School Finance Reform and the Distribution of Student Achievement*

February 2017

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 absolute and relative spending and achievement in low-income school districts. Using an event study research design that exploits the apparent randomness of reform timing, we show that reforms lead to sharp, immediate, and sustained increases in spending in low-income school districts. Using representative samples from the National Assessment of Educational Progress, we find that reforms cause increases in the achievement of students in these districts, phasing in gradually over the years following the reform. 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, Zach Liscow, Richard Rothstein, Max Schanzenbach, and conference and seminar participants at APPAM, AEFP, Bocconi, Brookings, Chicago, Erasmus, Wisconsin (IRP), LSE, New York University, Northwestern, Princeton, RAND, Teachers' College, Texas A&M, Warwick, and the 2015 Stavanger-Bergen-Berkeley workshop for helpful comments and discussions.

Economists have long been skeptical of resource-based education policies, based in part on observational studies showing small or zero effects of additional funding (see, e.g., Coleman et al. 1966, Hanushek 1986, Hanushek 2006).[2] Hanushek, for example, writes: "Simply providing more funding or a different distribution of funding is unlikely to improve student achievement (even though it may affect the tax burdens of school financing across the citizens of a state)" (1997, p. 153). Accordingly, recent policy discussions have focused on ways to improve the productivity of existing inputs rather than on changes in school resource levels.

Nevertheless, states have continued to implement aggressive resource-based policies, aimed in part at reducing achievement gaps. Figure 1 shows the evolution of average revenues per pupil, in 2013 dollars, in the lowest- and highest-income school districts in each state (defined as the bottom and top fifths of the state's district-level mean household income distribution).[3] Between 1990 and 2012, real per-pupil revenues rose by roughly 30% in the highest-income districts, and by over 50% in the lowest-income districts. Thus, while low-income districts collected about 20% less than high-income districts in 1990, they have been in rough parity since around 2001.

Much of this change came via reforms to state education funding formulas, many implemented in response to court orders. Figure 2 shows revenues of low-income districts relative to high-income districts, each defined as in Figure 1, separately for the 26 states that have implemented—or at least been ordered to implement by courts—school finance

---

[2] There are also observational (Card and Krueger 1992a) and experimental (Krueger 1999; Dynarski, Hyman & Schanzenbach 2013) studies pointing to positive school resource effects. There is no consensus about how to reconcile these (see, e.g., Burtless 1996; Hanushek 2003; Krueger 2003).

[3] Hawaii and the District of Columbia are excluded. Districts are weighted by log enrollment in computing state quintile means, which are then averaged without weights in Figure 1. We discuss data sources and definitions in Section III.

reforms since 1990 and for 23 states that have not. Growth in low-income districts' relative revenues has been more than twice as rapid in the former states than in the latter.

There are two primary types of school finance reforms (SFRs). In the 1970s and 1980s, SFRs were primarily "equity" reforms, aimed at reducing resource disparities across districts. Since 1990, the pace of reforms has quickened, and most have been "adequacy" reforms, aimed at achieving sufficient funding in low income districts regardless of implications for equity.[4]

SFRs are arguably the most substantial national policy effort aimed at promoting equality of educational opportunity since the turn away from school desegregation in the 1980s. But there is little evidence about their effects on student achievement. What evidence there is derives from non-representative data on students who took the SAT college entrance exam (Card and Payne 2002); from long-run outcomes measured in the relatively small Panel Study of Income Dynamics sample (Jackson, Johnson, and Persico 2016); or from case studies of individual reforms (Clark 2003; Hyman 2013; Guryan 2001).[5] These studies primarily examine pre-1990, equity-based SFRs, and generally find positive effects on student outcomes. But funding levels were much higher by 1990 than earlier, and the most severe inequities in school resources had been addressed. Thus, there may have been less scope for more recent, adequacy-based SFRs to benefit students.

The impacts of SFRs on student achievement are closely related to the impact of additional resources. The literature regarding whether "money matters" in education (Card

[4] Studies of the implications of SFRs for school finance, mostly examining equity reforms, include Hanushek and Lindseth (2009); Card and Payne (2002); Murray, Evans, and Schwab (1998); Ladd and Goertz (2015); Sands (2015); and Berry and Wysong (2012).
[5] Cascio and Reber (2013) and Cascio, Gordon, and Reber (2013) examine the introduction of federal Title I funding to low-income schools via the 1965 Elementary and Secondary Education Act.

and Krueger 1992a; Hanushek 1986, 2003, 2006; Burtless 1996) is contentious and does not offer clear guidance. State funding formulas are the main policy tool available to address inequities in academic outcomes, so funding shifts deriving from changes in these formulas are the most policy-relevant variation in school resources.

We provide the first evidence from nationally representative data regarding the impact of SFRs on student achievement. We exploit little-used data from the National Assessment of Educational Progress (NAEP), also known as "the Nation's Report Card." State-representative samples of 100,000-200,000 students in the fourth and eighth grades have taken math and reading tests every two to four years since 1990. Importantly, the tests have been uniform across states and over time, facilitating comparisons.

We use the NAEP data to construct a state-by-year panel of relative achievement in low-income school districts, covering 1990 to 2011. Conveniently, the beginning of our NAEP panel coincides with the onset of the adequacy era of school finance, which dates to the 1990 Kentucky Education Reform Act (KERA).

To distinguish the causal impacts of SFRs from other potential determinants of spending and test score trends, we use an event study framework, taking advantage of plausibly random variation in the location and timing of post-1990 SFRs.[6] We find no sign of systematic changes in either funding or test scores in the period leading up to a reform, supporting our assumption that reform timing is exogenous. Following reforms, we document sharp increases in state revenues, with larger increases in low-income districts

---

[6] A simple long-difference analysis of test score gaps between low-income and high-income districts, similar to the analysis of finance in Figure 2, shows that gaps have shrunken in states that implemented reforms relative to states that have not. See Figure A2 in the online Appendix.

and smaller but still positive increases in high-income districts.[7] These changes occur quickly after reform events, persist for many years, and are not offset by reductions in local revenues. Absolute and relative funding in low-income districts rises by approximately $1,200 and $700 per pupil per year, respectively. We find that, on average, schools use the additional funds on instructional spending, to reduce class size, and for capital outlays.

We also find clear changes in achievement trends following events. These cumulate over subsequent years: Ten years after a reform, relative achievement of students in low-income districts has risen by roughly 0.1 standard deviation, approximately one-fifth of the baseline gap between high- and low-income districts. The implied impact is between 0.12 and 0.24 standard deviations per $1,000 per pupil in annual spending. This is at least twice the impact per dollar that is implied by the Tennessee Project STAR class size experiment.[8] Given existing estimates of the relationship between test scores and students' subsequent earnings, our results imply that a $1 increase in funding to low-income school districts will raise students' eventual earnings by more than $1 in present value.

Nevertheless, we find no discernable effect of reforms on statewide achievement gaps between high- and low-income students or between minority and white students. This is not inconsistent with our results on the impacts on scores in low-income districts, nor does it indicate that only the high-income students in those districts benefit. Rather, we show that low-income and minority students are not very highly concentrated in school districts with low mean incomes. As a result, SFRs lead to only small increases in the

[7] Anecdotally, legislators facing court orders to increase funding to low-income districts often respond by increasing overall funding, as a way of disguising the resulting redistribution. Reforms are associated with sharp increases in total state education expenditures and tax collections.

[8] STAR raised costs by about 30% in K-3, and raised early grade test scores by 0.17 SDs (Krueger and Whitmore 2001; Krueger 1999, 2003). Current spending per pupil in Tennessee is around $9,000, so comparable proportional class size reductions would cost around $2,700 per pupil per year. The implied effect is thus around 0.06 SDs per $1,000 per (early elementary) pupil per year.

funding to which the average low-income or minority student is exposed. Thus, while our analysis suggests that finance reforms can be quite effective at reducing between-district inequities, other policy tools aimed at closing *within*-district achievement gaps will be needed to address overall equity concerns.

I. School finance reforms[9]

Historically, American public schools were locally managed and financed primarily via local property taxes. As school districts vary widely in both their tax bases and their voters' willingness to tax themselves to fund schools, this meant that school spending and quality varied substantially across districts.

In the 1960s, a group of legal scholars argued that local school finance violates federal and state constitutional provisions that guarantee equal access to public services (see, e.g., Wise 1967; Horowitz 1966; Kirp 1968; and Coons, Clune, and Sugarman 1970). Advocates brought and won suits in many states demanding more equitable school finance systems; in other states, legislatures acted without court decisions, often to stave off potential rulings.[10] The resulting finance regimes often involved substantial increases in state transfers to districts with low property tax bases. An extensive "fiscal federalism" literature examines the effects of these reforms on the distribution of school funding (see, e.g., Hanushek and Lindseth, 2009; Corcoran and Evans, 2015; Card and Payne, 2002; Murray, Evans, and Schwab, 1998).

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

[10] The U.S. Supreme Court held in 1973 that education is not a fundamental right under the U.S. Constitution (*San Antonio Independent School District v. Rodriguez,* 411 US 1, 1973). Subsequent suits focused on state constitutions, which often mandate adequate and/or equitable systems of public education.

We focus on a second wave of finance reforms, which began with a 1989 Kentucky Supreme Court ruling that the state constitution, which as in many other states dictates an "efficient system" of public schools, requires that "[e]ach child, *every child*, … must be provided with an equal opportunity to have an adequate education" (*Rose v. Council for Better Education*[11]; emphasis in original). The Court emphasized that equal funding was not sufficient, and articulated a standard closer to equality of outcomes for students in low-income districts ("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"). The Kentucky legislature responded with the Kentucky Education Reform Act of 1990 (KERA), which revamped the state's educational finance, governance, and curriculum. Clark (2003) and Flanagan and Murray (2004) find KERA substantially increased spending in low-income districts.

Since 1990, courts in many other states have found adequacy requirements in their own constitutions. In many cases reforms have aimed at *higher* spending in low-income than in high-income districts, to compensate for the out-of-school disadvantages that low-income students face.[12]

We have attempted to identify all major SFRs between 1990 and 2011. We began with lists of court-ordered reforms compiled by Jackson et al. (2016) and Corcoran and Evans (2015). We supplemented these with our own research into case histories, and updated them through 2011. We also tabulated major legislative SFRs. In some important

[11] 790 SW 2d 186. *Rose* was not the first adequacy ruling, but earlier rulings attracted less attention.
[12] A small industry has developed to calculate the spending level needed to satisfy an adequacy standard. See, e.g., Downes and Steifel (2015) and Duncombe, Nguyen-Hoang, and Yinger (2015). Sims (2011a) and Corcoran and Evans (2015) contrast fiscal effects of adequacy and equity reforms. Each relies on a sample ending in 2002, early in the adequacy era.

cases (e.g., Colorado, California), legislatures reformed finance systems without prior court decisions, often to forestall adverse judgments in threatened or ongoing lawsuits. Our primary analyses include these, though we also present results that focus exclusively on court orders. Some of the reforms were accompanied by governance, curriculum, or accountability changes, though our assessment is that these additional changes were typically not very important or impactful.

Appendix Table A1 presents a complete list of our events and compares it to those used in other studies. We identify a total of 64 school finance reform events in 26 states between 1990 and 2011.[13] 39 (61%) involve court orders; the remainder are legislative actions without a major court order in the same year. States with events are quite geographically diverse, though reforms are rare in the Deep South and upper Midwest.

18 states had multiple events in our period. These were generally closely spaced: 60% were three or fewer years apart. In these cases, we suspect that only one generated a major change in the state's finance rules and that others were procedural steps (e.g., court orders that were disregarded or legislation changes that were later found inadequate). Our analytical strategy is built with this idea in mind, though our results are robust to alternative models of the impact of multiple reform events in the same state.

II. Analytic approach

To identify the causal effect of school finance reforms, we leverage variation in the timing of reform events in 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

[13] Our panel excludes the 1989 *Rose* decision but includes KERA, the legislature's response in 1990.

have events in that year, after accounting for fixed differences between the states and for common time effects. The key assumption is that the exact timing of events is as good as random. We think this is plausible, given the idiosyncrasies of judicial processes. An attractive feature of our approach is that it builds in placebo tests that should identify likely violations of this assumption.

Our simplest event study specification models events as permanent, immediate shifts in outcomes relative to other states:

$$(1) \qquad \theta_{st} = \delta_s + \kappa_t + 1(t > t_s^*)\beta^{jump} + \varepsilon_{st}.$$

Here, θ_{st} represents some summary of the distribution of funding or achievement in state s in year t . We discuss our particular measures below. δ_s and κ_t represent state and year effects, respectively. t_s^* is the date on which state s's event occurred. (For now, we assume that each state has just one event; this term is set to zero for states without events.) The coefficient estimate β^{jump} represents the change in the outcome following the event. In all of our analyses, we use standard errors that are clustered at the state level to allow for arbitrary dependence of ε_{st} across t within s.

SFRs may not affect θ_{st} immediately, but may develop more gradually. This is particularly true for student achievement outcomes, as the achievement of a student in year t likely depends in part on the quality of the schooling she received in prior years. In addition, if event timing is non-random, states with events may diverge from states without events even before the date of the event. To accommodate these ideas, we add two trend terms to (1):

$$(2) \qquad \theta_{st} = \delta_s + \kappa_t + 1(t > t_s^*)\beta^{jump} + 1(t > t_s^*)(t - t_s^*)\beta^{phasein} + (t - t_s^*)\beta^{trend} + \varepsilon_{st}.$$

$\beta^{phasein}$ captures delayed event effects and represents the annual change in outcomes in state s after t_s^*, relative to the same state prior to the event. β^{trend}, which is identified from changes in s relative to other states in years *prior* to t_s^*, represents a falsification test: $\beta^{trend} \neq 0$ would indicate that event timing is meaningfully non-random.

