prior school finance reform literature has focused on court-ordered reforms; we are able to draw on lists in Jackson et al. (2016), Hanushek and Lindseth (2009), and Corcoran and Evans (2015), 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.[12]

We use an inclusive definition of events, including many court orders that were subsequently reversed or were ignored by the legislature. In contrast to some prior work, we do not restrict attention to initial orders, though we also try not to label every procedural ruling a separate event. 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 reforms resulted from court orders. In some important cases (e.g., 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 in our event list major legislative reforms that change school finance systems, whether compelled by court orders or not.

As shown in Figure 1, we identify a total of 68 events in 27 states between 1990 and 2013. 51% are court orders and 40% are legislative actions; in 9% of cases, we identify one of each in the same year, and count them as a single event. A complete list of our events, along with a comparison to those used in other studies,

[12] Note that the 1990 start date encompasses KERA, the initial adequacy reform in Kentucky.

is presented in Appendix Table A1.[13] There have been more court-ordered finance reforms during the adequacy era than in the prior equity era.[14] Figure 2 shows the geographic distribution of events, using shading to represent the date of the first post-1989 event and numerals to indicate the number of events. Reform events are geographically dispersed, though rare in the deep South and upper Midwest. In states with multiple events, they were generally quite closely spaced: 59% were three or fewer years apart. In these cases, we suspect that the final event in the series was the direct cause of changes in the state's finance rules and that the prior events are best seen as procedural steps, but we do not impose this in the analysis.

### B. Measuring school finance systems and student outcomes

Next we turn to the measurement of the dependent variables of interest, beginning with the state finance regime. Here, a challenge is how to summarize the distribution of school resources.[15] Corcoran and Evans (2015), for example, examine the standard deviation of spending per pupil and other summaries of the univariate distribution. But this approach does not account for the relationship of spending to area economic resources. Since the central issues in school finance reform 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

---

[13] We have conducted all of our analyses using alternative event definitions (e.g., counting only initial events or only court orders), with qualitatively similar results. See Appendix Table A3.

[14] Although our database begins in 1990, Jackson et al. (2016) code 15 court-ordered reforms from 1971 through 1989, and 48 since then.

[15] Some authors categorize school finance systems by the form of the finance formula itself (e.g., minimum foundation plan, power equalization, etc. – see Hoxby, 2001 and Card and Payne 2002). But finance formulas do not always conform to these categories, and even two states with formulas of the same type may vary substantially in the extent of intended or actual redistribution.

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 (dis)advantage is the average family income in the district in 1990, relative to the state average.[16] We use two measures of finance equity. The first is the difference in average per-pupil revenue – either in total or from the state – 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. We estimate the following regression separately for each state and year:

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

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 controls for log enrollment and district type (elementary, secondary, or unified).[17] A more positive $\theta_{st}$ coefficient means a greater gap in funding between high- and low-

[16] The Appendix reports analyses using alternative measures (e.g., mean home values, or the share of families under 185% of poverty), with similar results. Much school finance litigation has focused on disparities in property tax bases, which are imperfectly 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 the variation in the definition of the tax base or in taxable non-residential property.
[17] We weight by mean log enrollment in the district across the entire sample, to reduce volatility in $\theta_{st}$ from changing enrollment over time. 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.

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 revenues are negatively correlated with mean incomes across districts in the state.

When we turn to our examination of student outcomes, we use parallel measures to those used in our finance analysis: The mean test scores of students at districts in the bottom quintile of the family income distribution, 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.[18] Each is estimated separately for each available state-year-subject-grade combination.

### C. Ohio Case Study

To illustrate these measures and their relationships to school finance reform events, we present Ohio as a case study. Figure 3 shows the relationship between district income and state revenues in Ohio in 1990 and 2010. On the horizontal axis is the log of the average household income in a school district in 1990. On the vertical axis, we show state revenues per pupil, in inflation-adjusted 2013 dollars, in 1990 (left panel) and 2011 (right panel). (We discuss the data sources at greater length in Section III.) In each panel, we overlay a regression line with slope $\theta_{st}$ as well as a step function showing mean revenues by district income quintile. In 1990, 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. The $\theta_{st}$ slope is negative in both years, indicating progressive state funding to districts, but is much

---

[18] When estimating test score-district income slopes, we drop the controls for district type from (1) and weight by NAEP sample weights rather than district enrollment.

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.

Figure 4a presents the scatterplot of state revenue-income slopes, $\theta_{st}$, 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.

Figure 4b shows the corresponding scatterplot for the slope of total revenues per pupil, inclusive of state revenues, local tax collections, and federal transfers, 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 again see declining gradients over time in most states.

Figure 3 shows that school finance changed substantially in Ohio 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 a "complete systematic overhaul" of the school funding system. In 2000, the 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, with a few minor changes, satisfied constitutional requirements. This decision was reversed by the same Court – with new judges since the previous year – in 2002. To our knowledge, there have not been substantial reforms to the finance system since then. We code Ohio as having judicial reform events in 1997 and 2002 and a joint statutory-judicial event in 2000.

Figure 5a shows the estimated state revenue-income and total revenue-income slopes $\theta_{st}$ over time for Ohio. Vertical lines indicate the reform events. The figure shows a clear effect of the 1997 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 events, which do 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. This mirrors the slope trends, with the expected vertical flip.

### 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. The non-parametric model specifies the outcome for state s in year t as:

$$(2) \qquad \theta_{snt} = \delta_{sn} + \kappa_t + \sum_{r=k_{min}}^{k_{max}} 1(t = t_{sn}^{*} + r)\, \beta_r + \varepsilon_{snt}.$$

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. $\beta_r$ 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 $k_{min}$=-5, so $\beta_{-5}$ represents average outcomes five or more years prior to an event, relative to those in the event year. $\kappa_t$ is a calendar year effect that is constant across states, while $\delta_{sn}$ represents a fixed effect for (each copy of) each state's data.[19]

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

[19] Equation (2) is weighted by the inverse estimated sampling variance of $\theta_{snt}$ except in analyses of quintile means or gaps in district revenues, where it is unweighted. In each case, standard errors are clustered at the state level.

outcomes in the years leading up 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 {$\beta_{-k}$, …, $\beta_{-1}$} terms capturing pre-event dynamics is designed to capture this. Non-zero coefficients would suggest that we are unable to distinguish the causal effects of events from the prior dynamics that led to them.

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:

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

Here, $\beta^{jump}$ 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. $\beta^{trend}$ 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. 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. The event study literature has not converged on a universally accepted way of handling repeated events. In our preferred estimates, we treat each of several events in a state separately. 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.) 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.

We have verified via Monte Carlo analyses (available upon request) that this stacked-panel approach recovers true event effects, so long as event timing does not depend on prior $\varepsilon_{st}$ realizations. We have also verified that our main results are qualitatively unchanged under alternative ways of handling multiple events – for example, using only the first event in each state, or using a single panel per state and allowing multiple events to have additive effects. See Appendix Table A3.

Our decision to treat each of several events in a state separately affects the interpretation of the post-event coefficients. The coefficient $\beta_r$, 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 principle, gradual increases in $\beta_r$ might 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

[20] See the related discussion in Cellini, Ferreira, and Rothstein (2010).

outcomes appear almost immediately following our designated events, and persist without growing thereafter. This is in part because events tend to be closely clustered in time—the median gap between consecutive events is only 3 years—and, we believe, in part because in most states there is generally only one "real" event in most states, with other events representing procedural maneuvering (including court orders that went unenforced) prior to or subsequent to that event.

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 $\kappa_t$ 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 $\beta_r$s.

## 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 4th and 8th 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). Census data are available annually since school year 1994-95, as well as in 1989-90 and 1991-92. We supplement this with sample data from the Census Bureau's Annual Survey of Government Finances for 1992-93 and 1993-94. 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, share of low-income students (defined as those eligible for free or reduced-priced lunch), and pupil-teacher ratios.

We draw district-level mean household income from the 1990 Census School District Data Book. We drop districts below the 2nd or above the 98th percentile of their state's (unweighted) distribution.

Finally, our student outcome measures come from the restricted-use NAEP microdata. We use the "State NAEP," which is designed to produce representative samples for each participating state. This began in 1990, with 8th grade math and 42 states participating, and has 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

[21] We exclude districts with highly volatile enrollment (year-over-year changes of 15% or more in any year, or with enrollment more than 10% off of a log-linear trendline in over one-third of years) and those with revenue per pupil below 20% or above 500% of the (unweighted) state-year mean.
[22] The NAEP also tests 12th graders, but high school dropout makes the samples nonrepresentative. We use only math and reading assessments, which are administered most frequently.

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 the first year that the test was given for the grade and subject, but allow both the mean and variance to evolve afterward. We then aggregate to the district-year-grade-subject level and merge to the CCD and SDDB.[23] We estimate separate quintile mean scores and score-income slopes for each state-year-subject-grade in our sample. Our event study sample thus consists of state-subject-grade-event number-year cells.

Table 2a presents district-level summary statistics, pooling data from 1990-2011. 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. 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. This gradient is roughly stable in the years leading up to a finance reform event, but declines by roughly $500 (scaled as 2013 dollars per pupil per one-unit change in log mean income) in the three years following the event. The gradient continues to decline thereafter, reaching a minimum total effect of -$937 in the 11$^{th}$ year after the event before rebounding somewhat, but is roughly stable from about year seven onward. Dotted lines in the

[23] The pre-2000 NAEP 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.

graph show pointwise 95% confidence intervals. 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 and statistically insignificant pre-event trend, a sharp downward jump following the event, and no change in the trend following this jump. 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 (3). In column 1, we include only state and year effects and the post-event indicator (i.e., we constrain $\beta^{trend} = \beta^{phasein} = 0$). Column 2 adds the phase-in effect, while column 3 also adds the trend term. (This third specification is shown in Figure 6.) The table also reports tests of the joint hypothesis that $\beta^{jump} = \beta^{phasein} = 0$. 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 log district income of about $500 per pupil, or about 5% of mean total revenues per pupil in our sample.

Figure 7 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). In the first quintile districts, state revenues increase sharply after events; fifth quintile districts see smaller but still substantial increases. The former effects grow over time, while the latter erode. As a result, the effect on the between-quintile gap is small at first but grows 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.

Many of our reform events do not – because of subsequent judicial reversals or legislative foot-dragging – ever lead to implemented changes in school finance. We thus view our estimates as intention-to-treat (ITT) effects, representing an average of the effects of implemented finance reforms with 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, and changes in the state aid formula may induce property value changes that affect local revenues even with fixed rates (Hoxby 2001). We thus turn next to models for total revenues per pupil, inclusive of state and local components. Models for the district income slopes are presented in Figure 8 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 (p<0.001). 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 (Table 3).

Figure 9, panels A-C, and Table 4B repeat the quintile mean analyses for total revenues. These are much more precise than the slope results. 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. This is about twice as large is implied by the (insignificant) total revenue-income slope results.

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 Figures 7D and 9D and Table 5. Average state revenues per pupil rise by about $760 following an event, with no sign of meaningful pre-event trends or phase-in effects. The increase in total revenues is smaller, around $550, but equally sharp and also highly significant.

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 latter 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, perhaps reflecting extra "bite" of adequacy reforms.[24] 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 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.