We also estimate non-parametric models that do not constrain the phase-in and prior trend effects to be linear:

$$(3) \qquad \theta_{st} = \delta_s + \kappa_t + \sum_{r=k_{min}}^{k_{max}} 1(t = t_s^* + r)\, \beta_r + \varepsilon_{st}.$$

Here, β_r represents the effect of an event in year t_s^* 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 β_{-5} represents average outcomes five or more years prior to an event, relative to those in the event year.

Comparisons of the parametric and non-parametric estimates indicate that the simple specification (2) does a good job of capturing dynamics in finances and student achievement surrounding events, though the post-event "jump" is sometimes spread out over a few years following the event. In only one of the specifications that we estimate do we reject the null hypothesis that the pre-event coefficients (β^{trend} in (2) and $\{\beta_{-k}, \ldots, \beta_{-1}\}$ in (3)) are all zero, and in this case it appears to be an idiosyncratic blip in a single β_{-r} coefficient (see Figure 7, below). This supports our identifying assumption.

When we examine finance outcomes, all of the post-event effect appears to be nearly immediate, so we focus on the simpler specification (1). By contrast, in our student achievement analysis, the "jump" is never distinguishable from zero, and all of the effect that we estimate operates through the $\beta^{phasein}$ coefficient. We thus emphasize specifications that allow for a phase-in effect but no post-event jump. In each case, these simple specifications fit the non-parametric results quite well.

Our event study methodology is a form of difference-in-differences (DD). The identifying assumption is that without finance reforms, outcomes would have moved in parallel in treated and untreated states. While we view this as plausible, it may not be correct (Hanushek et al. 1996a,b). We can weaken the assumption by shifting our focus from the absolute level of test scores to the relative scores of different students in the same states. Given the emphasis in adequacy rulings on districts serving disadvantaged students, a natural contrast is between students in high- and low-income districts. When we use as a dependent variable the gap in test scores between low-income and high-income districts in a state, the event study strategy is robust to arbitrary state-by-year shocks to achievement, so long as they have similar effects on districts at different income levels. The identifying assumption is that the *relative* outcomes of low-income districts would have followed parallel trends across states in the absence of SFRs.

We consider two measures of relative outcomes in low-income districts. First, we use the gap between districts in the top and bottom quintiles of the state income distribution. These quintile gaps can be noisy, in part because they discard information on the middle 60% of districts. We thus emphasize a second measure, the slope of district-level outcomes with respect to log average income across all districts in the state.[14] A more negative slope corresponds to higher relative outcomes in low-income districts. For both

[14] Specifically, we regress district-level spending per pupil or mean achievement on log mean income, controlling for log enrollment. The regression is estimated separately for each state and year, and in achievement models for each subject and grade. The district log income coefficients are used as θ_{st} for subsequent analyses at the state-year-(subject-grade) level. See the Appendix for further detail.

finance and achievement outcomes, the slope and quintile gaps are highly (negatively) correlated, and all of our results are robust to the choice of relative outcome measure.

Event studies with multiple events

Many states had multiple events (court orders or legislation) over our period. Unfortunately, there is no accepted strategy for conducting event studies with multiple events per unit. Our primary estimates are based on a single event in each state. The intuition here is that when states have multiple events, they often represent jockeying between the legislature and the courts with only minor changes in school finance until the legislature finally enacts a major reform, and then continued jockeying afterward as advocates continue to push for additional changes. To identify the most consequential reform, we use data on state aid to districts to identify a regime change in the progressivity of a state's finance system, relying on methods for the identification of change points in time series data (e.g., Bai 1997; see also Card, Mas, and Rothstein 2008). We then use that as the date of the event for our analyses of student achievement.

Specifically, let θ_{st} be our slope measure of the progressivity of state aid. For each state and each potential event date t_s^* -- that is, each year that we observe a major court order or legislative change – we estimate a time series regression using as the only explanatory variable an indicator for observations after that date:

$$(4) \quad \theta_{st} = \alpha + 1(t > t_s^*)\kappa + \varepsilon_{st}.$$

We select the event date that yields the largest t statistic for κ – or, equivalently, the smallest mean squared error – for this time series regression.[15] We treat the selected date as the single event in state s.

Bai (1997) shows that if there really is a structural break in the time series (with a non-zero true κ) this method is super-consistent for the location of the break, permitting inference regarding κ to treat its location as known. However, in the event that there is no structural break (i.e., that each court order and legislative change in the state was ineffective, with $\kappa = 0$), our method will nevertheless pick one of the potential events. This could lead us to overstate the effect of a true reform on the progressivity of state aid. Our main outcome, however, is student achievement, and we do not use achievement data in selecting events. Thus, the potential inclusion of some non-reforms in our event study analysis might lead us to understate the effect of a true SFR on student achievement, since our estimates would combine the effects of true reforms with those of spurious non-events.

We also present estimates from two additional approaches to multiple events. One includes all events, without judgment about their relative importance. To implement this approach, we create a separate copy of the time series for the state for each apparent event, using a different value of t_s^* for each copy. We then stack the copies, replacing the state effects in equations (1)-(3) with state-by-event effects.[16] In Monte Carlo simulations (see Appendix), this method works well to identify the average effect of events both when each event has the same effect and when only one event in a state has a non-zero effect. Our final approach follows the prior literature—which generally emphasizes simple specifications

[15] We restrict attention to t^* for which the estimated κ has the expected sign.
[16] Results are unchanged when data are reweighted to offset the overrepresentation of states with multiple events.

analogous to (1)—by focusing on the initial court order in each state, even if this was not implemented for many years. Here, we treat states without court orders as untreated, though in some cases they saw legislative reforms. Results are extremely similar across all three methods. Accordingly, we do not view multiple events as a major issue in practice.

III. Data

Our analysis draws on data from several sources. We begin with our database of state SFR events, discussed above. We merge this to district-level finance data, from the National Center for Education Statistics' (NCES) annual census of school districts and the Census of Governments; mean household income by district from the 1990 Census; and the NAEP achievement measures, aggregated to the district-year level.

The district finance data report enrollment, revenues and expenditures annually for each local education agency.[17] We convert all dollar figures to 2013 dollars per pupil, and exclude very small districts and those with highly volatile enrollment or implausible per-pupil funding. Details are in the appendix.

We construct student achievement measures from the restricted-use "State NAEP" microdata. The state NAEP began in 1990, with 42 states participating. It has been administered roughly every two years since. Since 2003, all states have participated in 4th and 8th grade assessments in math and reading in every odd-numbered year.[18] Table 1 shows the schedule. Tests are administered to around 100,000 students (more in later years) in each subject-grade-year. These consist of representative samples of about 3,500 students per state, spread across about 140 schools in 80 districts.

---

[17] Census data are available in 1989-90 and 1991-92, and annually since 1994-95. We use samples from the Census Bureau's Annual Survey of Government Finances for 1992-93 and 1993-94.

[18] The NAEP also tests 12th graders, but samples are smaller, and other subjects.

The NAEP uses a consistent scoring scale across years for each subject and grade in order to permit time-series comparisons. 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.[19]

Table 2 presents district-level summary statistics, pooling data from 1990-2011. The rightmost columns show means for districts in the top (Q5) and bottom (Q1) quintiles by average family income in each state.

IV. Finance reforms and school finance

We begin our empirical analysis by documenting the implications of SFR events for school finance. We use the approach discussed in Section II to select a single SFR event that best explains the time series of the state aid – log district income slope in each state.

Figure 3 graphs event study results for state transfers per pupil in the lowest-income (Q1) quintile of districts. We present several plots of this basic form. The solid line represents estimates from the non-parametric event study specification (3), while dotted lines show pointwise 95% confidence intervals. The dashed line shows the parametric specification (2). There is a small upward trend in state revenues prior to the finance reform events, but this is not statistically significant in either the parametric or the nonparametric specification. Following reforms, state revenues increase substantially, by roughly $1,300 in the 4th post-event year. Though out-year estimates are noisy, impacts appear to persist through the end of our sample. Figure 4 repeats the same analyses for the

[19] The pre-2000 NAEP data do not use the same district codes as the CCD. We are grateful to Bruce Kaplan, Kate Pashley, and Fatih Unlu for their assistance in locating the crosswalk from the older NAEP data to schools and districts.

highest income (Q5) districts. Estimated changes in funding following reforms are much smaller here; while the nonparametric post-event effects are jointly significant, the parametric estimates are not and in any event the magnitudes are quite small

We report coefficients from our parametric specifications for state revenues in the lowest and highest income districts in columns 1 and 2 of Table 3; column 3 shows estimates for average revenues across all districts for comparison. In Panel A, we report the simple specification (1), while Panel B adds the pre-event and post-event trends from specification (2). (It is these that are shown in Figures 3 and 4.) The former indicates that average state funding rises by $1,225 following events in first quintile districts and by $527 (not significant) in fifth quintile districts. The upward trends preceding events seen in Figures 3 and 4 are reflected in the point estimates in Panel B, but are small and not distinguishable from zero. Similarly, point estimates indicate that the post-event jumps fade slightly over subsequent years, but these trends are again small and insignificant.

Panels C and D of Table 3 repeat the specifications from Panels A and B, this time taking total district revenues, inclusive of state aid and other revenues, as the dependent variable. These are quite similar to those for state revenues in both low- and high-income districts. There is no indication that declines in local revenues offset increases in state funding in low income districts, nor in (panel D) of pre-trends or erosion of initial impacts. The more flexible nonparametric specifications (Appendix Figure A3) are also similar.

In additional analyses of state budgets (Appendix Table A2), we have found no indication that growth in educational spending following events crowds out state spending on other programs; rather, SFRs are associated with increases in state tax collections large enough to fully fund the increase in state transfers to districts.

As noted above, our analysis of student achievement impacts of SFRs focuses on contrasts between low- and high-income districts, to abstract from unrelated shocks to overall average achievement that might be correlated with the timing of these reforms. Columns 4 and 5 of Table 3 show estimates for these contrasts, first using the difference in funding between bottom- and top-quintile districts (column 4) and then the slope of funding with respect to log district income (column 5; this is shown graphically in Figure 5). Using each measure, we see sharp increases in relative state funding for low-income districts following events that show no sign of eroding thereafter. In no case is there any sign of a pre-event trend that would suggest a violation of our quasi-random timing assumption, nor is there any sign that increased progressivity of state aid is offset by local revenues.[20]

Table 3 makes clear that SFRs are associated with large increases in funding in low-income school districts. A natural question is how the additional funds are spent. Table 4 presents event-study coefficients from our simple model (1) for per-pupil revenues and spending in various categories. There is no apparent impact of SFRs on local or federal revenues. We see substantial impacts of SFRs on average instructional spending, both overall and in Q1 districts (columns 2 and 3). We also see effects on teachers per pupil and total teacher salaries but not on average teacher pay, suggesting that districts use additional funds to reduce class size.[21] Finally, we see large effects on non-instructional expenditures, particularly capital outlays.

---

[20] When we estimate specifications similar to Card and Payne's (2002) closely related analysis of earlier SFRs (Appendix Table A3), estimated SFR effects are slightly larger but imprecise, and well within the earlier confidence intervals. Where Card and Payne find that total revenues rise by about $0.50 per extra $1 in state aid, our estimates indicate much more stickiness for the recent reforms.

[21] Using a different research design, Sims (2011b) finds effects of SFRs on teacher pay.

Columns 4 and 5 show results for *relative* spending in low-income districts. Little of the increase in relative funding goes to instructional expenditures, while roughly half goes to capital spending. The capital spending effect is not surprising; many lawsuits specifically concern dreadful conditions in low-income schools, and SFR remedies often created funds to support renovation of schools in poor shape.[22]

V. Finance reforms and district-level student achievement

The above results establish that reform events are associated with sharp, immediate improvements in the progressivity of school finance, with absolute and relative revenue increases in low-income school districts. We now turn to our main analysis, examining the effect of SFRs on student achievement.

Where the θ_{st} school finance measures formed a state-by-year panel, for test scores we have two additional dimensions: Grade and subject. We replace the year fixed effects (κ_t) in (1)-(3) with subject-grade-year effects. These capture any differences in tests between administrations, as well as changes in student performance by grade and/or subject that are common across states. To avoid confounding from state-level shocks, we focus on DDD specifications that use the achievement gap between low- and high-income districts as the dependent variable.

Sharp, permanent changes in funding, if used productively, should increase the flow of educational services. Achievement is cumulative, so these services are unlikely to have immediate impacts on test scores, but should raise scores gradually as students are exposed for longer. Effects should grow at least until students have been exposed to the

[22] Neilson and Zimmerman (2014) find that school reconstruction causes increases in student achievement. Cellini et al. (2010) and Martorell, Stange, and McFarlin (2015) fail to find significant effects, but each study is under-powered to detect effects of plausible magnitude.

new funding levels for their entire careers. They may even continue to grow beyond this point. For example, consider a state that responds to a court order by creating a new permanent facility to fund several school renovation and construction projects each year. Initially, only a few students benefit, but over time growing shares of students are exposed to funded projects. Insofar as better facilities promote student learning, achievement effects would continue to grow until several years after the last project is complete, potentially decades after the initial policy change. We thus emphasize the phase-in coefficient from equation (2) as the primary measure of SFR effects on test scores.

Figure 6 presents our event-study analysis of the slope of achievement with respect to district income. Recall that improvements in the relative achievement of students in low-income districts reduce this slope. As before, we present non-parametric results (equation 3) as a solid line and estimates of our three-parameter model (equation 2) as a dashed line. As before, there is no indication of a differential trend in reform states prior to events. Following events, the non-parametric series does not react immediately, but begins trending noticeably downward starting in about the fifth post-event year (though the immediate trend break encoded in (2) fits the data nearly as well). The downward trend continues through the end of our sample.[23]

Table 5 presents the parametric estimates. We begin in Column 1 with our three-parameter model, as shown in Figure 6. The estimated pre-event trend is essentially zero and the post-event jump is also small, but the post-event change in trend is large and statistically significant. Column 2 presents a specification that discards the other two coefficients. Results are quite similar. The estimated change in the slope is -0.010 per year.

[23] The sawtooth pattern at the end of the sample likely reflects the biannual NAEP testing schedule.

This implies that each year after an event, a district with log mean income one unit (about two-thirds) below the state average sees its scores rise relative to the state average by 0.010 standard deviations, accumulating to 0.10 SDs over ten years. This is quantitatively meaningful – on average in our sample the slope of test scores with respect to log income is 0.96 so SFRs reduce this gradient by approximately one-tenth within ten years.

As discussed above, the pattern of gradually growing effects in Figure 6 is consistent with a view of achievement as a stock reflecting accumulated past input flows. 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.[24] This may reflect the use of some additional funds for durable investments, as discussed above. We do not have enough precision, however, to rule out a flattening of the effect at the expected time.

Figures 7 and 8 present estimated test score impacts for the lowest- and highest-income districts, respectively. The effects on the income gradient are driven by dramatic increases in test scores in the lowest-income districts.[25] In higher-income districts, there is little sign of a systematic post-event change. Parametric estimates are shown in Columns 3 and 4 of Table 5; Column 5 shows that the impact of events on the test score gap between bottom- and top-quintile districts is 0.008 SDs per year, or 0.013 SDs in the more flexible model (column 6). The gap in mean log incomes between the top and bottom quintiles averages 0.65, so the quintile point estimate is a bit larger than what we obtain for our

[24] We have estimated separate non-parametric models for 4th and 8th grade scores. Both sets of effects grow roughly linearly through the end of our panels. See Lafortune et al. (2016), Appendix Figure 4.

[25] For the lowest-income districts (Figure 7), we can reject the null hypothesis of zero pre-event effects. This is driven by a temporary drop two years prior to events. A similar, though statistically insignificant, blip is apparent for high-income districts in Figure 8. There is no sign of systematic pre-event trends.

income slope measure in columns 1-2. Our earlier finance analyses also indicated larger effects for quintile gaps than for slopes.

Table 6 presents estimates separately by subject and grade. We cannot reject the null hypothesis of equal effects across each dimension. Appendix Figure A4 presents estimates of the phase-in coefficient for all five quintiles. Only the first quintile effect is large or distinguishable from zero. The ratio of test score effects to spending effects is larger at the bottom of the income distribution, consistent with the idea that funding is more productive in low-income districts, but equal ratios cannot be ruled out.

Robustness

Table 7 presents estimates of our key specifications from our two alternative approaches to event multiplicity. Column 1 repeats the estimates from our preferred approach from Tables 3 and 5. In Column 2, we include all identified events, creating separate panels for each; in Column 3, we focus only on the first court order in each state. Results are similar to those from our main specifications, though the initial court order approach yields less precise, insignificant estimates of finance effects in panel B.

One potential explanation for the achievement impacts that we identify is that they reflect changes in population stratification rather than changes in educational production. SFRs that flatten the gradient of school funding with respect to district income and that reduce the local share of school finance reduce the value of living in a high-income district, and may lead some high-income families to relocate to previously low-income districts. This could lead to rising achievement in these districts with no change in school effectiveness.

We assess this possibility in three ways. First, we have tested whether between-district income gaps narrow in the years following SFRs. We have found no evidence for this—district log incomes in 2011 are highly correlated with those in 1990, and there is no sign that gaps narrow in states that had reforms relative to those that didn't. Second, we have conducted event study analyses, parallel to those for test scores, for district income or the district non-white or free- or reduced-price lunch eligible share (Appendix Table A4). In only one specification—for the between-quintile gap in the free lunch share—do we find evidence that the demographic composition of (initially) low-income districts changes following SFRs. This result is not robust, and is small relative to the test score impacts that we estimate.

Third, we decompose test scores into two components, and estimate separate SFR effects on each. Specifically, we estimate an individual-level regression of test scores on student demographic characteristics, pooling NAEP data across years for each grade-subject pair and including year fixed effects. We then construct separate achievement-log district income gradients from the fitted values (excluding the fixed effects) for this regression, representing student characteristics that would be affected by SFRs only through changes in sorting, and from the residuals. We find no evidence that reforms affect the demographic component of our test score progressivity measures, supporting our interpretation that our results primarily reflect changes in educational production in low-income school districts (see Appendix Table A7).

As a final robustness exercise, we have tested whether the SFR effect on achievement is sensitive to including controls for the presence of a school accountability policy in a state, or whether the SFR effect varies with school accountability. We found evidence for neither.

VI. Finance reforms and statewide achievement gaps

The final topic 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 most inequality is within districts, changes in the distribution of resources across districts may not be well enough targeted to meaningfully close these gaps.

Table 8 presents estimates of effects on mean test scores across different subgroups of interest. The first panel shows a DD estimate of the effect on mean (pooled) test scores. The point estimate (not significant) implies a smaller impact per dollar than do our between-district contrasts, though we cannot rule out comparable effect sizes. In any event, our research design is more credible for outcome *disparities* than for the *level* of outcomes, as the latter would be confounded by unobserved shocks to average outcomes in a state that are correlated with the timing of school finance reforms (Hanushek, Rivkin, and Taylor 1996a,b). For example, if SFRs follow negative shocks to mean student achievement, this effect would be downward-biased. Another interpretation is that the marginal productivity of revenues is in fact higher in low-income districts.

The second panel shows impacts on the standard deviation or interquartile range of achievement within states, while the third and fourth panels present results by race and income, respectively. There is no discernible effect on achievement gaps by race or income or on the overall dispersion of test scores. Point estimates are all roughly a full order of magnitude smaller than the earlier estimates for district-level progressivity of mean scores.

Appendix Tables A5 and A6 resolve the discrepancy. While non-white, low-income, and low-scoring students are more likely than their white, higher-income, and higher-scoring peers to attend school in low-income school districts, the differences are not very large. Roughly one-quarter of non-white and low-scoring students, and one-third of low-income students, live in first-quintile districts, while about 10% of each live in fifth-quintile districts (Appendix Table A5). This leaves little room for SFRs to substantially affect the relative resources to which the typical minority, low income, or low scoring student is exposed.

To assess this more carefully, we assigned each student the mean revenues for his/her district and estimated event study models for the black-white, income, or test score gap in these imputed revenues. Results, in Appendix Table A6, indicate that finance events raise relative per-pupil revenues in the average black student's school district by only $195 (S.E. 164), in the average low-income student's district by $23 (S.E. 195), and in the average low-scoring student's district by 193 (S.E. 101). Even if funding was much *more* productive than the average effect implied by our analysis, the funding changes seen here 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—in the distribution of funding or in other policies that reduce achievement gaps—would be necessary to have dramatic impacts on the average low-income, minority, or low-scoring *student*.

VII. Discussion

After 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 early reforms on school finance have been well studied, there is little evidence about the finance effects of more recent “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. Schools used these additional funds to increase instructional spending, reduce class size, and for capital outlays. Using nationally representative data on student achievement, we find that these reforms were productive: Reforms increased the absolute and relative achievement of students in low-income districts.

Some SFRs were accompanied by other policy changes—e.g., new curricula, accountability provisions, or new pre-kindergarten programs—that may have contributed to the achievement effects, though our impression is that for the typical reform the main change was in funding.[26] We thus interpret our estimates as reflecting the productivity of additional resources, though other interpretations cannot be ruled out.

The different 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. To assess the magnitude of the impacts we estimate, we focus on estimated effects on student achievement ten years after an event. Because effects on school resources are stable in the years following events, these can be interpreted as the impact of a change in resources for every year of a student's career (through 8th grade). Nevertheless, the focus on the r=10 estimate is arbitrary. We would obtain larger estimates of the achievement effect per dollar if we used impacts more than ten years after events, or smaller effects with a shorter window.

Our preferred estimates, based on the gradient of student achievement with respect to district income, indicate that an SFR raises achievement in a district with log average income one point below the state mean, relative to a district at the mean, by 0.1 standard deviations after ten years. Our finance estimates indicate that this district saw an increase in relative state aid of $622 per pupil for each of those ten years, and an increase in total revenues of $424 per pupil.

[26] We used our event-study framework to estimate the association of SFRs with changes in state accountability policy, using various measures of accountability rules, and found no relationship. We also investigated specifications that allowed for interactions between finance reform events and the accountability regime, but found no evidence for this either. We are not aware of a systematic classification of other aspects of state policy that might have been affected by SFRs.

An increase of $424 per pupil in spending each year from kindergarten through grade 8, discounted to the student's kindergarten year using a 3% rate, corresponds to a present discounted cost of $3,400. 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. This implies a benefit-cost ratio of 1.5, even when only earnings impacts are counted as benefits.[27]

This ratio is not wholly robust. Our quintile analysis shows larger revenue effects, implying a benefit-cost ratio below one, while Jackson et al.'s (2016) study 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.

It is important to note that 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 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.

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

[27] The earnings effects of increases in 8th grade test scores are likely larger than those of increases in Kindergarten scores, so using estimates of the latter biases our benefit calculation downward. We do not count the cost of increased spending in grades 9-12, as we have no way to capture its benefits.

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 VI, 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 would require policies aimed at the distribution of achievement *within* school districts, something that was generally not a focus of the reforms that we study.

References

Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics, 79*(4), 551-563.

Berry, C., & Wysong, C. (2012). Making courts matter: Politics and the implementation of state supreme court decisions. *The University of Chicago Law Review 79*(1), 1-29.

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.; Mas, A., & Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics*, *123*(1), 177–218.

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.
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: Routledge.
Downes, T., Stiefel, L. (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., Nguyen-Hoang, P., & J. Yinger (2015). Measurement of cost differentials. In H. F. Ladd, & M.E. Goertz, eds., *Handbook of Research in Education Finance and Policy*, 2nd edition. New York, NY: Routledge.
Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, *32*(4), 692-717.
Flanagan, A. E., and Murray, S. E. (2004). A Decade of Reform: The Impact of School Reform in Kentucky. In J. Yinger, ed., *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity* (pp. 165-213). Cambridge, MA: MIT Press.
Guryan, J. (2001). *Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts.* National Bureau of Economic Research Working Paper No. 8269.
Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, *24*(3), 1141-1177.
Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, *19*(2), 141-164.
Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113, F64-F98.
Hanushek, E. A. (2006). School resources. In E. A. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*, vol. 2. Elsevier.

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

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

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

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

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

Hyman, J. (2013). Does money matter in the long run? Effects of school spending on educational attainment. Unpublished manuscript.

Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics 131*(1), 157-218.

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: Routledge.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The 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. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. *The Economic Journal* 111, 1-28.

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: Brookings Institution Press.

Ladd, H. F., & Goertz, M. E. (Eds.). (2015). *Handbook of Research in Education Finance and Policy,* 2nd Edition. New York, NY: Routledge.

Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2016). *School Finance Reform and the Distribution of Student Achievement.* National Bureau of Economic Research Working Paper No. 22011.

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.

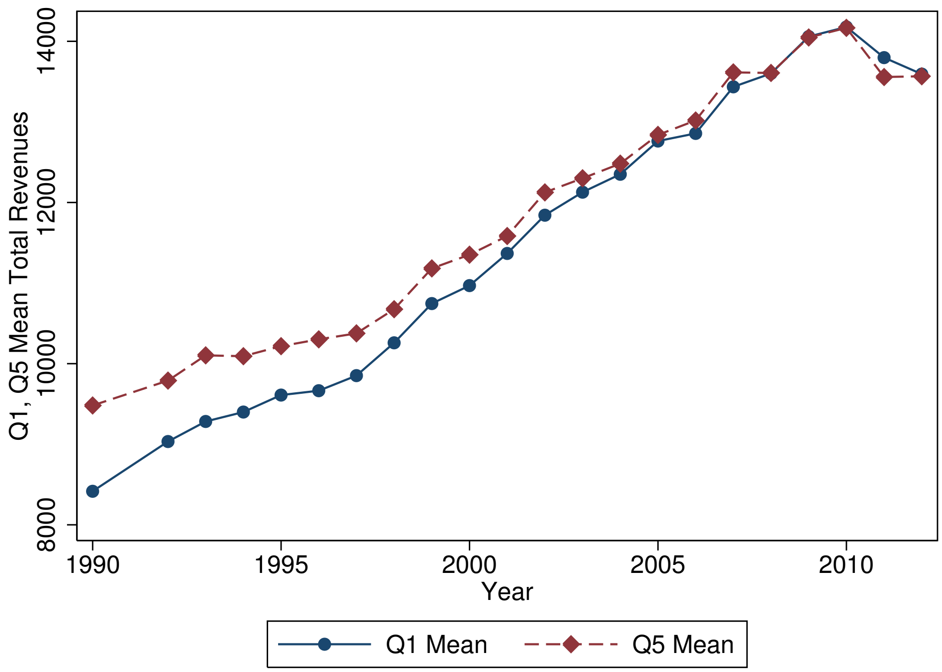
Nielson, C., & Zimmerman, S. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics 120*.

Sands, M. 2015. The distributive politics of education policy: Party control of state government and transfers to localities. Unpublished manuscript, October.

Sims, D. P. (2011a). Lifting all boats? Finance litigation, education resources, and student needs in the post-*Rose* era. *Education Finance and Policy, 6*(4), 455-485.
Sims, D. P. (2011b). 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.