[24] Corcoran and Evans (2015) find that adequacy reforms have larger effects on spending levels than equity reforms, but smaller effects on between-district inequality. Their inequality measures, however, do not take account of district income or other measures of local resources. When we examine similar univariate inequality measures, we find no effect of adequacy reforms. Our income gradient and inter-quintile gap measures are closer to Sims' (2011) analysis, which finds that adequacy reforms lead to higher relative revenues in districts with greater student need.

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 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 smooth, nearly linear decline in the test score gradient following an event, indicating gradual increases in relative scores in low-income districts. This is exactly the pattern one would expect, as test scores are cumulative outcomes that presumably reflect not only current inputs 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. [25]

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 (but cluster standard errors at the state level). Columns 1-3 include state-event and subject-grade-year effects, while 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

[25] 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.

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 0.08 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 0.008 coefficient in Table 7 indicates a somewhat larger effect than 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 8$^{th}$ grade), an interpretation that is facilitated by the apparent lack of dynamics in the revenue effects. Nevertheless, the focus on the r=10 estimate is arbitrary. We would obtain larger estimates of the achievement effect per dollar if we used estimates for more than ten years after events (perhaps reflecting the time it takes to implement successful new programs after funding increases), or smaller effects with a shorter window.

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.[26]

## 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 latter than the former. 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 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

[26] In separate non-parametric models for scores by grade, akin to Figure 10, we find no indication that the effect on 4th grade scores stops growing five years after the event – both 4th and 8th grade effects appear to grow roughly linearly through the end of our panels. See Appendix Figure A3.

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.[27]

Other rows of Table 9 show proxies for classroom quality: The average pupil-teacher ratio and teacher salary. There are no significant effects on 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.[28] While neither of these is obvious as the most efficient route to increased learning, each may be productive at some margins (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 allows us only to say that the average finance reform—which we interpret to involve roughly unconstrained increases in

---

[27] We also find no relationship between events and the change in district income between 1990 and 2011. See Appendix Table A3.

[28] 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 capital spending increases.

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.[29]

## 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 than our slopes and quintile gaps of the overall effectiveness of a state's educational system at delivering equitable, adequate services to disadvantaged students (Krueger and Whitmore 2002; Card and Krueger 1992b). However, because only a small portion of income or other inequality is between districts, changes in 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 (pooled) test scores 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

---

[29] 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.

state that are correlated with the timing of school finance reforms (Hanushek, Rivkin, and Taylor (1996).

The second and third panels present results by race and income, respectively. There is no discernible effect on mean scores for any group, or on achievement gaps by race or income. Point estimates are roughly a 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 low-income students are more likely than their white and higher-income 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 low-income 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 income gap in these imputed 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 $221 (S.E. 167) and in the average low-income student's district by only $86 (S.E. 161). Even if this funding was *more* productive than the average effect implied by our pooled analysis, it would still not be enough to yield effects on black or low-income students' average test scores large enough to detect with our research design. Thus, while reforms aimed at 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 dramatic impacts 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 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: Reforms also led to increases in the absolute and relative achievement of students in low-income districts. Our estimates thus complement those of Jackson et al. (2016), 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 (2016) 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. (2016). School finance reforms are blunt tools, and some critics (Hanushek, 2006; Hoxby, 2001) have argued that they will be offset by changes in district or voter choices over tax rates or that funds will be spent so inefficiently as to be wasted. Our results do not support these claims. Courts and legislatures can evidently 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 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 resource or achievement gaps between 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**


Baker, B. D., & Green, P. C. (2015). Conceptions of Equity and Adequacy in School Finance. In H.F. Ladd and M.E. Goertz, eds., *Handbook of Research in Education Finance and Policy*, 2nd edition. New York, NY: Routledge.

Barrow, L., & Schanzenbach, D. W. (2012). Education and the Poor. In P. Jefferson, ed., *The Oxford Handbook of the Economics of Poverty*, 316-343. Oxford: Oxford University Press.

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. (1992a). 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. (1992b). 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., Ferreira, F., & Rothstein, J. (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.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., & Yagan D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR. *Quarterly Journal of Economics*, *126*(4), 1593-1660.

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. (2015). Equity, Adequacy, and the Evolving State Role in Education Finance. In H.F. Ladd and M.E. Goertz, eds., *Handbook of Research in Education Finance and Policy*, 2nd edition. New York, NY: 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 (2015). Measuring equity and adequacy in school finance. In H.F. Ladd & M. E. Goertz, eds., *Handbook of Research in Education Finance and Policy,* 2nd Edition. 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*, 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. The Netherlands: North Holland.

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. (1996). Aggregation and the estimated effects of school resources. *The Review of Economics and Statistics* 78(4), 611-627.

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. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics*, *131*(1).

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

Koski, W.S., & Hahnel, J. (2015). The past, present and future of educational finance reform litigation. In H.F. Ladd and M. E. Goertz, eds., *Handbook of Research in Education Finance and Policy*, 2nd Edition. New York, NY: Routledge.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics 114*(2), 497-532.

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 J. E. Chubb and T. Loveless, eds., *Bridging the Achievement Gap*. Washington, DC: Brookings Institution Press.

Ladd, H. F., & Fiske, E. B. (Eds.). (2015). *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.

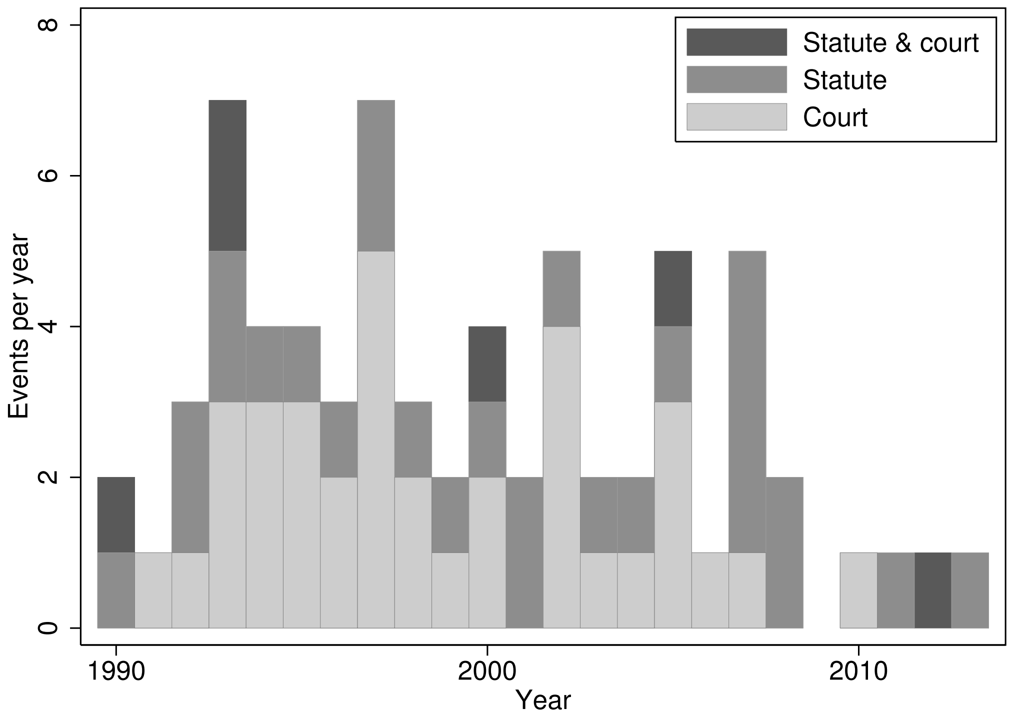
Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. Duncan and R. Murane (Eds.), *Whither opportunity*, 91-116. New York, NY: Russell Sage Foundation.

Sims, D. P. (2011). Suing for your supper? Resource allocation, teacher compensation and finance lawsuits. *Economics of Education Review*, *30*(5), 1034-1044.

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

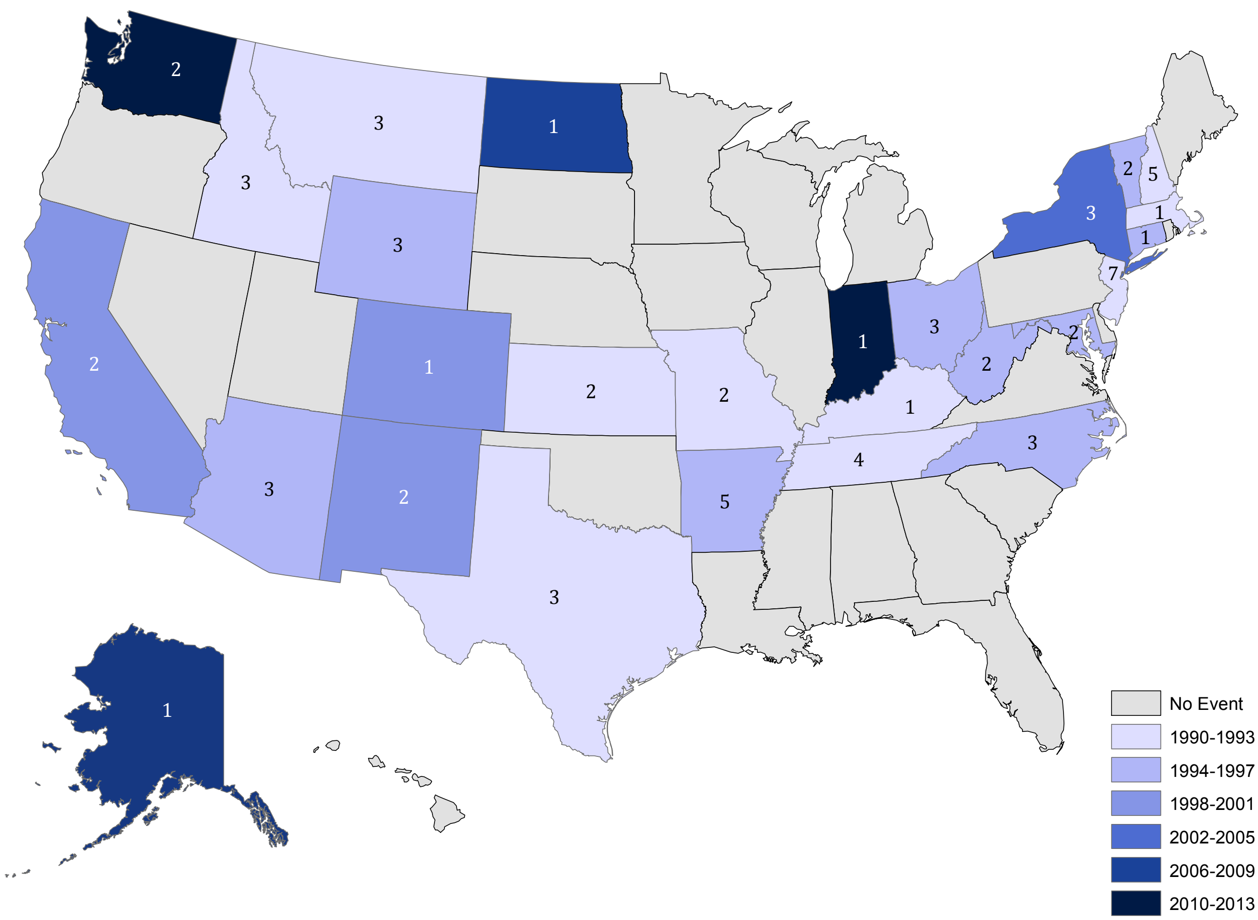
## Figures

Figure 1: Timing of school finance events



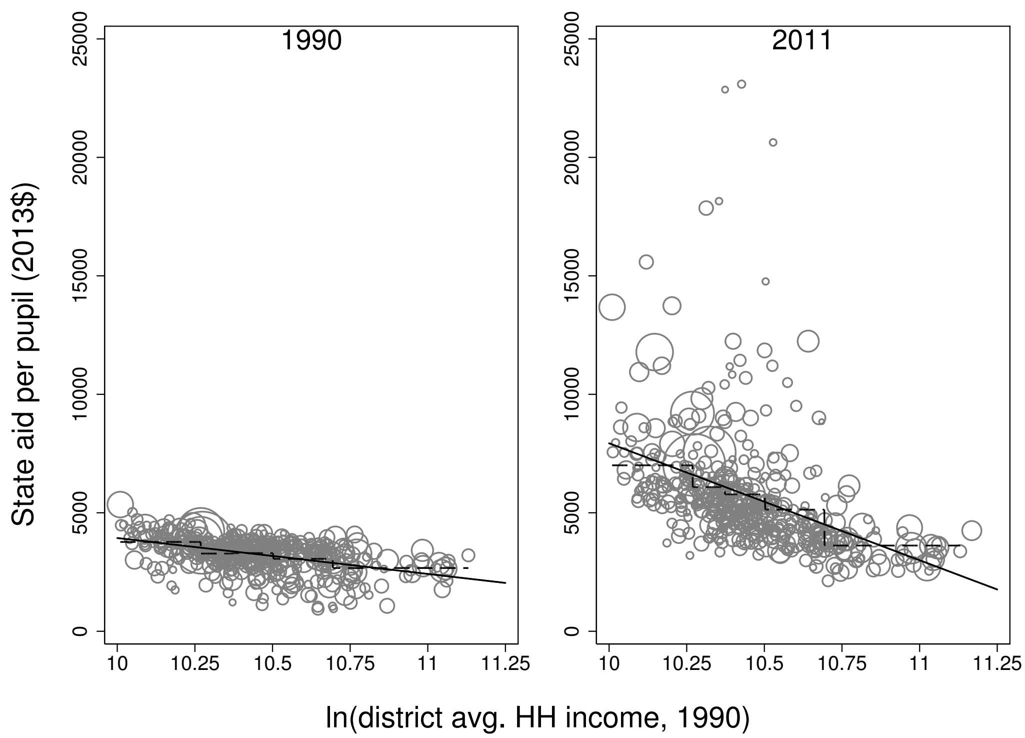

Notes: When multiple events occur in a state in a given year, they are combined into a single event for this chart.

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



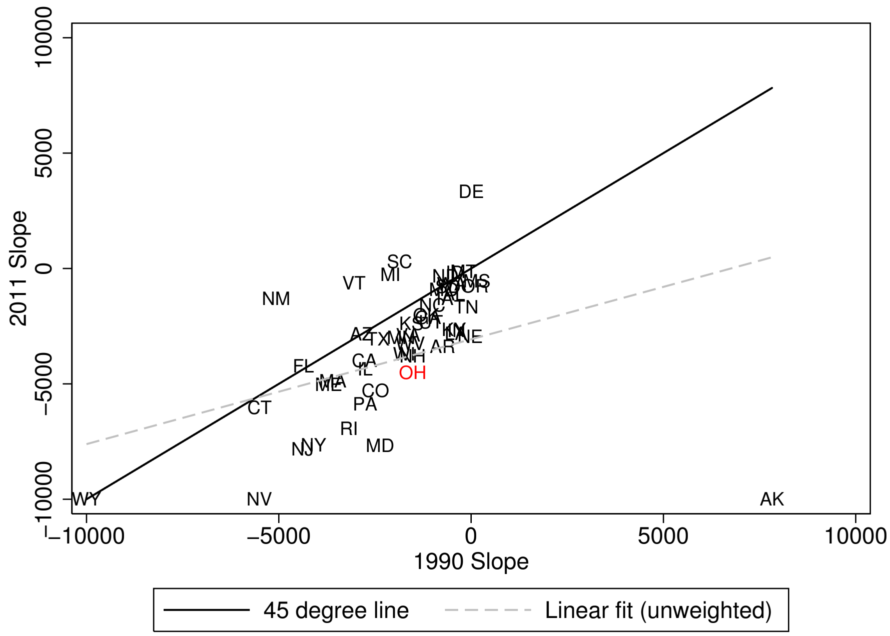

Notes: Colors correspond to the date of the first post-1989 school finance event. Numbers indicate the number of events in that period.

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

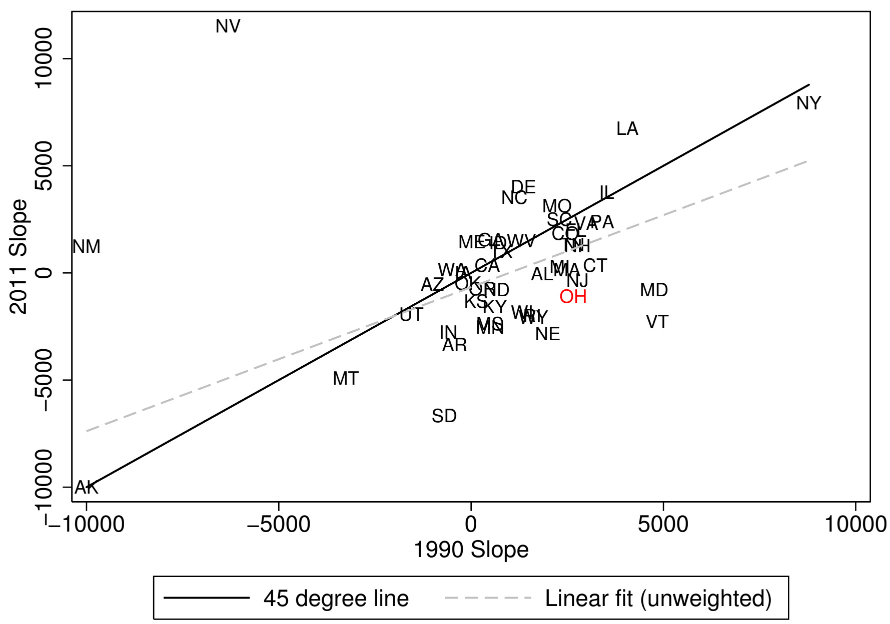



Notes: Each point represents one district. Circle sizes are proportional to average district enrollment over 1990-2011. Solid lines have slope equal to $\theta_{st}$ from equation (1) and correspond to predicted values for a unified district of average log enrollment. Dashed lines represent means among 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



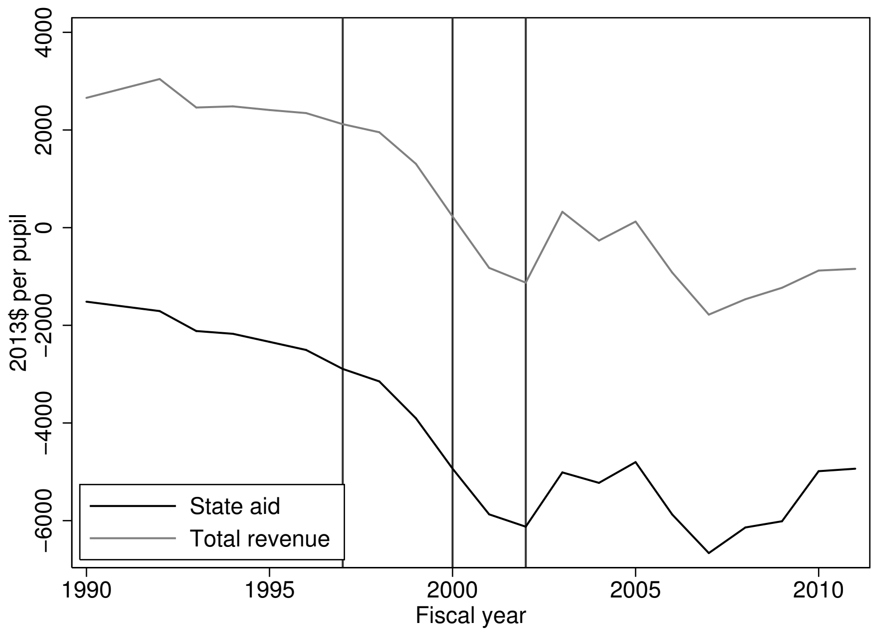

(a) State revenue per pupil



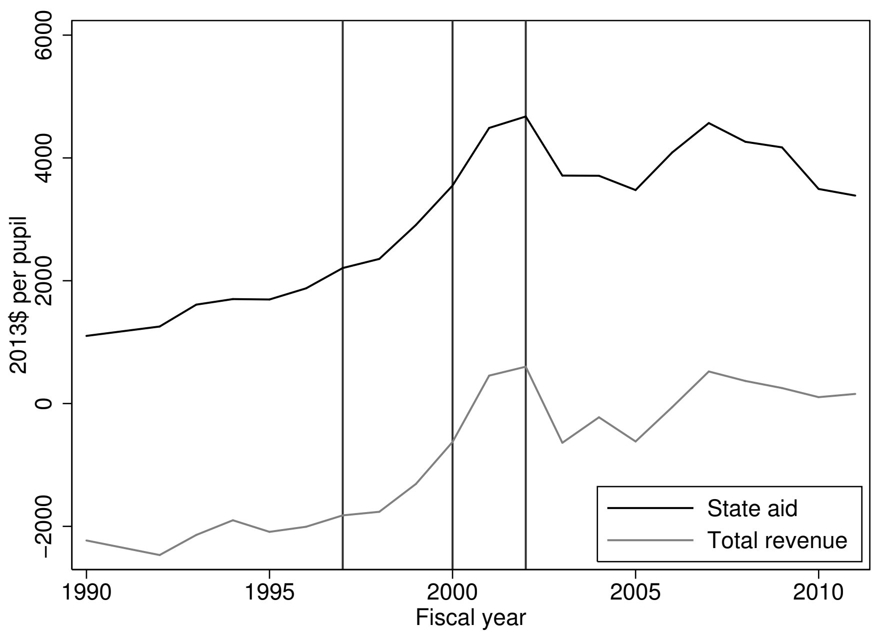

(b) Total revenue per pupil

Notes: Points indicate $\theta_{st}$ 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