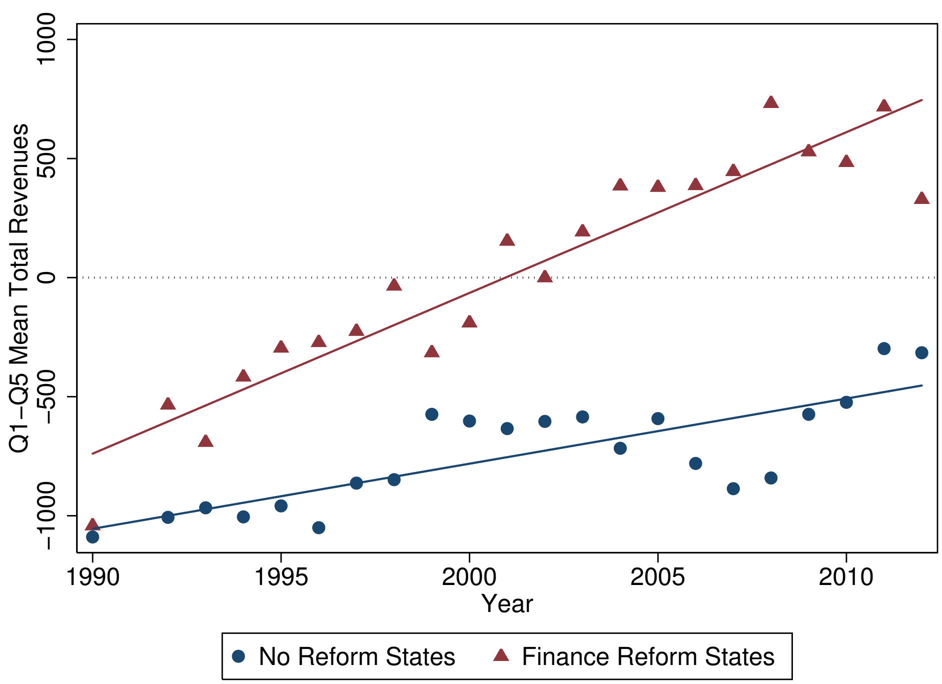
Figures

Figure 1: Mean revenues per pupil for highest and lowest income school districts, 1990-2012



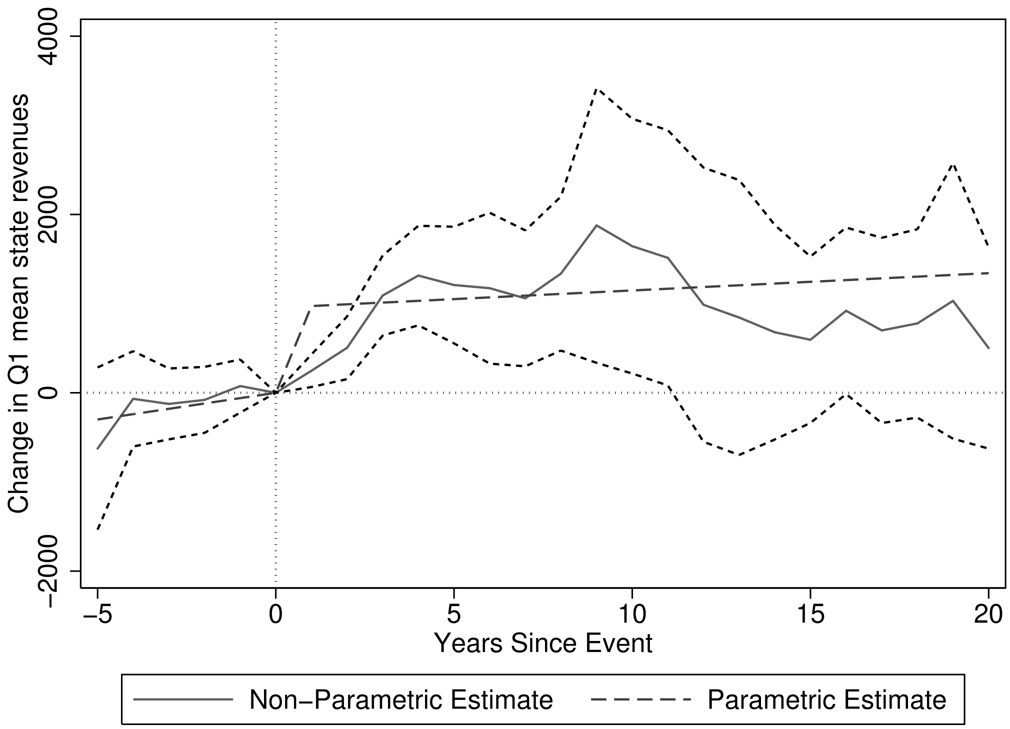
Notes: Highest (lowest) income districts are those in the top (bottom) 20% of their states' district-level distributions of mean household income in 1990, and are labeled as "Q5" and "Q1", respectively. See appendix for details of quintile classifications. Revenues are expressed in real 2013 dollars. Districts are averaged within states, weighing by log district enrollment; states are then averaged without weights. Hawaii and the District of Columbia are excluded.

Figure 2: Gap in revenues per pupil between lowest and highest income districts, by state finance reform status, 1990-2012



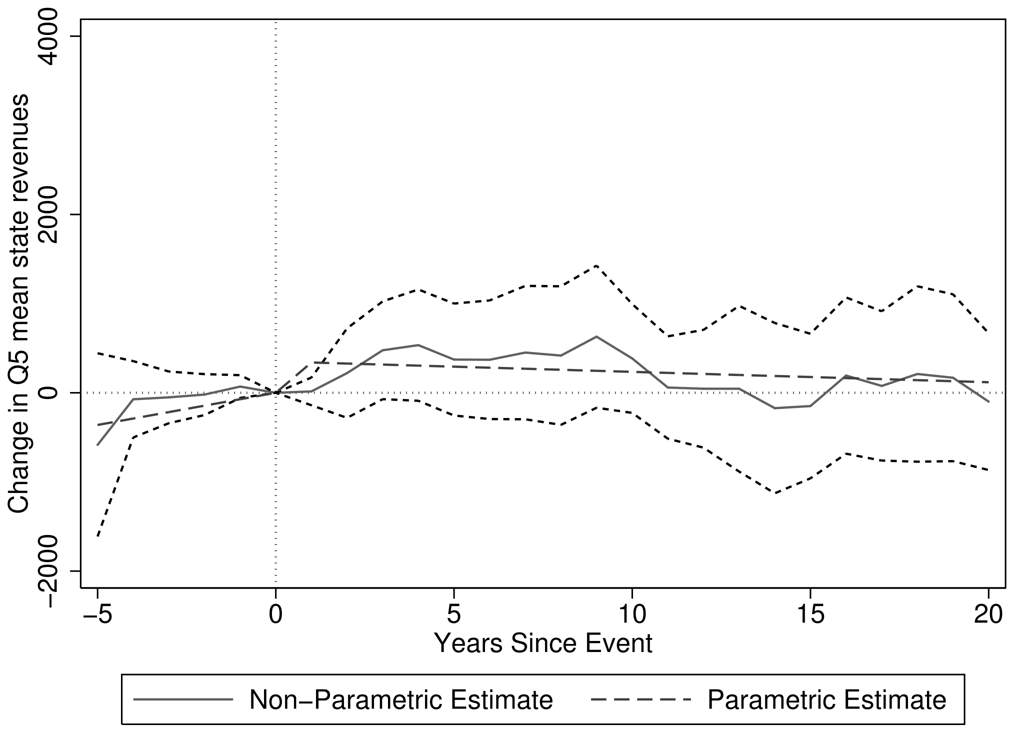
Notes: See notes to Figure 1. Finance reform states are those with school finance reforms between 1990 and 2011, as listed in Appendix Table A1. Lines show unweighted best linear fit to time series.

Figure 3: Event study estimates of effects of school finance reforms on mean state revenues in lowest income districts



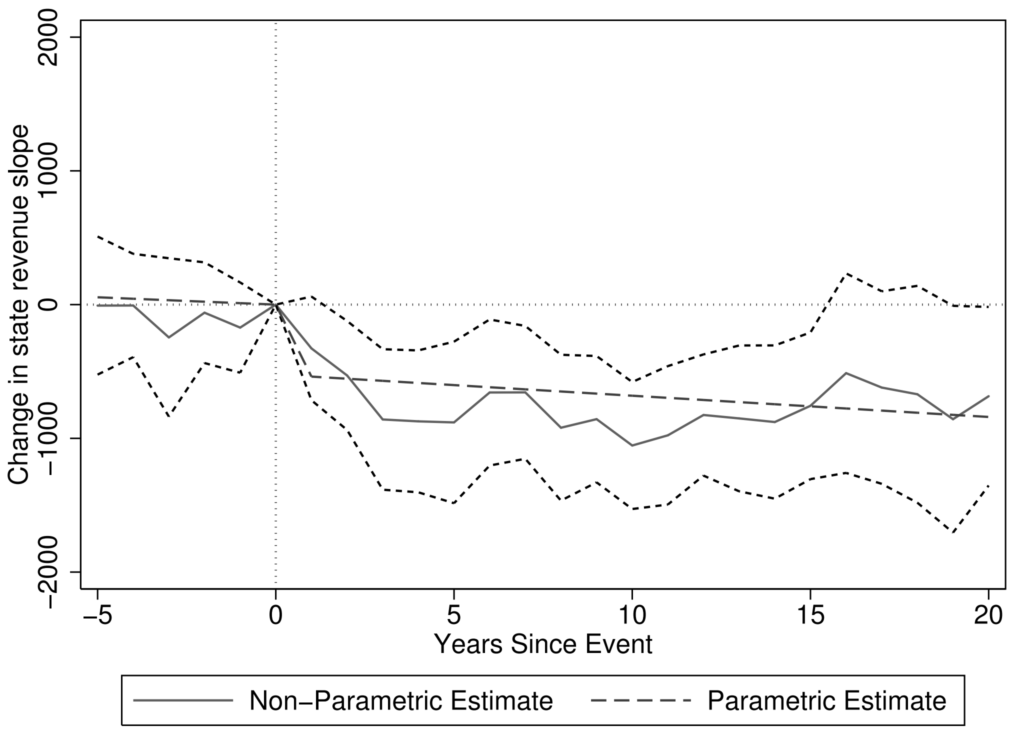
Notes: Figure displays coefficients from event study regressions. Dependent variables are mean state revenues in the lowest income quintile of districts, measured in 2013 dollars per pupil. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 3, column 2, Panels B and C. The p value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model in is 0.53; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.24; for the test that the post-event jump and change in trend is zero it is 0.01. Standard errors are clustered at the state level.

Figure 4: Event study estimates of effects of school finance reforms on mean state revenues in highest income districts



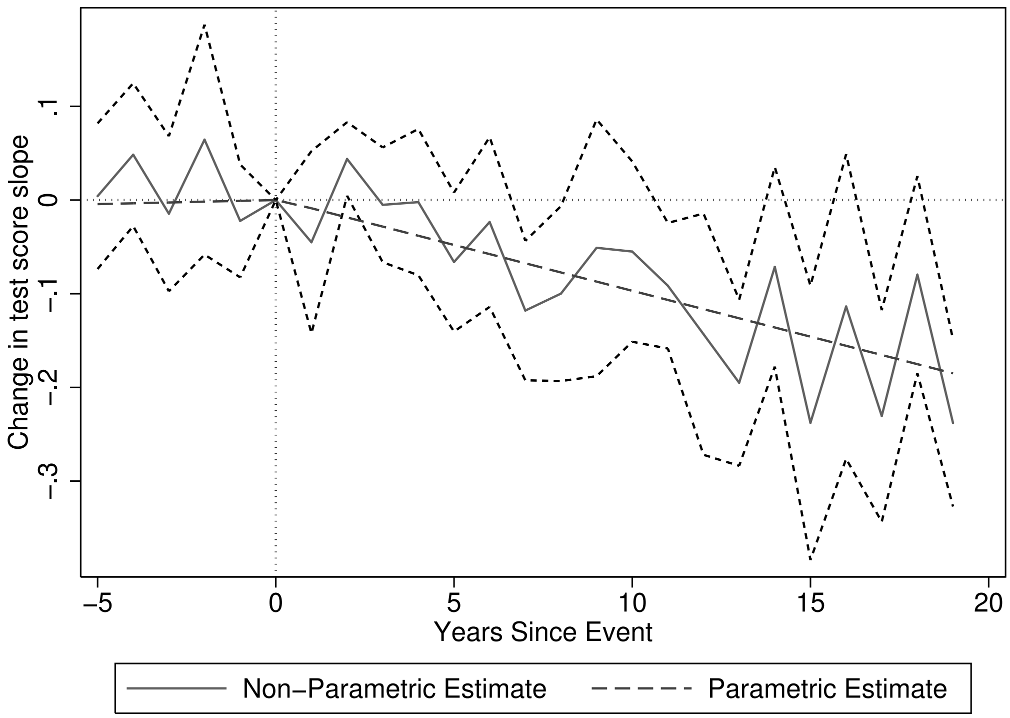
Notes: Figure displays coefficients from event study regressions. Dependent variables are mean state revenues in the highest income quintile of districts, measured in 2013 dollars per pupil. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 3, column 2, Panels B and C. The p value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.41; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.21; for the test that the post-event jump and change in trend is zero it is 0.30. Standard errors are clustered at the state level.