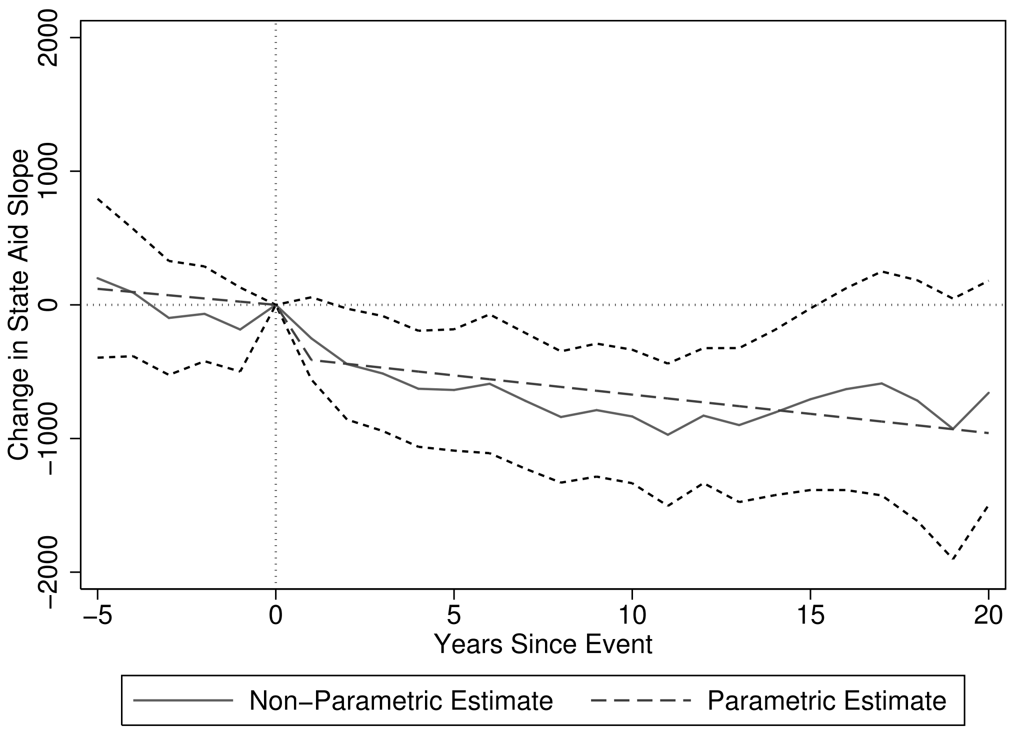

(a) Log income gradients




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

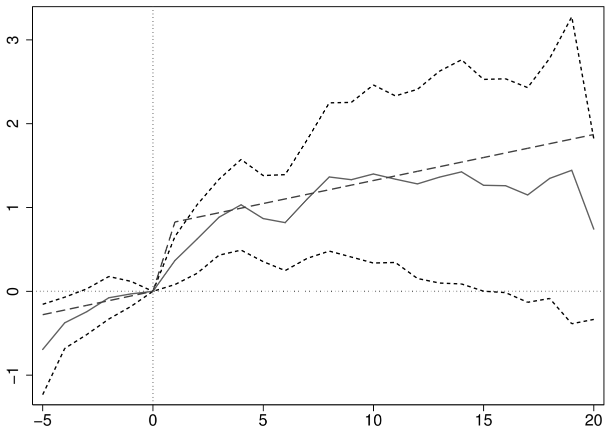
Notes: In panel (a), series represent $\theta_{st}$ from equation (1), varying the dependent variable, with 95% confidence intervals. In panel (b), series are the difference in the mean of the relevant revenue variable between districts in the first and fifth quintiles of the district mean income distribution. 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

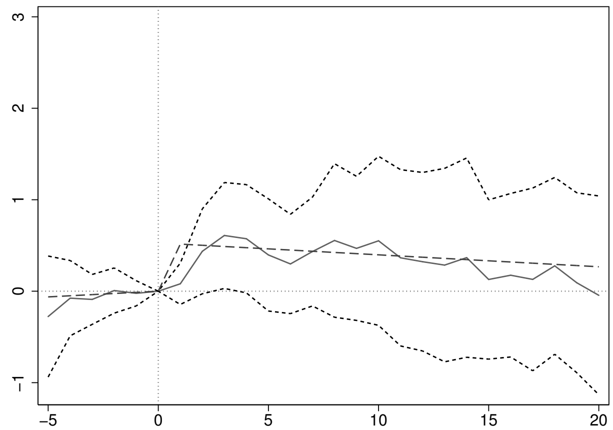



Notes: Dependent variable is the slope of state revenue per pupil with respect to ln(district income) in the state-year cell. Figure shows parametric and non-parametric estimates of the effect of a finance event on this slope, by years since (or prior to) the event, along with 95% confidence intervals 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.

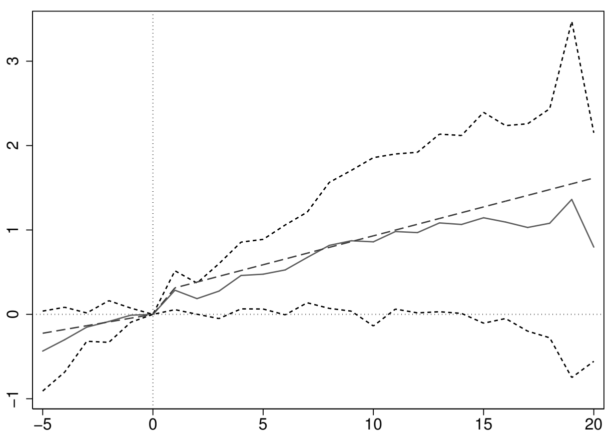
Figure 7: Event study estimates of effects of reform events on mean state revenues per pupil by district income quintile



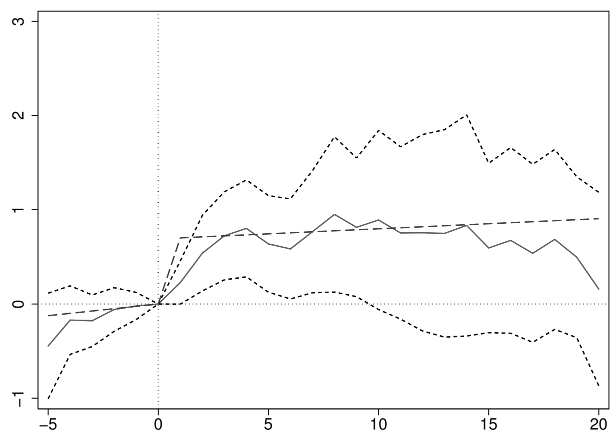

(a) Q1




(b) Q5



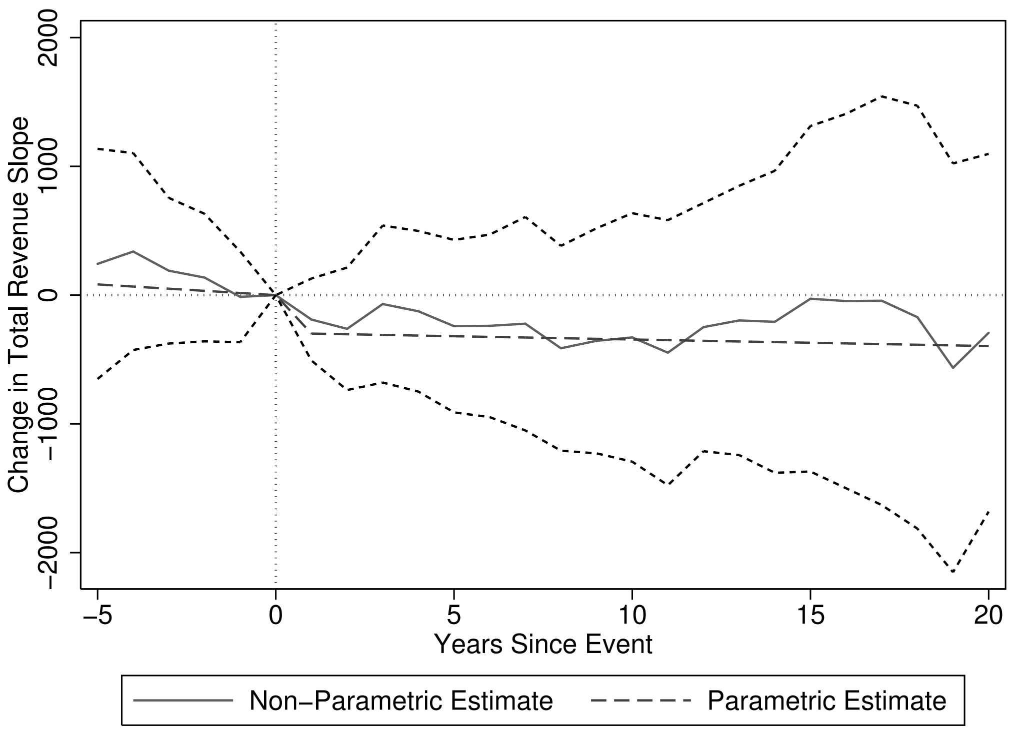

(c) Q1 - Q5




(d) Overall

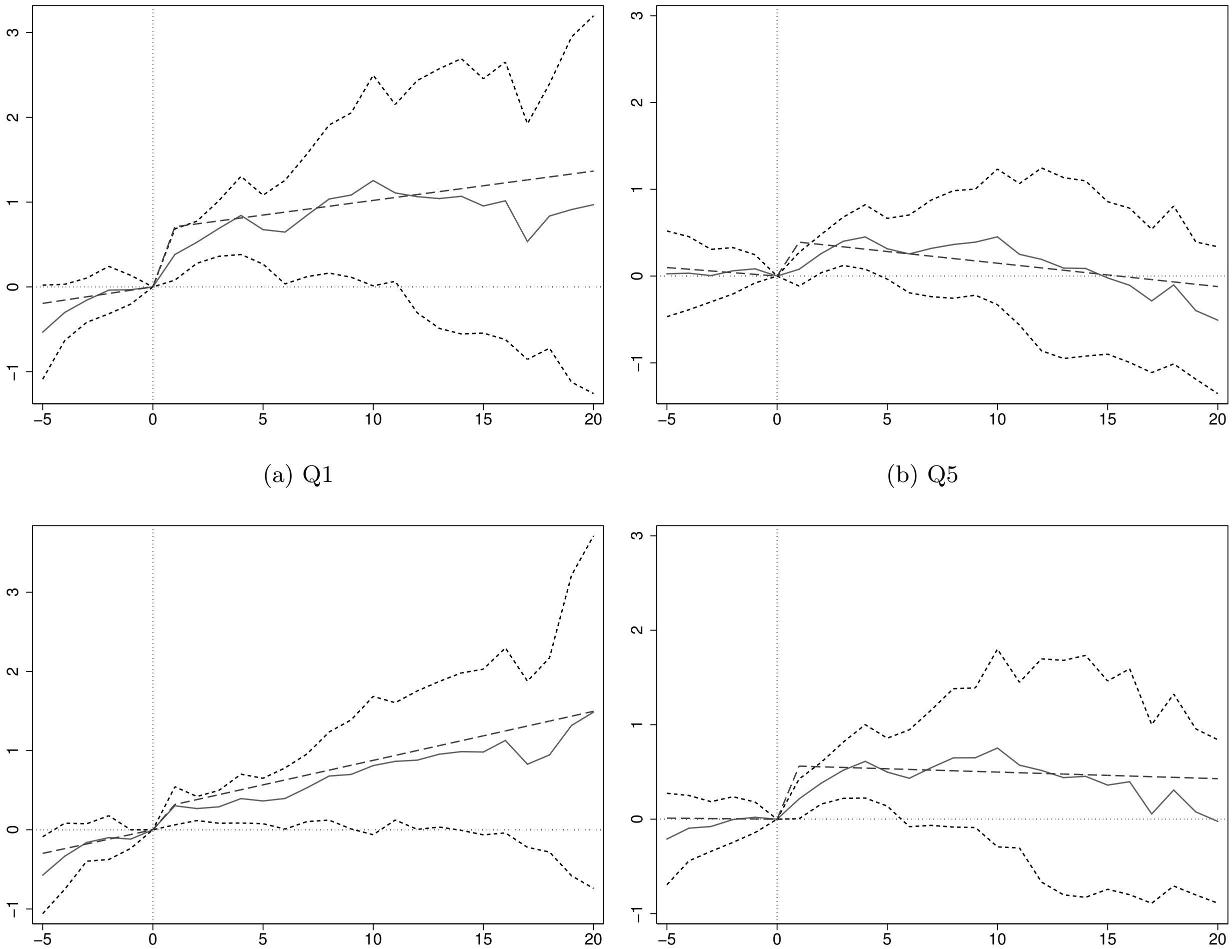
Notes: Dependent variable is mean state aid per pupil in the relevant subgroup of districts. In Panels A and B, the mean is for districts in the bottom fifth and top fifth, respectively, of the district mean income distribution (unweighted). In Panel C, the dependent variable is the difference between 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 8: 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 9: Event study estimates of effects of reform events on mean total revenues per pupil by district income quintile