Figure 5: Event study estimates of effects of school finance reforms on progressivity of state revenues



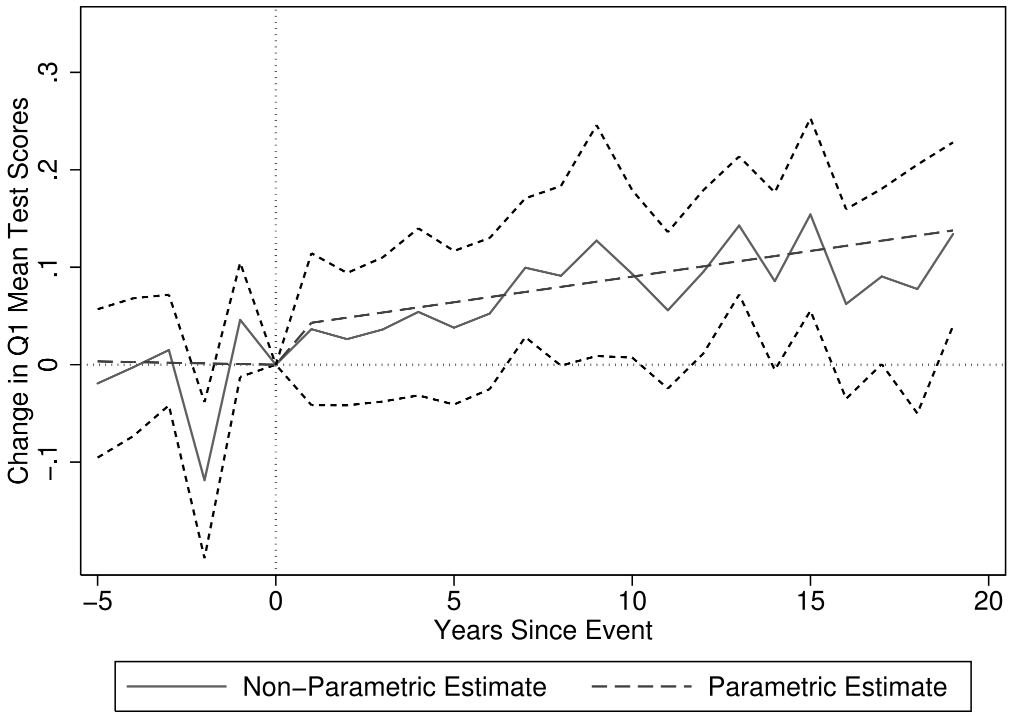
Notes: Figure displays coefficients from event study regressions. Dependent variable is the slope of state per-pupil revenues (in 2013$) with respect to log mean family income, controlling for log enrollment and district type. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 3, Panels A and B, columns 5. The p-value for the omnibus hypothesis tests of zero pre-event effects in the non-parametric model is 0.73; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.67; for the test that the post-event jump and change in trend is zero it is 0.05. Standard errors are clustered at the state level.

Figure 6: Event study estimates of effects of school finance reforms on progressivity of test scores



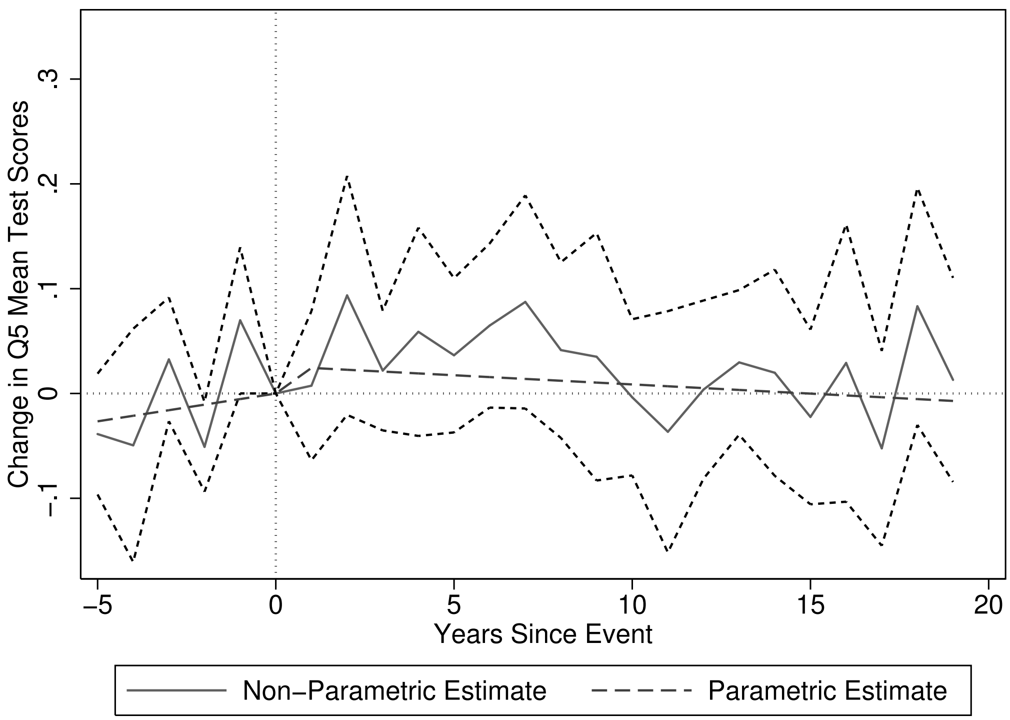
Notes: Figure displays coefficients from event study regressions. Dependent variable is the slope of mean test scores with respect to log mean family income, controlling for log enrollment. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Both event study regressions include state and subject-grade-year fixed effects. Estimates for the parametric models are reported in Table 6, Column 1. The p-value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.43; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.80; for the test that the post-event jump and change in trend is zero it is 0.02. Standard errors are clustered at the state level.

Figure 7: Event study estimates of effects of school finance reforms on mean test scores in lowest income school districts



Notes: Figure displays coefficients from event study regressions. Dependent variables are mean test scores for students at districts in the bottom quintile of the state's distribution of 1990 district mean household incomes. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Both regressions include state and subject-grade-year fixed effects. The p-value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.01; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.86; for the test that the post-event jump and change in trend is zero it is 0.01. Standard errors are clustered at the state level.

Figure 8: Event study estimates of effects of school finance reforms on mean test scores in highest income school districts



Notes: Figure displays coefficients from event study regressions. Dependent variables are mean test scores for students at districts in the top quintile of the state's distribution of 1990 district mean household incomes. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Both regressions include state and subject-grade-year fixed effects. The p-value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.02; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.15; for the test that the post-event jump and change in trend is zero it is 0.25. Standard errors are clustered at the state level.

Tables

Table 1: NAEP Testing Years

Year	Subjects and grades covered				Number of	Number of
	Math G4	Math G8	Reading G4	Reading G8	States	Students
1990		X			38	97,900
1992	X	X	X		42	321,120
1994			X		41	104,890
1996	X	X			45	228,980
1998			X	X	41	206,810
2000	X	X			42	201,110
2002			X	X	51	270,230
2003	X	X	X	X	51	691,360
2005	X	X	X	X	51	674,420
2007	X	X	X	X	51	711,360
2009	X	X	X	X	51	775,060
2011	X	X	X	X	51	749,250

Notes: In final column, students are cumulated across all tested subjects and grades, and rounded to the nearest 10.

Table 2: Summary statistics (district-year panel)

	Overall			Mean by subgroup	
	N	Mean	SD	Q1	Q5
Enrollment	229,386	67,523	181,811	13,537	31,403
Log(mean income, 1990)	223,334	10.53	.2935	10.21	10.9
Total revenue p.p.	229,386	11,087	3,489	10,809	11,871
State	229,386	5,135	2,291	6,371	4,003
Local	229,386	5,094	3,273	3,258	7,349
Federal	229,386	858.2	641.4	1,180	518.4
Expenditures p.p.	229,386	11,264	3,685	10,837	12,116
Instructional	229,386	5,845	1,953	5,659	6,167
Non-instructional	229,386	5,419	2,221	5,178	5,949
NAEP scores	49,867	.2559	.4578	.02925	.5884

Notes: Table reports summary statistics at the district by year level, weighted by district enrollment for the financial variables and by the sum of the student weights for the mean NAEP score.

Table 3: Event study estimates of effects of school finance reforms on revenues per pupil

	Q1	Q5	All districts	Q1-Q5 difference	Slope
A: State revenue (1 parameter)					
Post Event	1,225***	527	912**	711**	-622***
	(343)	(378)	(359)	(316)	(223)
B: State revenue (3 parameter)					
Post Event	954***	351	672**	606***	-522**
	(302)	(325)	(320)	(231)	(209)
Trend	60	72	68	-10	-11
	(50)	(56)	(50)	(25)	(25)
Post Event * Yrs Elapsed	-40	-84	-61	42	-5
	(70)	(61)	(60)	(36)	(21)
C: Total revenue (1 parameter)					
Post Event	1,233***	544**	829***	701**	-424
	(370)	(277)	(302)	(309)	(304)
D: Total revenue (3 parameter)					
Post Event	1,164***	471*	839***	696***	-469**
	(287)	(277)	(269)	(243)	(233)
Trend	16	9	9	9	-25
	(39)	(32)	(32)	(24)	(45)
Post Event * Yrs Elapsed	-11	2	-17	-14	53
	(70)	(41)	(52)	(44)	(61)
Observations	1,078	1,076	1,078	1,076	1,078

Notes: Table reports estimates of the parametric event study models, equations (1) (panels A and C) and (2) (panels B and D). In columns 1-3, dependent variables are mean state (panels A and B) or total (panels C and D) revenues per pupil, weighting districts by their log enrollment; each is computed separately for each state and year. In columns 1 and 2, means are computed over the bottom and top, respectively, quintiles of the states' district 1990 mean household income distributions; in column 3 means are computed over all districts in each state. In column 4, the dependent variable is the gap in state (panels A and B) or total (panels C and D) revenues per pupil between districts in the bottom and top quintiles of the states' district 1990 mean household income distributions. In column 5, the dependent variable is the coefficient from a district-level regression of the state (panels A and B) or total (panels C and D) per-pupil revenue measure on the log of the district's 1990 mean household income, controlling for district log enrollment and district type (elementary / secondary / unified) and weighting by the district's average log enrollment over time. Event study regressions include state and year fixed effects, and are unweighted. Standard errors are clustered at the state level.

Table 4: Event study estimates of effects of school finance reforms on components of district finance

	Mean of depvar	Mean	Q1 Mean	Q1-Q5 Mean	Slope
Revenue Effects:					
Total revenue	11,593	829***	1,233***	701**	-424
		(302)	(370)	(309)	(304)
State revenue	5,449	912**	1,225***	711**	-622***
		(359)	(343)	(316)	(223)
Local revenue	5,238	-146	-126	-126	90
		(307)	(233)	(235)	(339)
Federal revenue	907	63	134	116	34
		(83)	(143)	(116)	(33)
Expenditure Effects:					
Total expenditures	11,595	907***	1,377***	753**	-449
		(290)	(367)	(309)	(309)
Current instructional exp.	6,000	443***	604***	243*	-161
		(134)	(155)	(127)	(208)
Teacher salaries + benefits	5,533	339**	449***	143	-103
		(153)	(169)	(117)	(189)
Mean teacher salary	63,425	-286	-245	272	-259
		(1,024)	(1,107)	(947)	(1,122)
Pupil teacher ratio	15.40	-0.61***	-0.68***	0.03	0.22
		(0.19)	(0.19)	(0.20)	(0.17)
Non-instructional exp	5,595	464**	773***	511**	-232
		(186)	(257)	(235)	(176)
Student support	3,426	221**	299**	100	-81
		(102)	(119)	(83)	(88)
Total capital outlays	1,076	272**	486***	369**	-87
		(114)	(177)	(181)	(78)
Other current exp.	431.0	7.9	9.2	-2.5	-2.9
		(12.4)	(14.5)	(13.3)	(12.1)

Notes: Each entry in columns 2-5 represents the coefficient from a separate event study regression, using the one-parameter specification in equation (1). Dependent variables are constructed from district-level finance summaries indicated by row headings and expressed in per-pupil terms; means across districts are reported in column 1. Specifications correspond to columns 1 and 2 of Table 3, panels A (column 2) and B (column 3), and Table 4, panels A (column 4) and B (column 5). See notes to Tables 3 and 4. Standard errors are clustered at the state level.

Table 5: Event study estimates of effects of school finance reforms on student achievement

	Slopes		Q1	Q5	Q1-Q5	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Event * Yrs Elapsed	-0.011**	-0.010***	0.007**	-0.001	0.008**	0.013**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)
Trend	0.001					-0.006
	(0.003)					(0.005)
Post Event	0.001					0.011
	(0.023)					(0.024)
Observations	1498	1498	1509	1506	1504	1504
p, total event effect=0	0.02	0.01	0.02	0.69	0.04	0.07
State FEs	X	X	X	X	X	X
Sub-gr-yr FEs	X	X	X	X	X	X

Notes: Each column represents a separate event study regression, using specification (2) and, in columns 2-5, constraining $\beta^{jump} = \beta^{trend} = 0$. Dependent variable in columns 1-2 is the slope of test scores with respect to log mean 1990 income in the district, using NAEP weights and controlling for log district enrollment. In columns 3-4, dependent variable is the weighted mean score in districts in the bottom or top quintile, respectively, of the state district-level income distribution. In columns 5-6, dependent variable is the difference between the bottom and top quintiles. All are computed separately for each state-year-subject-grade cell with available data. All event study specifications include state and subject-grade-year fixed effects, and are weighted by the inverse squared standard error of the dependent variable. p-values for total event effect in columns 1 and 6 test the hypothesis that the β^{jump} and $\beta^{phasein}$ coefficients are both zero; in columns 2-5, the p-value is for the hypothesis that $\beta^{phasein} = 0$, with β^{jump} constrained to zero. Standard errors are clustered at the state level.

Table 6: Event study estimates of effects of school finance reforms on student achievement by subject and grade

	Test Score Slope	Q1-Q5 Mean
Pooled	-0.010***	0.008**
	(0.003)	(0.004)
By Subject:		
Math	-0.012***	0.007*
	(0.003)	(0.004)
Reading	-0.006	0.009**
	(0.005)	(0.004)
Difference	-0.006	-0.002
p-value	0.09	0.46
By Grade:		
G4	-0.010**	0.009*
	(0.005)	(0.005)
G8	-0.010**	0.007**
	(0.004)	(0.004)
Difference	0.000	0.001
p-value	0.93	0.72

Notes: First row repeats specifications from Table 6, columns 2 and 5. See notes to that table for details. Subsequent models restrict the event study sample to slope and quintile gaps computed in specific subjects or grades. Difference entries report the difference in coefficients between math and reading or grade 4 and grade 8 specifications, with p-values for the hypothesis that the event study coefficient is equal in the two subsamples. Standard errors are clustered at the state level.