(a) Q1

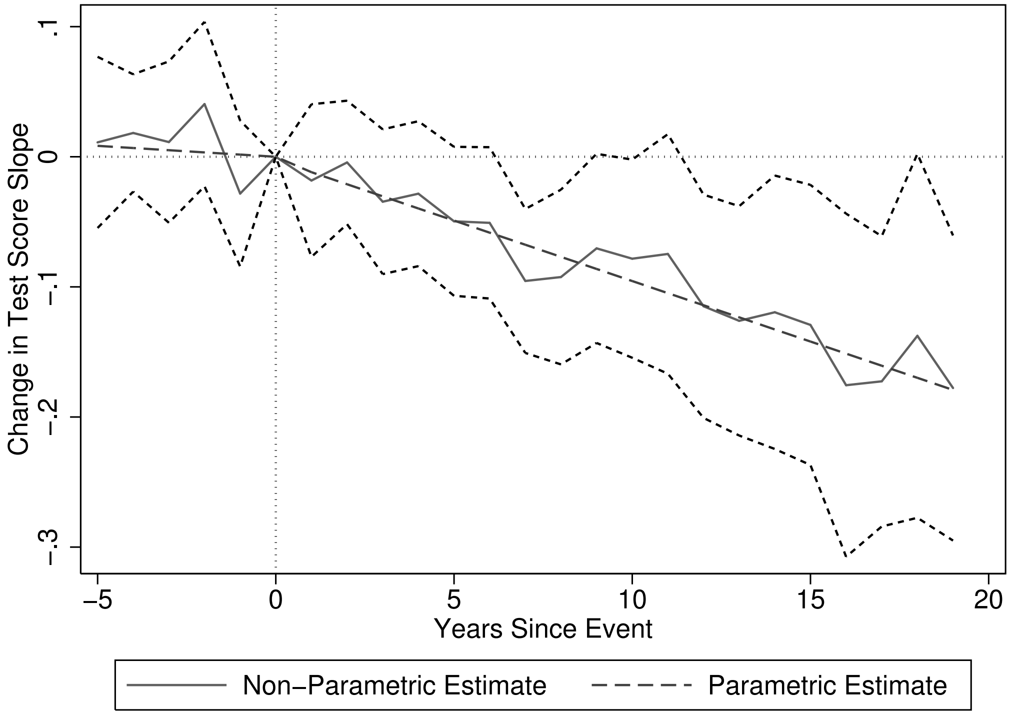
(b) Q5

(c) Q1 - Q5

(d) Overall

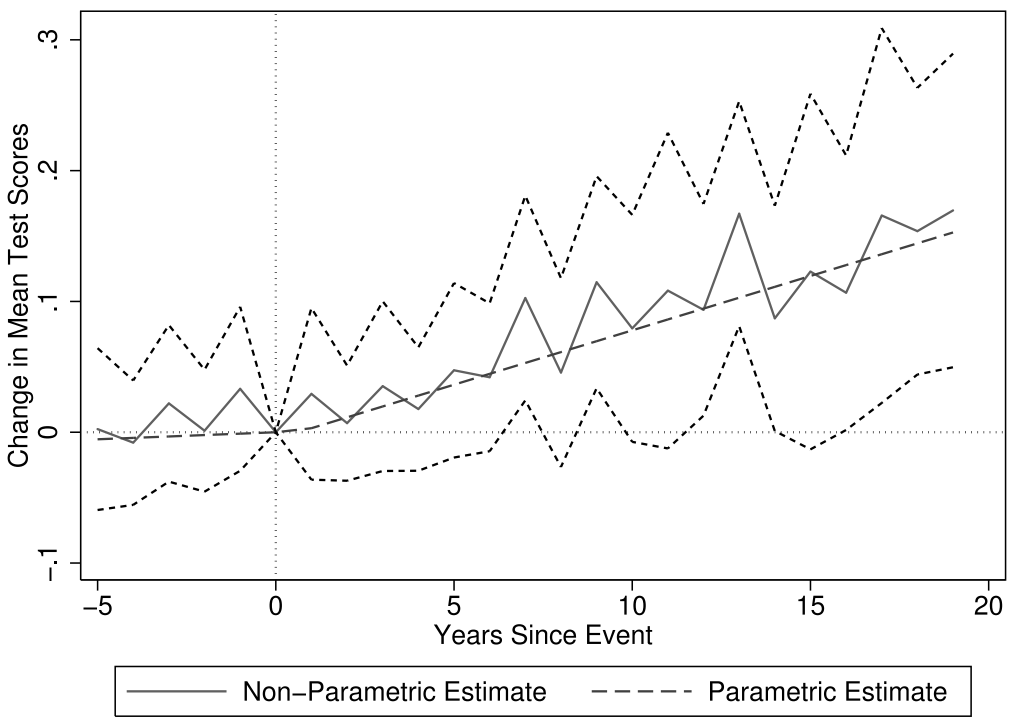
Notes: Dependent variable is mean total revenues per pupil in the relevant subgroup of districts. In Panels A and B, the mean is for districts in the bottom fifth and top fifth, respectively, of the district mean income distribution (unweighted). In Panel C, the dependent variable is the difference between 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.

# Tables

Table 1: NAEP Testing Years

| Year | Subject(s) | Grade(s) | Number of States | Number of Students Tested |
|---|---|---|---|---|
| 1990 | Math | G8 | 38 | 97,900 |
| 1992 | Math, Reading | G4, G8 | 42 | 321,120 |
| 1994 | Reading | G4 | 41 | 104,890 |
| 1996 | Math | G4, G8 | 45 | 228,980 |
| 1998 | Reading | G4, G8 | 41 | 206,810 |
| 2000 | Math | G4, G8 | 42 | 201,110 |
| 2002 | Reading | G4, G8 | 51 | 270,230 |
| 2003 | Math, Reading | G4, G8 | 51 | 691,360 |
| 2005 | Math, Reading | G4, G8 | 51 | 674,420 |
| 2007 | Math, Reading | G4, G8 | 51 | 711,360 |
| 2009 | Math, Reading | G4, G8 | 51 | 775,060 |
| 2011 | Math, Reading | G4, G8 | 51 | 749,250 |

Notes: Number of students tested is rounded to the nearest 10 to satisfy disclosure prevention rules.

Table 2: Summary statistics

(a) District-Year Panel

| | mean | sd | N |
|---|---|---|---|
| Total revenue per pupil | $10,979 | (3,376) | 208,207 |
| State revenue per pupil | $5,155 | (2,234) | 208,207 |
| Local revenue per pupil | $4,971 | (3,184) | 208,207 |
| Federal revenue per pupil | $853 | (625) | 208,207 |
| Log(Mean income) - 1990 | 10.51 | (0.27) | 208,207 |
| Unfied district | 0.93 | (0.25) | 208,207 |
| Elementary district | 0.05 | (0.21) | 208,207 |
| Secondary district | 0.02 | (0.14) | 208,207 |
| Total expenditure per pupil | $11,149 | (3,582) | 208,212 |
| Total instructional expenditure per pupil | $5,804 | (1,915) | 208,212 |
| Total non-instructional expenditure per pupil | $5,346 | (2,151) | 208,212 |
| Enrollment (student weighted) | 70,973 | (188,868) | 208,207 |
| Enrollment (unweighted) | 4,006 | (16,378.2) | 208,207 |

(b) State-Year Panel

| | mean | sd | N |
|---|---|---|---|
| State revenue slope | -3,164 | (3,512) | 4,116 |
| Total revenue slope | 326 | (3,666) | 4,116 |
| Test score slope | 0.95 | (0.36) | 1,498 |
| Dist income Q1 mean: state revenue | $6,430 | (2,856) | 4,264 |
| Dist income Q1 mean: total revenue | $11,462 | (3,798) | 4,264 |
| Dist income Q5 mean: state revenue | $4,410 | (2,278) | 4,256 |
| Dist income Q5 mean: total revenue | $11,554 | (3,358) | 4,256 |
| Dist income Q1-Q5 mean: state revenue | $2,012 | (2,094) | 4,256 |
| Dist income Q1-Q5 mean: total revenue | $-103 | (2,028) | 4,256 |
| Dist income Q1 mean: test scores | 0.07 | (0.37) | 1,574 |
| Dist income Q5 mean: test scores | 0.48 | (0.41) | 1,571 |
| Dist income Q1-Q5 mean: test scores | -0.41 | (0.30) | 1,569 |
| Num events to date | 0.77 | (1.29) | 5,100 |

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

| | (1) St. Rev. | (2) St. Rev. | (3) St. Rev. | (4) Tot. Rev. | (5) Tot. Rev. | (6) 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 |

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. All specifications include state-event and year fixed effects. Standard errors are clustered at the state level. See text for further specification details. P-values are for the joint hypothesis test that all after-event coefficients equal zero

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

(a) State Revenue

| | (1) Q1 | (2) Q1 | (3) Q5 | (4) Q5 | (5) Q1-Q5 | (6) Q1-Q5 |
|---|---|---|---|---|---|---|
| Post Event | 1022.7*** | 772.8*** | 510.2 | 528.1** | 517.5** | 245.7** |
| | (279.9) | (249.4) | (328.6) | (255.9) | (210.5) | (119.3) |
| Post Event * Yrs Elapsed | | -0.815 | | -25.48 | | 23.73 |
| | | (46.53) | | (23.81) | | (34.61) |
| Trend | | 55.73 | | 12.44 | | 44.75 |
| | | (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

| | (1) Q1 | (2) Q1 | (3) Q5 | (4) Q5 | (5) Q1-Q5 | (6) Q1-Q5 |
|---|---|---|---|---|---|---|
| Post Event | 838.0*** | 674.7*** | 307.5 | 417.8** | 534.4*** | 257.7** |
| | (236.8) | (209.8) | (220.9) | (193.1) | (179.5) | (123.0) |
| Post Event * Yrs Elapsed | | -4.258 | | -7.276 | | 2.316 |
| | | (58.69) | | (31.20) | | (38.73) |
| Trend | | 38.83 | | -19.68 | | 59.62* |
| | | (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 | 0.119 |

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 and are unweighted. Standard errors are clustered at the state level. See text for further specification details. P-values are for the joint hypothesis test that all after-event coefficients equal zero.

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

| | (1)<br>St. Rev. | (2)<br>St. Rev. | (3)<br>St. Rev. | (4)<br>Tot. Rev. | (5)<br>Tot. Rev. | (6)<br>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 |

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. All specifications include state-event and year fixed effects and are unweighted. Standard errors are clustered at the state level. See text for further specification details. P-values are for the joint hypothesis test that all after-event coefficients equal zero.

Table 6: Event study estimates for test score slopes

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| 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 | | | |

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-event and subject-grade-year fixed effects; columns 4-6 include state-event-grade-subject and year fixed effects. Regressions are weighted by the inverse of the estimated sampling variance of the dependent variable. Standard errors are clustered at the state level. See text for further specification details. P-values are for the joint hypothesis test that all after-event coefficients equal zero.

Table 7: Event studies for mean subgroup scores

| | (1) Q1 | (2) Q1 | (3) Q5 | (4) Q5 | (5) Q1-Q5 | (6) Q1-Q5 |
|---|---|---|---|---|---|---|
| Post Event * Yrs Elapsed | 0.00759*** | 0.00477 | 0.000386 | -0.00173 | 0.00780*** | 0.00724** |
| | (0.00265) | (0.00381) | (0.00181) | (0.00219) | (0.00269) | (0.00291) |
| Post Event | | -0.00378 | | -0.00418 | | -0.00521 |
| | | (0.0145) | | (0.0158) | | (0.0117) |
| Trend | | 0.00401 | | 0.00344 | | 0.00108 |
| | | (0.00450) | | (0.00254) | | (0.00312) |
| Observations | 2833 | 2833 | 2828 | 2828 | 2820 | 2820 |
| p, total event effect=0 | 0.00606 | 0.378 | 0.832 | 0.715 | 0.00554 | 0.0404 |

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. Each regression includes state-event and subject-grade-year fixed effects. Regressions are weighted by estimates of the sampling variance of the dependent variable: In columns 1-4 by the number of NAEP students in the state-subject-grade-year-quintile cell, and in columns 5 and 6 by the harmonic mean of the two sample counts. Standard errors are clustered at the state level. See text for further specification details. P-values are for the joint hypothesis test that all after-event coefficients equal zero.

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

| | Test Score Slope | Q1-Q5 Mean |
|---|---|---|
| Pooled | -0.00882*** | 0.00780*** |
| | (0.00313) | (0.00269) |
| *By Subject*: | | |
| Math | -0.0106*** | 0.00826*** |
| | (0.00340) | (0.00300) |
| Reading | -0.00653* | 0.00645** |
| | (0.00383) | (0.00268) |
| *By Grade*: | | |
| 4th | -0.0106*** | 0.00866*** |
| | (0.00396) | (0.00322) |
| 8th | -0.00724** | 0.00735*** |
| | (0.00341) | (0.00284) |

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 and column 5 of table 7, respectively. Each regression includes state-event and subject-grade-year fixed effects; prior trends and post event indicators are not included. Regressions are weighted by the inverse of the estimated sampling variance of the dependent variable and standard errors are clustered at the state level. See text for further specification details. None of the differences between math and reading or between 4th and 8th grade coefficients are statistically significant.

Table 9: Mechanisms: Teacher and student variables

| | Post Event (1 para) | Post Event (3 para) | Post * Yrs Elapsed (3 para) | $p$ (3 para) |
|---|---|---|---|---|
| *Slopes*: | | | | |
| Share black/hispanic | -0.00180 | -0.00191 | 0.0000956 | 0.669 |
| | (0.00207) | (0.00235) | (0.000481) | |
| Share free/reduced price lunch | -0.0207 | -0.0296 | 0.00439 | 0.456 |
| | (0.0189) | (0.0241) | (0.00489) | |
| Mean teacher salary | -235.2 | -22.09 | -11.58 | 0.990 |
| | (921.0) | (748.1) | (147.0) | |
| Pupil teacher ratio | 0.198 | 0.179 | -0.0282 | 0.431 |
| | (0.136) | (0.138) | (0.0338) | |
| *Q1-Q5 Means*: | | | | |
| Share black/hispanic | -0.00789 | -0.00671 | -0.000644 | 0.474 |
| | (0.00861) | (0.00640) | (0.00141) | |
| Share free/reduced price lunch | 0.00612 | 0.00538 | -0.00143 | 0.778 |
| | (0.0104) | (0.0103) | (0.00210) | |
| Mean teacher salary | 309.2 | -46.02 | 97.69 | 0.588 |
| | (678.5) | (429.2) | (98.88) | |
| Pupil teacher ratio | -0.112 | 0.0613 | -0.000854 | 0.839 |
| | (0.138) | (0.104) | (0.0182) | |

Notes: In column 1, estimates of the post-event coefficient are shown for parametric event study models which include only the post event variable (plus state-event and year effects). In columns 2 and 3, estimates of the post-event coefficients are shown for the three-parameter parametric event study model, including a post-event indicator, a time trend (not shown), and a post-event change in the time trend. P-values for the joint test of both post event coefficients in the 3 parameter model are shown in column 4.

Table 10: Mechanisms: Revenue and expenditure variables

| | Post Event (1 para) | Post Event (3 para) | Post * Yrs Elapsed (3 para) | $p$ (3 para) |
|---|---|---|---|---|
| *Slopes*: | | | | |
| Total revenue | -321.2 | -293.7 | 11.54 | 0.438 |
| | (285.1) | (228.1) | (40.18) | |
| State revenue | -501.4$^{***}$ | -383.9$^{**}$ | -4.760 | 0.0339 |
| | (187.6) | (153.8) | (19.25) | |
| Local revenue | 44.34 | -31.08 | -6.270 | 0.896 |
| | (209.9) | (165.2) | (20.45) | |
| Federal revenue | 35.43$^{*}$ | 28.47$^{*}$ | 2.733 | 0.0325 |
| | (21.51) | (16.41) | (5.119) | |
| Total expenditures | -374.4 | -333.2 | 13.82 | 0.397 |
| | (284.5) | (252.7) | (49.44) | |
| Current instructional expenditure | -49.73 | -22.53 | -5.128 | 0.909 |
| | (138.3) | (108.8) | (21.04) | |
| Teacher salaries + benefits | -36.52 | -24.14 | -0.303 | 0.975 |
| | (141.0) | (125.3) | (19.93) | |
| Non-instructional expenditure | -236.0 | -282.7$^{*}$ | 20.53 | 0.264 |
| | (181.0) | (170.8) | (28.65) | |
| Student support | -69.77 | -49.08 | -6.202 | 0.465 |
| | (67.41) | (54.17) | (7.290) | |
| Other current expenditures | -0.862 | -7.769 | 1.129 | 0.517 |
| | (11.96) | (9.320) | (1.453) | |
| Total capital outlays | -78.37 | -94.84 | 3.487 | 0.584 |
| | (102.2) | (92.24) | (12.05) | |
| *Q1-Q5 Means*: | | | | |
| Total revenue | 534.4$^{***}$ | 257.7$^{**}$ | 2.316 | 0.119 |
| | (179.5) | (123.0) | (38.73) | |
| State revenue | 517.5$^{**}$ | 245.7$^{**}$ | 23.73 | 0.0910 |
| | (210.5) | (119.3) | (34.61) | |
| Local revenue | -45.79 | 2.717 | -14.39 | 0.548 |
| | (175.5) | (134.9) | (13.37) | |
| Federal revenue | 63.07$^{**}$ | 9.558 | -7.025 | 0.795 |
| | (31.92) | (24.22) | (11.30) | |
| Total expenditures | 541.0$^{***}$ | 257.4$^{**}$ | 5.249 | 0.128 |
| | (161.4) | (127.3) | (36.15) | |
| Current instructional expenditure | 163.6$^{*}$ | -0.383 | 10.95 | 0.726 |
| | (99.27) | (66.69) | (13.63) | |
| Teacher salaries + benefits | 103.5 | -9.957 | 11.58 | 0.673 |
| | (80.04) | (60.05) | (13.03) | |
| Non-instructional expenditure | 377.4$^{***}$ | 257.8$^{***}$ | -5.703 | 0.0117 |
| | (92.10) | (83.52) | (27.13) | |
| Student support | 114.7$^{*}$ | 43.81 | 2.681 | 0.522 |
| | (60.94) | (38.22) | (10.08) | |
| Other current expenditures | -1.924 | 1.493 | -3.021$^{**}$ | 0.0666 |
| | (6.464) | (5.535) | (1.268) | |
| Total capital outlays | 200.0$^{***}$ | 156.4$^{**}$ | -12.37 | 0.0501 |
| | (72.99) | (62.79) | (23.10) | |

Notes: All revenue and expenditure variables are per pupil. See notes to table 9 for a description of each column.

Table 11: Event studies for mean subgroup scores

| | Post Event * Yrs Elapsed | |
|---|---|---|
| Pooled | 0.00125 | (0.00208) |
| 25th percentile | 0.00151 | (0.00258) |
| 75th percentile | 0.000507 | (0.00169) |
| *By Race*: | | |
| White | 0.00159 | (0.00179) |
| Black | 0.000601 | (0.00265) |
| White-black gap | -0.000534 | (0.00211) |
| *By Free Lunch Status*: | | |
| No Free Lunch | 0.00108 | (0.00184) |
| Free Lunch | 0.000733 | (0.00271) |
| No free lunch-free lunch gap | -0.00216 | (0.00182) |

Notes: White-black gap corresponds to the mean white score minus mean black score in each state-subject-grade-year cell, using NAEP sample weights for each. The no free lunch-free lunch gap is analogously defined. Regressions of mean score effects are weighted by the number of NAEP observations used to compute the subgroup mean. Regressions with test score gaps as the dependent variable are weighted by the harmonic mean of the two subgroup sample sizes, and thus do not correspond exactly to the difference between the two subgroup models.

# Appendix

We present in this appendix several additional analyses not included in the main paper.

### Sample construction

Figure A1 and Table A1 give additional background on the sample construction in the event study analysis. Table A1 lists the events included in our database and compares them to those used in prior studies. A more complete discussion of event classification is given in section II.A. of the text. Figure A1 shows the distribution of state-year (in the finance analyses) or state-grade-subject-year (in the NAEP analyses) observations in event time. Both the finance and NAEP panels are fairly balanced in event time, at least up to 10 years after the initial event.

### Number of events

During the period we study, many states had several court-ordered and legislative school finance reforms, which complicate analysis and interpretation using traditional event study methods. To empirically address the magnitude of potential biases from overlapping event-time windows within certain states, we estimated our non-parametric event study model, taking the total number of events to date in the state as the dependent variable. Figure A2 shows the estimates. The average state experiences about one additional statutory or court-ordered event in the five years prior to an event, and another one in the ten subsequent years. This suggests that the preferred reduced form estimates of the effects of finance reforms on realized financial and student achievement outcomes reflect the combined effects of more than one event, typically relatively closely clustered in time.

### Heterogeneity by grade

One might expect that the effects of funding changes would accumulate over a child's time in school. If effects were otherwise constant, this would imply that effects of discrete finance changes on 4th grade scores would phase in over 5 years, while those on 8th grade scores would phase in over 9 years. Figure A3 reports separate parametric and non-parametric event study estimates for the slopes of 4th and 8th grade NAEP scores with respect to log mean district incomes. We find no evidence of differential impacts or timing between 4th and 8th grade students. The nonparametric 4th grade test score estimates are almost all greater (in absolute value) than the 8th grade estimates, though not significantly so, and this gap appears to grow slightly rather than shrink over time.