Table 7: Sensitivity of event study estimates to the treatment of states with multiple events

	Selected Events	All events (stacked)	Initial court events
Panel A: Gradients			
State revenue p.p.	-622***	-479***	-432*
	(223)	(160)	(222)
Total revenue p.p.	-424	-197	-399
	(304)	(269)	(292)
NAEP scores	-0.010***	-0.009***	-0.009***
	(0.003)	(0.003)	(0.003)
Panel B: Q1-Q5 differences			
State revenue p.p.	711**	463**	516
	(316)	(191)	(354)
Total revenue p.p.	701**	448**	584
	(309)	(195)	(398)
NAEP scores	0.008**	0.011***	0.008**
	(0.004)	(0.004)	(0.004)

Notes: Column 1 repeats estimates of the one-parameter parametric event study models from Table 4, columns 1 and 3, and Table 6, columns 2 and 5. See notes to those tables for details. In column 2, each potential event in each state is included, with a separate copy of the state's finance or test score panel for each event. Event study specification is modified to include state-by-event (-by-grade-by-subject) fixed effects. Column 3 returns to the single-event specification, but uses the first post-1990 court order in each state as its event; states without judicial events are treated as not having finance reforms. Standard errors are clustered at the state level.

Table 8: Event study estimates for mean NAEP scores by subgroup

	Post Event * Yrs Elapsed	
Overall mean	0.004	(0.003)
Spread of distribution:		
Std Dev.	-0.000	(0.001)
25th percentile	0.004	(0.003)
75th percentile	0.003	(0.002)
P75 - P25	-0.001	(0.002)
By race:		
Black	0.001	(0.003)
White	0.004*	(0.003)
White - black	0.002	(0.002)
By free lunch status:		
Free lunch	0.001	(0.003)
No free lunch	0.004	(0.003)
No free lunch - free lunch gap	-0.000	(0.002)

Notes: Table reports event study specifications, using equation (3) with β^{jump} and $\beta^{phasein}$ constrained to zero. Dependent variables are the indicated summaries of the state-level student achievement distribution: The mean score; the standard deviation of scores; the 25th and 75th percentile scores; the interquartile range; mean scores for black and white students, respectively; the white-black mean score gap; mean scores for free/reduced-price lunch and non-free/reduced-price lunch students; and the gap between these. Standard errors are clustered at the state level.

School Finance Reform and the Distribution of Student Achievement: Appendix

February 2017

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

Appendix A. Event database

Appendix Table A1 lists all of the events in our database, covering the period from 1990 forward. As noted in the text, we include both court orders and major legislative reforms. Columns of the table indicate how our list compares with those of Jackson, Johnson, and Persico (2016) and Corcoran and Evans (2015).

Appendix D, below, discusses in depth each case where our list differs from that of Jackson, Johnson, and Persico (2016).

Appendix Figure A1 shows which states have and have not had events since 1990.

Appendix B. Details of samples and empirical specifications

The primary sources for our outcome measures are the Common Core of Data Local Education Agency (LEA) finance survey (also known as the F-33), for finance outcomes, and the NAEP, for test scores. Each is matched to district mean household income from the 1990 School District Data Book (SDDB), a tabulation of Decennial Census data at the district level. Mean incomes pertain to all households in the district, with and without children and without regard to public school attendance.

Our analysis relies on collapsing the district- and student-level measures to summaries at the state-by-year level. We use three types of summaries: Means for districts in each quintile of family income, the difference in between the first and fifth quintile means, and slopes with respect to log district income. Our methods differ slightly among these; we describe them here.

Sample definitions

All of our samples exclude Hawaii and the District of Columbia, each of which has only one school district.

Our finance analyses exclude district-year observations with enrollment of fewer than 100 students. This removes 8% of district-year observations, with only 0.1% of total enrollment.

We make two additional exclusions aimed at reducing volatility in the per-pupil funding measures. Both total funding and enrollment can vary dramatically from year to year in a district, particularly in small districts, creating enormous swings in per-pupil revenues. We view this variability as likely to reflect measurement error; it is particularly problematic when it derives from large proportional swings in enrollment with more stable funding.

We begin by computing each district's average enrollment over our sample, as well as its average growth rate over our sample period. We exclude from our sample any district-year observation with enrollment (a) more than double the district's average enrollment; (b) more than 15% above or below the prior year or the subsequent year's enrollment; or (c) more than 10% above or below the district's constant-growth-rate trend. In addition, for any district for which more than one-third of annual observations are excluded under these criteria, we exclude all remaining observations as well. Exclusion (a) in particular likely leads us to exclude districts serving newly developed areas, but we are not confident that 1990 incomes are reliable measures of population resources in these districts in any case. Together, these exclusions capture 18% of district-year observations, with 8% of enrollment.

To address volatility in the numerator of our revenue and expenditure measures, we exclude as well district-year observations with per-pupil revenues (respectively, expenditures) greater than 500% or less than 20% of the unweighted state-by-year mean. Only 0.02% of district-year observations are excluded by these rules.

When analyzing mean teacher salaries and pupil-teacher ratios, each of which is somewhat noisily measured, we exclude the top and bottom 2% of districts (unweighted) in each state-year cell.

Finally, our NAEP analyses exclude students in charter schools.

For many purposes, it is useful to have a weight for each district that does not vary over time. We use the geometric mean of the district's enrollment in all available years for this.

Definition of income quintiles

The basis for our income quintile calculations is the finance sample, as defined above. Districts' quintile assignments are treated as fixed over time.

To construct our income quintile cutoffs, we sort districts in a state by their 1990 mean family income and compute the 20^{th}, 40^{th}, 60^{th}, and 80^{th} percentile of the state distribution of 1990 mean family income. These percentiles are based on the districts in our finance sample in 1994 (the first year in which complete CCD data are available), weighted by our stable enrollment count. Districts spanning the quintile cutoffs are assigned with partial weights to each of the relevant quintiles.

Quintile means

Quintile means of our finance measures are computed as weighted averages over the districts in each quintile, weighting each district by its average log enrollment and including only districts that meet the criteria outlined above. The specific districts included can vary slightly over time due to differences in the availability of the dependent variable.

Our NAEP quintile means are similar but weight districts by the sum of the NAEP student weights in the district. Only districts meeting the finance sample restrictions for the relevant year are included.

Income slopes

To construct state-by-year slopes of revenues with respect to district income, we estimate a separate regression for each state and each year. Explanatory variables are the log mean income of the district, based on the 1990 data, the district's log enrollment in that year, and indicators for whether the district is an elementary or a secondary district (unified districts are the excluded category). These regressions are weighted by our stable log enrollment measure and samples are defined as above. The coefficient on the log mean district income is extracted, along with its standard error.

NAEP score-income slopes are computed similarly, using NAEP data aggregated to the district-year-subject-grade level. Separate regressions are estimated for each state-year-subject-grade combination. The district-level regression is weighted by the sum of NAEP student weights in the district, and does not include district type controls. Standard errors account for the NAEP plausible values methodology, as discussed below.

Event study regressions

Once quintile means, between-quintile gaps, and income slopes are constructed at the state-year level, we estimate event-study regressions as described in the main text.

For finance outcomes, where we have a census of school districts, the event study regressions using quintile means as dependent variables are unweighted. For test score outcomes, our quintile mean event study regressions are weighted by the sum of the NAEP student weights within the grade-subject-quintile-year cell. When we examine quintile gaps in scores, we weight by the harmonic mean of the two quintile weights.

All event studies with income slopes as the dependent variable are weighted by the inverse sampling variance of the state-year slope.

NAEP plausible values

NAEP does not report a single test score for each student, but rather reports five plausible values, random draws from the student's posterior distribution. We average these five plausible values before computing quintile means or income slopes. Our estimated standard errors for the income slopes account for the contribution of the sampling from the posterior distribution, following NAEP guidance.

Jacob and Rothstein (2016) point out that the use of NAEP plausible values as dependent variables may create biases, as the measurement error in PVs is not classical. The bias depends on the degree to which the "conditioning variables" in the NAEP model can predict the relevant explanatory variables in the research regression, but likely takes the form of attenuated treatment effects.

Jacob and Rothstein (2016) discuss methods for obtaining unbiased estimates from a single NAEP. We are not aware of methods for avoiding bias in analyses that pool across many NAEP samples. We have verified that cross-sectional regressions of NAEP PVs on measures of district finance, both across and within states, are minimally biased relative to unbiased marginal maximum likelihood estimates based on underlying item responses, and therefore conclude that the use of PVs is unlikely to meaningfully bias our results.

Appendix C. Robustness and additional analyses

Appendix figures and tables present numerous additional results.

Figure A2 shows the gap in average NAEP scores between the lowest- and highest-income districts, averaged separately across states that did and did not have reform events since 1990. This is analogous to Figure 2 of the main paper. Reform states had larger test score gaps in 1990, but while the gaps in non-reform states were stable from 1990-2011, the gap shrunk steadily in reform states.

Figure A3 displays coefficients from parametric (equation 2) and non-parametric (equation 3) event study regressions on mean total revenues per pupil (in 2013 dollars), including revenues from state, local, and federal sources. Dependent variables are mean total revenues per pupil (panel A), mean total revenues per pupil in the lowest income quintile of districts (panel B), mean total revenues per pupil in the highest income quintile of districts (panel C), and the difference in mean total revenues per pupil between districts in the bottom and top income quintile in the state (panel D). Parametric estimates correspond to those reported in panel D of Table 3. Mean total revenues per pupil increased significantly following reform events (panel A), with little sign of pre-trends. Revenue increases were concentrated in districts in the lowest income quintiles (panel B), whereas districts in the top income quintiles (panel C) saw only small increases. Thus, the difference in mean total revenues between districts in the bottom and top income quintiles in a state increased significantly following reform events (panel D). All of these effects persisted essentially unchanged for at least 10 years following the events.

Figure A4 plots our event study estimates of impacts on total revenues and mean test scores in each of the five quintiles, along with 95% confidence intervals. Estimates are consistent with effects on test scores that are proportional to effects on revenues across district groups, though the confidence intervals are wide.

Table A2 presents an event study analysis for state-level budgets, on both per-capita and per-pupil bases. Our one-parameter specification (equation 1) is used. Events are associated with sharp increases in both total expenditures and total revenues, though the latter are imprecisely estimated and not significant. There is some indication that non-education expenditures rose following reforms, but estimates are imprecise and this result is not robust. In any event, it does not appear to be concentrated in health and welfare budget items, which seem most likely to confound our estimates. The impacts on total per-pupil education expenditures are larger than in our district-level analyses, though confidence intervals are quite wide. Insofar as this is real effect, it may indicate that some of the new state expenditures involve re-labeling existing funds rather than allocating new funds to education.

Table A3 presents Card and Payne's (2002) analysis of the effect of SFRs on the slope of district revenues with respect to district income. Card and Payne use the income level, in dollars, rather than log income as the explanatory variable in this slope calculation. They compute a single long-difference of this slope for each state, from 1977 to 1992, and regress it on indicators for plaintiff and defense victories in court cases. We construct a similar slope with respect to income levels, and a similar 1990-2012 long difference. Columns 2 and 5 report regressions of this on an indicator for a post-1990 event. Columns 3 and 6 report estimates of our one-parameter event study model. Our estimates of the impacts of court rulings are directionally similar but somewhat larger than in Card and Payne (2002). We cannot rule out equal effects.

Table A4 reports event study estimates where the dependent variable is the slope of various demographic characteristics of a district's students with respect to district log income (panel A) or the between-quintile gap in mean demographic characteristics (panel B). Log mean income, used in columns 1-2, is measured in 1990, 2000, and 2011. (The slope of log mean income in 1990 with respect to a districts 1990 log mean income is by construction 1 in every state, but this slope can vary in subsequent years.) Minority and free lunch shares, used in columns 3-6, are measured in every year, though free lunch data are missing for some states and years. In five of six cases, we find no indication of an effect of SFR events on the level or trend of demographic composition. The one exception is the low- vs. high-income district gap in student free lunch share, where we find that the relative prevalence of poor students in low-income districts may have declined slightly following SFRs. The estimates are imprecise, however, and the point estimates are quite small: They indicate that the relative free lunch share in low-income districts might have fallen by 3-5 percentage points over the decade following an SFR, far too little to account for the 0.1 standard deviation effect that we find.

Table A5 reports the share of students of various characteristics who are in districts in each quintile. Rows sum to 1. Minority and low-income students are disproportionately represented in bottom-quintile districts, but there are substantial shares in even the highest-income districts.

Table A6 reports event study regressions where the dependent variable is the difference in the revenues of the district attended by the average black (or free lunch, or low scoring) student and the revenues of the district attended by the average white (or non-free-lunch, or high scoring) student. Estimates indicate that the average black student in a state was exposed to less than $200 in additional per pupil revenue relative to the average white student in the same state following an SFR, and that free lunch students got no additional revenues on average.

Table A7 further explores the potential impact of demographic changes on our test score results. We decompose test scores into two components, one capturing demographics and one not, and estimate separate SFR effects on each. Specifically, we estimate an individual-level regression of test scores on student demographic characteristics, pooling NAEP data across years for each grade-subject pair and including year fixed effects. We then construct separate achievement-log district income gradients from the fitted values (excluding the fixed effects) for this regression, representing student characteristics that would be affected by SFRs only through changes in sorting, and from the residuals. Table A7 presents results of our event study analyses of these gradients. We present two decompositions: The first panel uses only race and gender, which are consistently available in each NAEP wave, along with school means of these. The next uses additional covariates, parental education and free lunch status, that are less consistently available, including indicators for years in which each is unavailable. The first set of variables explains 22% of the variance in student test scores (net of the subject-grade-year effects), while the second set explains 28%.

We find no evidence that reforms affect the demographic component of our test score progressivity measures. Point estimates are less than half the size of our overall test score impacts, and are never significantly different from zero. By contrast, estimated effects on the residual component of test scores are all significant, and about two-thirds the size of the overall impacts. Thus, while we cannot rule out small effects of SFRs on student sorting, the robustness of effects on the residual component supports our interpretation that our results primarily reflect changes in educational production in low-income school districts.

Appendix D. Monte Carlo analyses of event studies with multiple events

Our analysis differs from many event study analyses in that states in our sample often experience multiple potential events. Our analytical strategy is predicated on the view that states typically implemented only one actual finance reform during our sample, and that other apparent events represent political and legal maneuvering with little consequence for school spending or student achievement. Accordingly, we develop a procedure for selecting a single potential event in each state and estimate our models as if that were the only event.

In Table 7, however, we present two additional approaches that would be more appropriate if states actually had multiple events. The first uses all identified potential events, creating separate copies of the state's time series for each and analyzing them independently (though clustering standard errors by state), while the second uses only the first court order in a state and ignores any subsequent events (as well as any prior legislative events).

In this Appendix, we present Monte Carlo analyses of the performance of these two estimators under two data generating processes (DGPs). In our first DGP, all events have equal effects; in the other DGP, one (randomly chosen) event in each state has an effect, while others do not. Other aspects of the DGP are quite simple: The distribution of potential events is as in the true data, and errors are independent over time within states. (We do not use these Monte Carlos to study our preferred event selection approach, as the performance of that approach is highly dependent on the specific time series properties of the errors.)

Specifically, suppose that state s has J_s potential events that take place in $t = \{t_{s1}, t_{s2}, \dots, t_{sJ_s}\}$. Our first DGP is

$$y_{st} = \frac{1}{J_s}\sum_{j=1}^{J_s} 1\left[t > t_{sj}\right] + \epsilon_{st}.$$

That is, the outcome rises, permanently, by $\frac{1}{J_s}$ following each event. In the second DGP, we choose for each state a random integer j_s^* between 1 and J_s; the DGP is then

$$y_{st} = 1\left[t > t_{sj_s^*}\right] + \epsilon_{st}.$$

Here, outcomes rise by a full unit following the randomly chosen true event, with no change following other apparent events. Thus, in each DGP the total effect across all events in the state is 1.

We model ϵ_{st} as i.i.d. standard normal random variables. We use the actual sequence of measured events in each state for $\{t_{s1}, t_{s2}, \ldots, t_{sJ_s}\}$. 53% of states have events; among those with at least one, the average state has 2.5 events, and 69% have more than one.

We simulate each DGP 5000 times. We apply our two estimators, one selecting just the first event and one using all events with stacked panels for each, to each simulated sample. In each case, we use the same specifications as are applied to actual data in Table 7, with state(-by-event-copy) and year fixed effects and a single jump following the event, but no state-specific trend terms.

Table A8 shows the results. The two rows correspond to the two DGPs, while the two columns show the two different estimators. The first thing to notice is that the rows are nearly identical: Each estimator yields very similar results whether the DGP includes multiple true events or just one. Second, the first column, in which the estimator is based solely on the first apparent event, yields estimates of the event effect of about 0.79. This is somewhat attenuated from the total event effect, but the bias is not severe. Intuitively, the post event period in this estimator includes some years in which event effects have not yet kicked in (or have only partially kicked in), reducing the contrast with the pre-event period.

In the second column, the estimator that stacks all potential events in a state yields somewhat lower estimates, around 0.58. Here, the intuition is that for any given potential event, some of the pre-event observations reflect the impact of earlier events, and some of the post-event observations do not reflect the impacts of events yet to come. Again, this attenuates the effect.

In general, the pattern of results in Appendix Table A8 mirrors that in Table 7. Estimates from the initial event and stacked specifications are both smaller than those from our preferred specification, which focuses on a single event selected as the true one, and (in most results) the attenuation is greater for the stacked specification than for the initial event specification. Overall, we conclude that the three approaches are all likely to work reasonably well in our setting, but that the more "hands off" estimators are likely to understate the true effects of events.

Appendix E. Reconciliation of school finance reform tabulations

The literature on school finance reforms has been plagued by a lack of authoritative tabulations of court-ordered reforms, with substantial discrepancies between the tabulations used by different authors. Our tabulation, too, differs from all previous listings.

In an effort to provide clarity to the literature going forward, in this appendix we discuss every case between 1990 and 2011 where our tabulation of court-ordered school finance reforms differs from that of Jackson, Johnson, and Persico (2016; hereafter *JJP*). Many of these discrepancies reflect judgment calls. We have estimated our main results with a number of variants of the event sample, and in general have found little sensitivity of the results; we nevertheless present the basis for our preferred tabulation for completeness.

The states and years for which the two tabulations disagree are:

- Alabama, 1993
- Arizona, 2007
- Connecticut, 1995 & 2010
- Idaho, 1993 & 1998
- Maryland, 1996 & 2005
- Michigan, 1997
- Montana, 1993 & 2008
- New Hampshire, 2006
- New Jersey, 1991, 1998 & 2000
- New Mexico, 1998 & 1999
- Oregon, 2009
- South Carolina, 2005
- Texas, 2004
- Washington, 1991, 2007 & 2010

This includes only cases in scope for both lists but coded differently. This in particular means that we do not discuss our tabulation of legislative school finance reforms, as these are out of JJP's scope. For each state, we discuss only the events where the two tabulations disagree; see Appendix Table A1 for a full listing of events in each state.

Alabama[1]

1993: JJP court order; Lafortune-Rothstein-Schanzenbach (LRS) no event

In 1993's *Alabama Coalition for Equity (ACE) v. Hunt*, the public school funding system was found inequitable, on both adequacy and equity grounds, by a lower court, and a remedy order was issued. The remedy negotiated among the parties and ordered by the court, in *ACE v. Folsom*, included equitable and adequate funding,

[1] Case histories from http://schoolfunding.info/2012/01/school-funding-cases-in-alabama/; http://www.encyclopediaofalabama.org/article/h-2045.

and in addition also covered performance-based education, professional development, early childhood programs, and inclusive special education— all to be fully funded within six years. No educational reform package made it through the legislature before the 1994 election season, and education reform became an important issue in the gubernatorial campaign, with incumbent governor Folsom promising reform and compliance with the remedy order and his opponent (and eventual winner) Forrest "Fob" James vowing to fight what he described as a usurpation of executive and legislative powers. Upon appeal, the Alabama Supreme Court decided in 1997 that while schools were inadequately funded, it would decline to issue a remedy order, leaving the funding system unchanged. Because the lower court was overturned and no school finance legislation was passed, we do not code this event as a school finance reform.

Arizona[2]

2007: JJP court order; LRS no event

Flores v. Arizona was first decided in 1992 under Federal law, in the United States District Court for the District of Arizona. The plaintiffs claimed that the state failed to adequately fund programs for English language learners (ELLs). In 2000, the district court found that the state's method and level of funding ELL programs was "arbitrary and capricious" and ordered that the level of state funding for ELL programs bear a rational relationship to the cost of those programs. The parties reached an agreement in 2002, and the court ordered a costing-out study. The state's repeated failure to comply led to a December 2005 order and daily fines that mounted to $21 million before the state enacted additional funding in early March 2006. In August 2006, the Ninth Circuit Court of Appeals, in *Flores v. Rzeslawski*, vacated the 2005 district court judgment and remanded the case so the district court could hold new hearings to determine whether circumstances had changed and required modification of the 2000 court order.

In March 2007 Judge Raner Collins of the U.S. District Court for the District of Arizona ruled that Arizona was in violation of the Equal Educational Opportunities Act (EEOA) by under-funding programs directed towards English learners, invalidating HB 2064, the funding formula passed by the Arizona legislature in response to the court's earlier decision. Judge Collins ordered the state to comply with the order by the end of the 2007 legislative session, but when the legislature failed to do so the judge issued a contempt order. The Ninth Circuit Court of Appeals upheld the ruling in 2008.

In September 2008, the defendants petitioned the U.S. Supreme Court to review the 9th U.S. Circuit Court of Appeals' holding. In 2009's *Horne v. Flores*, 129 S. Ct. 2579 decision, the Supreme Court reversed and directed the District Court to examine

[2] Case history from http://www.clearinghouse.net/detail.php?id=11194; http://www.schoolfunding.info/states/az/lit_az.php3.

several specific factors, including whether non-compliance was statewide.

On March 28, 2013, the plaintiff's statewide claims were dismissed, requiring district-by-district analysis instead. The dismissal was upheld by the Court of Appeals in June 2015.

Because the 2007 ruling was a Federal district court order that the state never complied with and was subsequently overturned by the Supreme Court, we do not code this event as a school finance reform.

Connecticut[3]
1995: JJP court order; LRS no event
2010: JJP court order; LRS no event

In 1996, the State Supreme Court ruled in *Sheff v. O'Neill* (coded by JJP as 1995) that the separation of suburban and Hartford students violated the segregation clause in the Connecticut Constitution, and ordered the State Legislature to take necessary measures to integrate schools and to provide equal educational opportunity to all children. This resulted in a plan by the 1997 State Legislature geared at promoting voluntary school desegregation and magnet schools. Though plaintiffs made adequacy-based arguments, the ruling and subsequent legislation focused on desegregation and not school funding.

In 2010, the Supreme Court ruled in *Coalition for Justice in Education Funding, Inc. v. Rell* that the state's constitution guaranteed all students an adequate education. It did not order changes in school finance, however, but rather sent the case back to a trial court to determine whether the appropriate standard had been met. As of 2013, the case was still pending. We therefore do not code it as a school finance reform order.

Idaho[4]
1993: JJP no event; LRS court order
1998: JJP court order; LRS no event

In the 1993 ruling on *Idaho Schools for Equal Educational Opportunity v. Evans* (ISEEO) (850 P.2d 724), the Idaho Supreme Court found that the state constitution required adequate (but not equitable) school spending. In 1994, the

[3] Case history drawn from https://www.jud.ct.gov/external/news/sheff.htm; http://connecticuthistory.org/sheff-v-oneill-settlements-target-educational-segregation-in-hartford/#sthash.6QnsSrbm.dpuf; https://scholar.google.com/scholar_case?case=10572221569547466633&q=Sheff+v+o+neill+199&hl=en&as_sdt=400006; http://schoolfunding.info/2012/01/school-funding-cases-in-connecticut-2/.
[4] Case history from http://www.educationjustice.org/states/idaho.html; https://nces.ed.gov/edfin/pdf/lawsuits/ISEEO_v_idaho.pdf.

legislature passed Senate Bill 1560 which revised the state funding formula in regard to teacher salaries, allocating more than $90 million to public schools fund this change. We code the 1993 decision as a court-ordered school finance reform.

After the legislative changes, the trial court declared the lawsuit moot, but this decision was overturned by the state Supreme Court, which concluded that whether a "thorough education" was being provided to students was still in question. In 1997, the trial court again dismissed the plaintiffs' claim. The state Supreme Court reversed in part in 1998, in *ISEEO v. State* (976 P.2d 913), remanding the facilities and capital funding portion of the case. The court held that "the Legislature has the duty to provide a means for school districts to fund facilities that offer a safe environment conducive to learning." In 2000 and 2001, the legislature passed minor facilities measures that help property-poor districts, but plaintiffs argued these measures were insufficient. We do not code this as an independent school finance reform, due in part to its limited scope.

Maryland[5]
1996: JJP no event; LRS court order
2005: JJP court order; LRS no event

The ACLU and Baltimore City alleged that Baltimore's students were not receiving an adequate education. In a 1996 summary judgment decision in the consolidated *Bradford v. Maryland State Board of Education* case, the trial court agreed, though the cause of the inadequacies was in dispute. The parties entered into a settlement that provided an increase in state funding for the Baltimore City Public Schools for the next five years. During this period, the "Thornton" Commission on Education Finance, Equity, and Excellence was established to address statewide adequacy in funding. We code the 1996 court order as a school finance reform, in part because Baltimore is such a large district.

In 2004, the Baltimore schools had an accumulated budget deficit of $58 million. In response to a new state law requirement, it cut its budget drastically, and planned a two-year paydown of the deficit. The ACLU returned to court in Bradford, trying to restore funding to Baltimore schools and stop cuts to academic programs impacting students. The Circuit Court ruled that the budget cuts had resulted in reduced educational opportunity to students and that $30 million to $45 million in funds should be restored, preferably with additional revenue from the city and state. The State appealed, arguing that education funding levels are outside court jurisdiction. In 2005, Maryland's highest court ruled against the State's attempt to strike the lower court order, but did not overturn the state law for the deficit paydown. As a result, the additional funding awarded under the Thornton commission would proceed, but since no additional funds were ordered at this time we do not code this as a separate court ordered school finance reform.

[5] Case history drawn from http://www.schoolfunding.info/states/md/lit_md.php3; http://www.aclu-md.org/uploaded_files/0000/0173/bradford_summary.pdf.

Michigan[6]
1997: JJP court order; LRS no event

Durant v. State of Michigan ("Durant I"), was filed in 1980 and decided in 1997. The major issue was state funding for special education mandates. In its ruling on *Durant I*, the Michigan Supreme Court unanimously held that state government had not properly financed three state-imposed mandates: special education, special education transportation and a school lunch program. The court split awarded monetary damages to local school districts to repay past costs of mandates. Due to the limited nature of the lawsuit, we do not code this as a school finance reform.

Montana[7]
1993: JJP court order; LRS legislative event but no court order
2008: JJP court order; LRS no event

Montana's Supreme Court ruled in *Helena Elementary School District No. 1 v. State* in 1989 that the state's school finance system was unconstitutional. This is outside the scope of our sample. The legislature responded in 1989, then overhauled the formula again in 1993 via House Bill 667. Earlier that year, the *Montana Rural Education Association v. State* case was tried but not decided. Following the legislative action, but still in 1993, the state's First Judicial District Court for Lewis and Clark County ruled that the case was moot due to the new law. It permitted the plaintiffs to argue that the new law remained unconstitutional, but to our knowledge the case ended then. JJP code this as a court order; we code House Bill 667 as a legislative action, but do not code the case as a court order.