### Change in district incomes

Finance reforms may impact the housing market, perhaps leading higher-income families to move into previously low-income districts experiencing infusions of state funds. This would represent an alternative explanation for our test score effects – they might reflect changing student composition rather than impacts on student learning. Table A2 investigates whether the finance reform events are associated with changes in district income composition. The table reports estimates from models where the difference in log district incomes between 1990 and 2011 (2011 income minus 1990 income) is the dependent variable. Independent variables are the number of years since the event, log of 1990 mean district income, and their interaction. When the interaction term is not included, there is a slightly positive and marginally significant relationship between time elapsed since the event and log district income changes in states that experienced an event.

When the interaction term is included, the years since event coefficient is still positive, while the interaction is negative. Neither coefficient is significant whether or not non-event states are included in the estimation as controls. When all states are included in the analysis (column 3), the sign on the coefficients is reversed. Both coefficients are insignificant in columns 2 and 3, where the interaction term is included. This provides suggestive evidence that finance reform events did not create confounding changes in district composition.

#### Robustness / alternative specifications

Table A3 reports alternative event study specifications aimed at gauging the sensitivity of results to the way that states with multiple events are handled. Panel (a) shows estimates from models where only the first event in a state is used. Results are largely consistent with those estimated using the full sample.

Panel (b) reports estimates from our event study models that use only court-ordered reform events. This specification avoids concerns about potential endogeneity of the timing of legislative reforms. Results are very similar to our baseline results. Including both court ordered and legislative reforms, as we do in our baseline specifications, does not appear to introduce meaningful biases in our estimates.

Panel (c) reports estimates where states are reweighted by $1/n$, where $n$ is the number of total events in a state during the sample period, to place equal total weight on each state. These are again qualitatively similar to baseline estimates. Panel (d) shows estimates from models where the independent variable is the number of events to date. This corresponds to a model where each event has a separate additive effect. In the NAEP analyses, we allow each event to have a separate additive effect on the trend, so the key independent variable is the sum across all events to date of the number of years elapsed since the event. Estimated finance effects are slightly smaller but qualitatively similar to the baseline results. Test score effects are significant but substantially smaller, but should be added across the number of events that occur in the state.

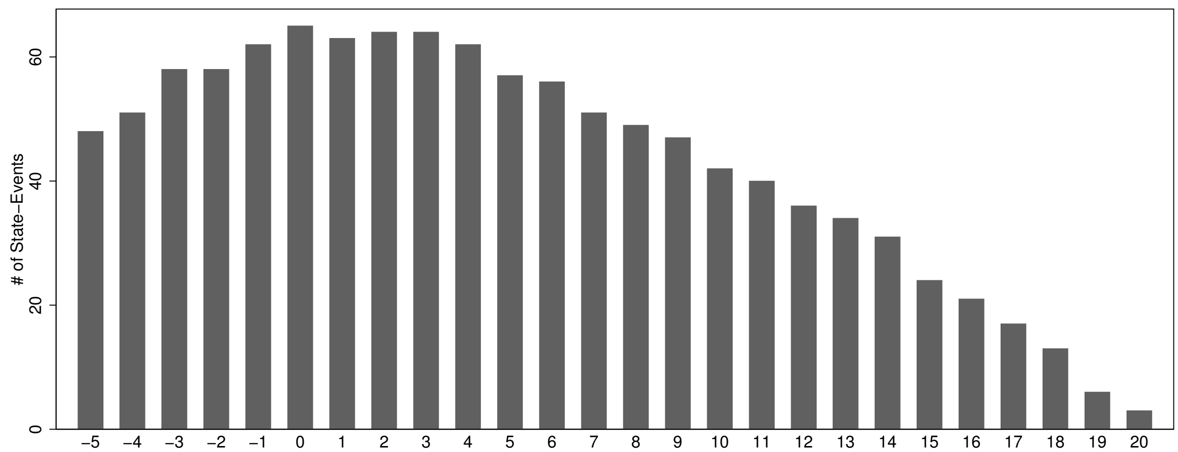
Table A4 reports estimates where we vary the measure used to classify district (dis)advantage. Our baseline models use 1990 mean household incomes. Panels (a) and (b) of Table A4 are computed using 2010 mean incomes and 1990 mean housing values (for owner-occupied housing), respectively. These are both very highly correlated with 1990 mean household income, and the results are not sensitive to the change. The final panel of table A4 uses the share of individuals in a district with income below 185% of the federal poverty line in place of the district's log mean income. Estimates computed using the poverty-based measure are qualitatively consistent with our baseline estimates, although none are statistically significant.

#### Income / race analysis

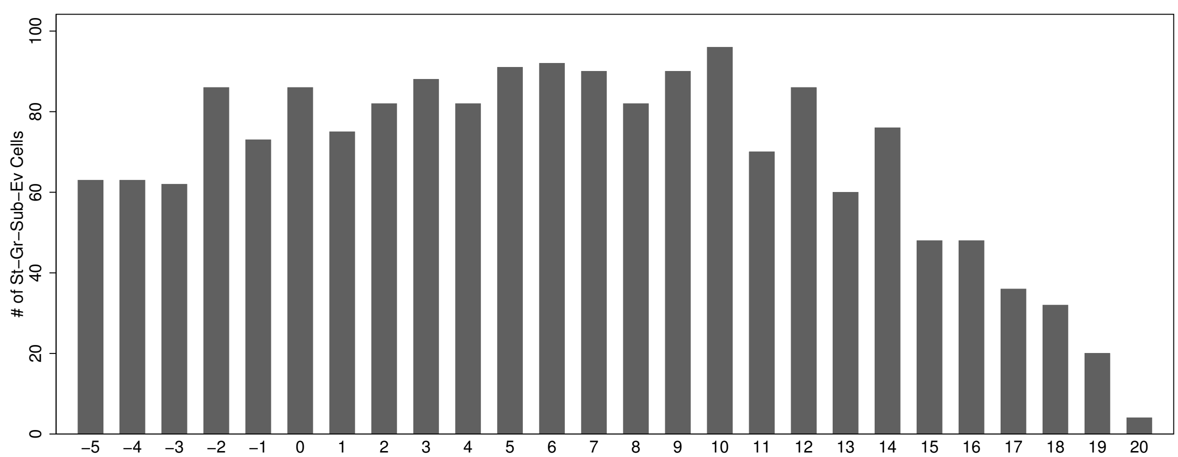
Tables A5 examines the racial composition of districts in different income quintiles. Black and Hispanic students and students from low socioeconomic backgrounds are somewhat concentrated in low income districts, but not overwhelmingly so. Reforms targeted to low income districts thus have limited ability to close achievement gaps between high and low income (or minority and non-minority) students. Table A6 shows parametric event study estimates of finance reforms on black-white and low income (free or reduced price lunch)-non low income funding gaps. These are computed by assigning to each student the average per-pupil funding in his/her district. The coefficients suggest small but positive post event effects (i.e. decreasing gaps), although only the coefficient on the black-white state funding gap is significant, and only at the 10 % level. See section VII in the text for further discussion.

## Appendix Figures

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



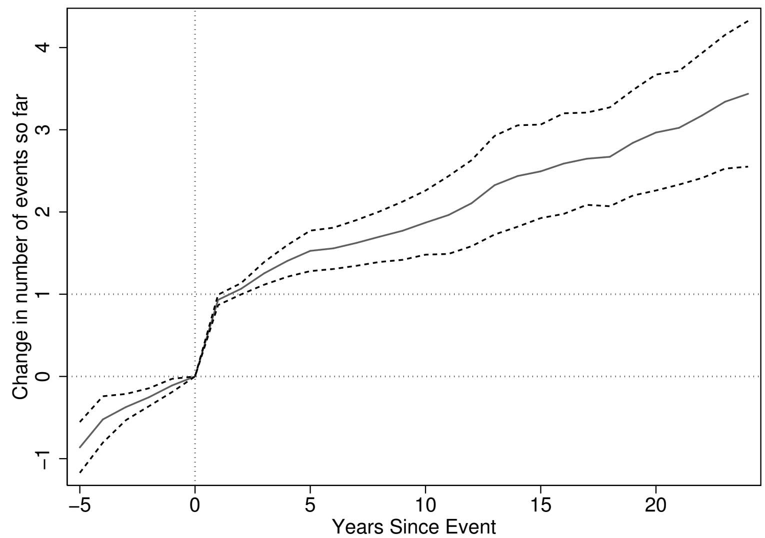

(a) State-year-event sample for finance analysis




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

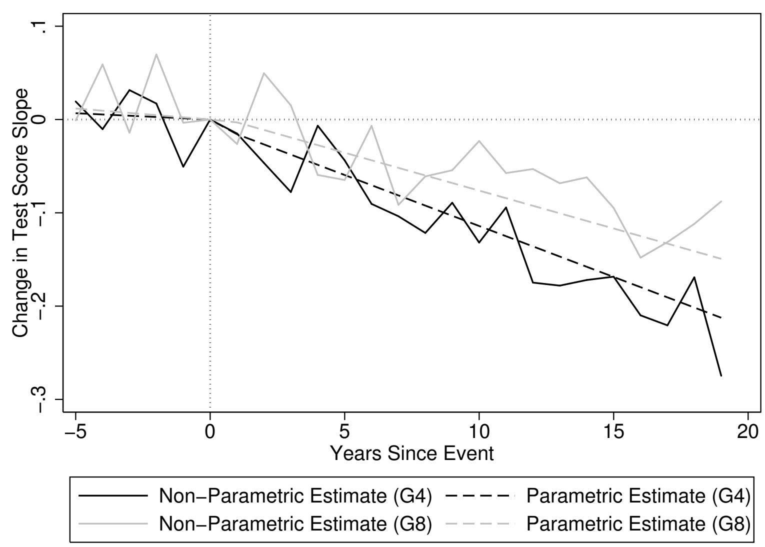
Notes: X-axis corresponds to "event-time" used in event study figures. States without events (there are 24) are not included in this figure. Panel A shows the number of state-event observations in the finance analysis. Panel B shows the number of state-subject-grade-event observations in the test score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

Figure A2: Event study of number of events to date




Notes: Dependent variable is the number of events to date. Nonparametric point estimates of event study time coefficients are shown, with 95% confidence intervals. Sample construction and fixed effects are identical to baseline specifications.

Figure A3: Event study estimates of effects of reform events on test score slope (G4 and G8 separately)




Notes: Dependent variables are slopes of 4th and 8th grade test scores w.r.t. log district income. Non-parametric and parametric estimates of event study coefficients (identical to specifications in figure 10) are shown for models run separately on 4th and 8th grade test scores.

## Appendix Tables

| Table A1. School Finance Events | | | | | | |
|---|---|---|---|---|---|---|
| | **Lafortune, Rothstein & Schanzenbach (through 2013)** | | **Hanushek & Lindseth (through 2005)** | **Jackson, Johnson, & Perisco (through 2010)** | **Corcoran & Evans (results through 2006)** | **Murray, Evans, & Schwab (through 1996)** |
| | *Court Order* | *Statute* | *Court Order* | *Court Order* | *Court Order* | *Court Order* |
| Alaska | 2007 | _ | Equity* | 1999 | _ | _ |
| Arizona | 1994<br>1997<br>1998 | | 1994<br>Equity* | 1994<br>1997<br>1998<br>2007 | _ | 1994 |
| Arkansas | 1994<br>2002<br>2005 | 1995<br>2007 | 2002<br>Equity* | 1994<br>2002<br>2005 | 2002<br>2005 | 1994 |
| California | _ | 1998<br>2004 | Equity* | 2004 | _ | _ |
| Colorado | _ | 2000 | _ | _ | _ | _ |
| Connecticut | 1996 | _ | Equity* | 1995<br>2010 | _ | 1995 |
| Idaho | 1993<br>2005 | 1994 | _ | 1998<br>2005 | _ | _ |
| Indiana | | 2011 | _ | _ | _ | _ |
| Kansas | 2005 | 1992<br>2005 | 2005<br>Equity* | 2005 | 2005 | _ |
| Kentucky | (1989) | 1990 | (1989) | (1989) | (1989) | (1989) |
| Maryland | 1996 | 2002 | _ | 2005 | _ | _ |
| Massachusetts | 1993 | 1993 | 1993 | 1993 | 1993 | 1993 |
| Missouri | 1993 | 1993<br>2005 | Equity* | 1993 | _ | _ |
| Montana | 2005 | 1993<br>2007 | 2005<br>Equity* | 1993<br>2005<br>2008 | 2005 | 1993 |
| New Hampshire | 1993<br>1997<br>2002 | 1999<br>2008 | 1997 | 1993<br>1997<br>1999<br>2002<br>2006 | 1997<br>1998<br>1999<br>2000<br>2002 | 1993 |
| New Jersey | 1990<br>1994<br>1998<br>2000 | 1990<br>1996<br>1997<br>2008 | 2002<br>Equity* | 1990<br>1991<br>1994 | 1990<br>1994<br>1997<br>1998 | 1990<br>1991<br>1994 |
| New Mexico | 1999 | 2001 | Equity* | 1998 | _ | _ |
| New York | 2003<br>2006 | 2007 | 2003 | 2003<br>2006 | 2003 | _ |

| | | | | | | |
|---|---|---|---|---|---|---|
| North Carolina | 1997<br>2004<br>2012 | 2012 | 2004 | 1997<br>2004 | 2004 | _ |
| North Dakota | _ | 2007 | _ | _ | _ | _ |
| Ohio | 1997<br>2000<br>2002 | 2000 | _ | 1997<br>2000<br>2002 | 1997<br>2000<br>2001<br>2002 | _ |
| Tennessee | 1993<br>1995<br>2002 | 1992 | Equity* | 1993<br>1995<br>2002 | 1993<br>1995<br>2002 | 1993<br>1995 |
| Texas | 1991<br>1992 | 1993 | Equity* | 1991<br>1992<br>2004 | 1991<br>1992<br>2005 | 1991<br>1992 |
| Vermont | 1997 | 2003 | 1997<br>Equity* | 1997 | 1997 | _ |
| Washington | 2010 | 2013 | _ | 1991<br>2007 | _ | _ |
| West Virginia | 1995<br>2000 | | _ | 1995 | _ | 1994 |
| Wyoming | 1995 | 1997<br>2001 | 1995<br>Equity* | 1995<br>2001 | 1995 | 1995 |

* No year given; plantiff victory

Table A2: Difference in log mean 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) |
| log(Dist avg. HH income) * Years Since Event | | -0.00277 | 0.00130 |
| | | (0.00328) | (0.00280) |
| Observations | 12527 | 12527 | 15576 |
| Event States | X | X | |
| All States | | | X |

Notes: Coefficients are reported for models with the long difference in district income (from 1990-2011) as the dependent variable. Standard errors are clustered at the state level.

Table A3: Alternative ways of handling event sample

(a) First event in each state

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Post Event | -460.8* | -194.9 | | 427.0 | 610.1** | |
| | (261.4) | (405.5) | | (316.8) | (245.1) | |
| Post Event * Yrs Elapsed | | | -0.00810** | | | 0.00492* |
| | | | (0.00332) | | | (0.00279) |
| Observations | 1029 | 1029 | 1498 | 1064 | 1064 | 1569 |
| p, total event effect=0 | 0.084 | 0.633 | 0.019 | 0.184 | 0.016 | 0.084 |
| State-Event FEs | X | X | X | X | X | X |
| Year FEs | X | X | | X | X | |
| Sub-Gr-Yr FEs | | | X | | | X |

(b) Court events only

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Post Event | -312.8** | -655.2** | | 87.07 | 130.2 | |
| | (127.9) | (270.8) | | (153.3) | (168.8) | |
| Post Event * Yrs Elapsed | | | -0.00885* | | | 0.00965** |
| | | | (0.00478) | | | (0.00374) |
| Observations | 861 | 861 | 1249 | 859 | 859 | 1253 |
| p, total event effect=0 | 0.023 | 0.025 | 0.078 | 0.576 | 0.449 | 0.017 |
| State-Event FEs | X | X | X | X | X | X |
| Year FEs | X | X | | X | X | |
| Sub-Gr-Yr FEs | | | X | | | X |

Table A3: (continued) Alternative ways of handling event sample

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

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Post Event | -563.9** | -317.7 | | 559.0** | 694.6*** | |
| | (250.0) | (353.0) | | (269.1) | (207.6) | |
| Post Event * Yrs Elapsed | | | -0.00771** | | | 0.00529* |
| | | | (0.00362) | | | (0.00279) |
| Observations | 1890 | 1890 | 2743 | 1924 | 1924 | 2820 |
| p, total event effect=0 | 0.029 | 0.373 | 0.039 | 0.043 | 0.002 | 0.064 |
| State-Event FEs | X | X | X | X | X | X |
| Year FEs | X | X | | X | X | |
| Sub-Gr-Yr FEs | | | X | | | X |

(d) Number of events to date

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Num Events to Date | -289.6*** | -180.7 | | 363.1** | 337.0** | |
| | (104.3) | (169.1) | | (144.3) | (129.6) | |
| Num Event-Years to Date | | | -0.00283*** | | | 0.00281*** |
| | | | (0.00101) | | | (0.000806) |
| Observations | 1029 | 1029 | 1498 | 1064 | 1064 | 1569 |
| 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 |

Notes: Tables A3 and A4 report variations on baseline specifications from columns 1 and 4 of tables 3, 4 (both panels), column 1 of table 6, and column 5 of table 7. State-event and year fixed effects are included in finance models, and state-event and grade-subject-year fixed effects are included in NAEP models. See notes to tables 3, 4, 6, and 7 for further specification details. In Table A3, panel (a), the sample is restricted to include only the first event in any state. In panel (b), only court ordered reform events are included in the estimation. Panel (c) shows estimates where states with multiple events are reweighted by $1/n$, where $n$ is the number of events in that state. In panel (d), results are shown for an alternative parametric model where the independent variable is either the number of years to date (in the finance analyses) or the sum across all events to date of the number of years elapsed since the event (in the NAEP analyses).

Table A4: Alternative income measures

(a) 2010: ln(district income)

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Post Event | -384.4** | -222.1 | | 410.0* | 394.7** | |
| | (172.1) | (225.5) | | (214.3) | (149.5) | |
| Post Event * Yrs Elapsed | | | -0.00977*** | | | 0.00829*** |
| | | | (0.00286) | | | (0.00211) |
| Observations | 1890 | 1890 | 2743 | 1924 | 1924 | 2828 |
| p, total event effect=0 | 0.030 | 0.330 | 0.001 | 0.061 | 0.011 | 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: ln(housing values)

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Post Event | -331.8** | -214.1 | | 517.5** | 534.4*** | |
| | (141.8) | (214.2) | | (210.5) | (179.5) | |
| Post Event * Yrs Elapsed | | | -0.00717*** | | | 0.00882*** |
| | | | (0.00201) | | | (0.00262) |
| Observations | 1890 | 1890 | 2743 | 1924 | 1924 | 2826 |
| p, total event effect=0 | 0.023 | 0.322 | 0.001 | 0.017 | 0.004 | 0.001 |
| State-Event FEs | X | X | X | X | X | X |
| Year FEs | X | X | | X | X | |
| Sub-Gr-Yr FEs | | | X | | | X |

(c) 1990: share < 185% poverty line

| | Slopes | | | Q1-Q5 Means | | |
|---|---|---|---|---|---|---|
| | (1) St. Rev. | (2) Tot Rev. | (3) NAEP | (4) St. Rev. | (5) Tot Rev. | (6) NAEP |
| Post Event | 807.3** | 572.8 | | -566.1* | -466.7** | |
| | (386.3) | (447.6) | | (286.8) | (206.0) | |
| Post Event * Yrs Elapsed | | | 0.0225*** | | | -0.00608** |
| | | | (0.00840) | | | (0.00268) |
| Observations | 1890 | 1890 | 2743 | 1911 | 1911 | 2793 |
| p, total event effect=0 | 0.042 | 0.207 | 0.010 | 0.054 | 0.028 | 0.028 |
| State-Event FEs | X | X | X | X | X | X |
| Year FEs | X | X | | X | X | |
| Sub-Gr-Yr FEs | | | X | | | X |

Notes: Tables A3 and A4 report variations on baseline specifications from columns 1 and 4 of tables 3, 4 (both panels), column 1 of table 6, and column 5 of table 7. State-event and year fixed effects are included in finance models, and state-event and grade-subject-year fixed effects are included in NAEP models. See notes to tables 3, 4, 6, and 7 for further specification details. In table A4, panel (a), 2010 log mean income is used to measure district resources (in place of 1990 log mean income in prior specifications). In panel (b), the 1990 log mean value of owner-occupied homes is used. In panel (c), the resource measure is the share of individuals in the district in 1990 with incomes below 185% of the federal poverty line.

Table A5: Fraction in each district income quintile

| | Q1 | Q2 | Q3 | Q4 | Q5 |
|---|---|---|---|---|---|
| Black | 0.242 | 0.241 | 0.223 | 0.169 | 0.125 |
| Black/Hispanic | 0.244 | 0.229 | 0.243 | 0.167 | 0.116 |
| White | 0.196 | 0.191 | 0.182 | 0.202 | 0.228 |
| Free/reduced-price lunch | 0.315 | 0.217 | 0.199 | 0.158 | 0.110 |

Notes: Table shows proportion of each racial/socioeconomic group in each district income quintile. The numbers in each row (i.e. across the 5 columns) add up to one.

Table A6: Event studies for per pupil revenue gaps

| | St. Rev | | Tot. Rev | |
|---|---|---|---|---|
| | Black/White | Free Lunch | Black/White | Free Lunch |
| Post Event | 277.1* | 65.14 | 221.1 | 86.27 |
| | (147.3) | (210.4) | (166.6) | (160.7) |
| Observations | 1810 | 1624 | 1810 | 1624 |

Notes: The *Post Event* coefficient shows estimated event effect from parametric event study model without controlling for prior trends. Per pupil gaps in state revenue are shown in columns (1) and (2), whereas columns (3) and (4) show gaps in total revenues. State-event and year fixed effects are included. Standard errors are clustered at the state level.