In 2005, in *Columbia Falls Elem. Sch. Dist. 6 v. State*, the trial court found that the state was not providing a "quality" education as mandated by the constitution, and in particular it had violated the provision of the state constitution requiring the state to commit to preserve the cultural heritage of American Indians. JJP and LRS each code this as a court-ordered school finance reform. A subsequent 2007 legislative reform (which we code as a legislative event) made substantial changes to the school finance system in light of this ruling.

---

[6] Case history from https://www.mackinac.org/8568.
[7] Case history from http://www.mqec.org/school-funding-history/; http://schoolfunding.info/2011/12/school-funding-cases-in-montana/; https://static1.squarespace.com/static/53ab63e1e4b0cb2b67560152/t/55ef5b40e4b064e46223df9f/1441749824419/CF-Decision-II.pdf; http://static1.squarespace.com/static/53ab63e1e4b0cb2b67560152/t/55ef3dcbe4b0adc4e323efbc/1441742283987/Rural_Ed_Assoc-v-State_District_Order_re_Mootness_Issue_1993.pdf; http://leg.mt.gov/content/committees/interim/2005_2006/edu_local_gov/minutes/02242006exhibits/ELG02242006_ex5.pdf.

Suit was filed in 2008 seeking supplemental monetary relief to help districts avoid funding shortfalls in 2009. In December 2008, the district court declined to award any supplemental relief, so we do not code 2008 as a court-ordered school finance reform.

New Hampshire[8]
2006: JJP court order; LRS no event

In September 2006 in *Londonderry School District v. State* the New Hampshire Supreme Court ordered the state to define a "constitutionally adequate education" by June 2007. After recounting the failure to establish this definition in several previous cases (both JJP and LRS code court orders in 1993, 1997, and 1999), the court concluded that it is willing to defer to the legislature one more time, and that "in the absence of action…, a judicial remedy is not only appropriate, but essential" in order to vindicate the constitutional rights of New Hampshire's students.

In the 2006 decision, the Court ordered the State to define a "constitutionally adequate education" by the end of the 2007 legislative session, but deferred to the legislature for appropriate action. We code the 2008 legislative action but not the 2006 court order.

New Jersey[9]
1991: JJP court order; LRS no event
1998: JJP no event; LRS court order
2000: JJP no event; LRS court order

New Jersey's school finance litigation history is extremely complex, with a decades-long exchange between the legislature and the courts. There have been many, many rulings in the *Abbott v. Burke* case in particular. The court ruled in *Abbott II* in 1990 (counted by both JJP and LRS) that state funding statutes failed to ensure adequate funding in the low-wealth "Abbott districts", and noted that students in these districts need programs and services beyond those provided to students in wealthier districts. In response, the legislature passed the Quality of Education Act of 1990 (QEA; LRS code this as a legislative event).

---

[8]Case history from https://www.nhbar.org/publications/display-journal-issue.asp?id=365.
[9] Case histories drawn from
http://www.schoolfunding.info/resource_center/legal_docs/New%20Jersey/Abbott%20Decisions/Abbott-SupremeCourt_May1997.PDF**;**
http://www.schoolfunding.info/resource_center/legal_docs/New%20Jersey/Abbott%20Decisions/Abbott-SupremeCourt-May-1998.PDF;
http://www.schoolfunding.info/resource_center/legal_docs/New%20Jersey/Abbott%20Decisions/Abbott-SupremeCourt-Feb-2002.PDF.

In 1991, the plaintiffs applied to the court to declare the QEA unconstitutional. The court declined to hear the motion at that time. JJP code this as a court order, but we do not. The court did find the QEA unconstitutional in the 1994 *Abbott III* ruling; both JJP and LRS count this event.

In 1998's *Abbott V* ruling, the court required the state to increase funding to ensure parity in per-pupil expenditures between the Abbott districts and the average of the state's 110 successful suburban school districts, and directed the state to conduct a study to determine the needs of Abbott students and the programs necessary to meet those needs. Based on the State's study, the court ordered additional remedial measures for the Abbott children, including preschool for all three- and four-year olds, adequate school facilities, and "supplemental" programs. We code this as a school finance reform, though JJP do not.

After plaintiffs brought another motion alleging the state did not comply with the *Abbott V* ruling, the court provided (in the 2000 *Abbott VI* ruling) more detail on the preschool requirements, including substantive educational standards, certified staff, and a maximum student/teacher ratio of 15:1. We code this as a school finance reform; again, JJP do not.

New Mexico[10]
1998: JJP court order; LRS no event
1999: JJP no event; LRS court order

In 1998, a number of districts brought a capital funding/facilities suit, *Zuni School District v. State,* CV-98-14-II (Dist. Ct., McKinley County Oct. 14, 1999), claiming that the funding system for capital items was unconstitutional. The trial court granted partial summary judgment in favor of plaintiffs and ordered the state to "establish and implement a uniform funding system for capital improvements . . . and for correcting existing past inequities."

The case was filed in 1998 but decided in 1999. JJP code it as a 1998 event, but we code it as a 1999 event based on the decision date.

Oregon[11]
2009: JJP court order; LRS no event

In January 2009, the Oregon Supreme Court found in *Pendleton School District 16R v. State* that the legislature had, in violation of a 2000 constitutional amendment, failed to fund the Oregon public school system at a level sufficient to meet the quality education goals established by law. However, it concluded that the state

---

[10] Case history from http://www.schoolfunding.info/states/nm/lit_nm.php3; https://nces.ed.gov/edfin/pdf/lawsuits/Zuni_v_%20nm.pdf; http://ielp.rutgers.edu/resources/New_Mexico.
[11] Case history from http://www.educationjustice.org/states/oregon.html.

constitution did not give the court authority to issue an injunction requiring the state to provide sufficient funding to reach those goals. Because the court ruled that the law was not judicially enforceable, and no subsequent legislative actions were taken, we do not code this event as a school finance reform.

South Carolina[12]

2005: JJP court order; LRS no event

In 1999, in *Abbeville County Sch. Dist. v. State*, the South Carolina Supreme Court held that plaintiffs had a valid claim under the state constitution's education clause, interpreted the clause to mean that the legislature must provide children with a "minimally adequate education," and remanded the case for trial. The lower court ruled in 2005 that the state's failed to meet its constitutional requirement by inadequately providing early education programs, but ruled against plaintiff claims requesting relief regarding school buildings and quality teaching. Because the court did not order substantial school finance reform, we do not code a 2005 event.

Both plaintiffs and defendants appealed to the South Carolina Supreme Court, which heard oral argument in 2008 and again in 2012. In 2014, the state supreme court held the state's school funding unconstitutional, declaring that "South Carolina's education funding scheme is a fractured formula denying students ... the constitutionally required opportunity." The court explained that the resources provided failed to produce sufficient educational opportunities. The court explicitly refrained from mandating how the state should remedy the system, but ordered the parties to work together to present a new funding system to the court "within a reasonable time." The 2014 court order meets our definition of a court-ordered school finance reform, but is outside of our sample period so is not included in our tabulation (or in JJP's).

Texas[13]

2004: JJP court order; LRS no event.

A trial court found in *West Orange-Cove Consolidated ISD v. Nelson* (2004) that the Texas school finance system failed to provide "an adequate, suitable and efficient education system" as required by the state constitution, and additionally found the state property tax to be unconstitutional.

In 2005, the state Supreme Court ruled in *Neeley v. West Orange-Cove Indep. Sch. Dist.* that the state property tax was unconstitutional, but held that despite funding inequities the state's education finance system did not violate the constitutional adequacy, efficiency, and suitability requirements. The court wrote that the school finance system displayed deficiencies that could in time render it unconstitutional

---

[12] Case history from http://www.educationjustice.org/states/southcarolina.html.

[13] Case history from http://caselaw.findlaw.com/tx-supreme-court/1153227.html; http://www.schoolfunding.info/states/tx/McCown.pdf.

under the education article. Because the supreme court did not order reform's, we do not include this case.

Washington[14]
1991: JJP court order; LRS no event
2007: JJP court order; LRS no event
2010: JJP no event; LRS court order

Seattle School District v. State, also known as *Seattle II*, was a 1983 trial court ruling following up on the 1978 *Seattle I* decision that prompted an overhaul of the school finance system and the introduction of the Basic Education Act. *Seattle II* expanded the definition of "basic education" in the state to include special education, and bilingual and remedial programs. The state did not appeal, and the legislature amended the school finance system to include funding for these programs. JJP date this case to 1991. To our knowledge, it occurred in 1983, so does not fall into our sample period.

In *Federal Way Sch. Dist. v. State*, filed in 2006, plaintiffs alleged that the state funding system failed to amply fund education in all school districts and was unconstitutional. In 2007, Judge Michael Heavey held in favor of plaintiffs, finding that the State's method of providing salary funding was unconstitutional. The state Supreme Court, however, issued a narrower ruling in 2009 that a "uniform system" of education governs educational content, teacher certification, instructional hour requirements and the assessment system, but does not require uniform funding of staff salaries. The court did not rule on whether the plaintiffs had "ample" funds under the state constitution. Because the 2009 Supreme Court ruling did not involve finances, we do not code this as an event.

McCleary v. State, filed in 2007, argued that although the state had developed standards for a constitutional "basic education," it was not fully funding that education. In 2010, the Superior Court held that the state funding system was unconstitutional because it neither determined the cost of nor provided the resources needed for a basic education for all children in the state. The court ordered the state to fund a constitutionally adequate education, using stable and dependable state sources. In response, the legislature enacted legislative reforms, and in early 2012 the Washington Supreme Court affirmed the Superior Court ruling. We code the 2010 event, as the legislature acted on it without waiting for it to be upheld by the Supreme Court.

---

[14] Case history drawn from
https://www.courts.wa.gov/opinions/pdf/843627.opn.pdf;
http://digitalcommons.law.seattleu.edu/cgi/viewcontent.cgi?article=2290&context=sulr;
https://www.courts.wa.gov/opinions/pdf/843627.opn.pdf;
http://www.schoolfunding.info/states/wa/lit_wa.php3.

References

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.

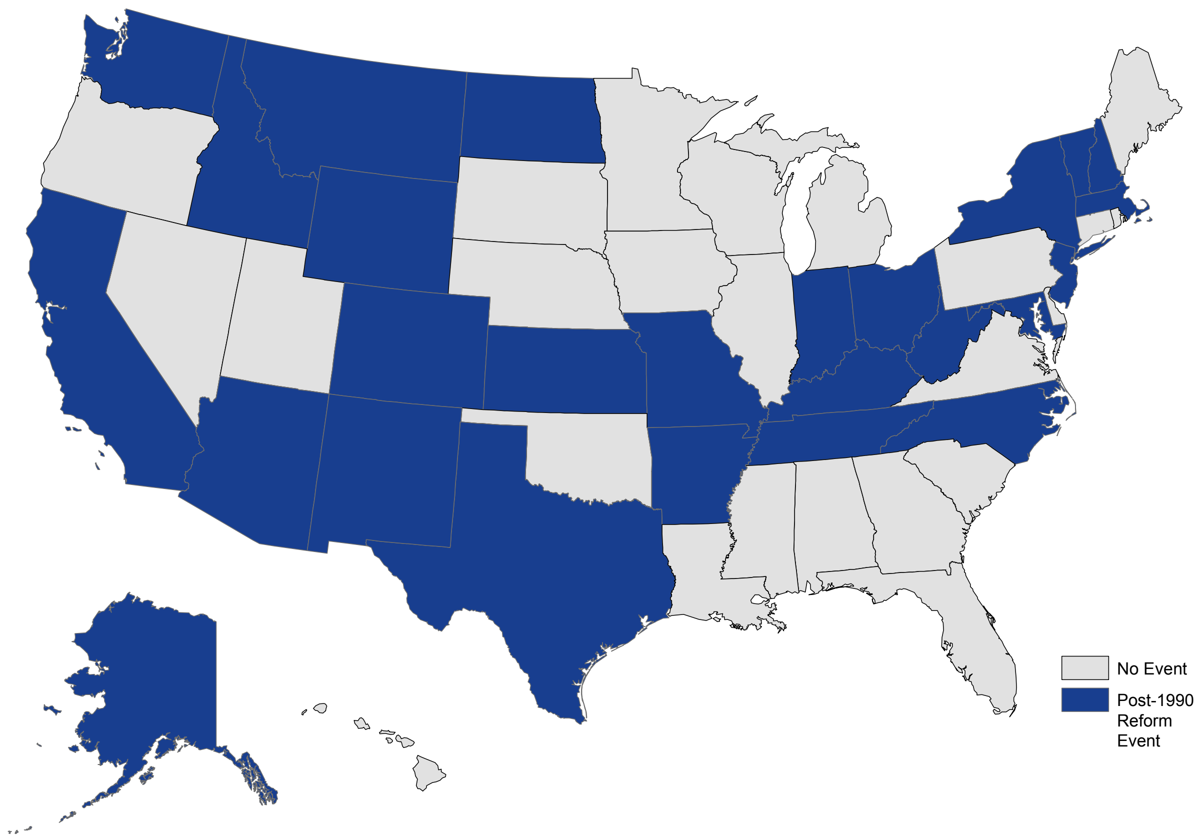
Corcoran, S. P., & Evans, W. N. (2015). Equity, adequacy, and the evolving state role in education finance. In H. F. Ladd and M. E. Goertz, eds., *Handbook of Research in Education Finance and Policy,* 2nd edition. New York: Routledge.

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), 157-218.

Jacob, B. & Rothstein, J. (2016). The measurement of student ability in modern assessment systems. *Journal of Economic Perspectives*, forthcoming.

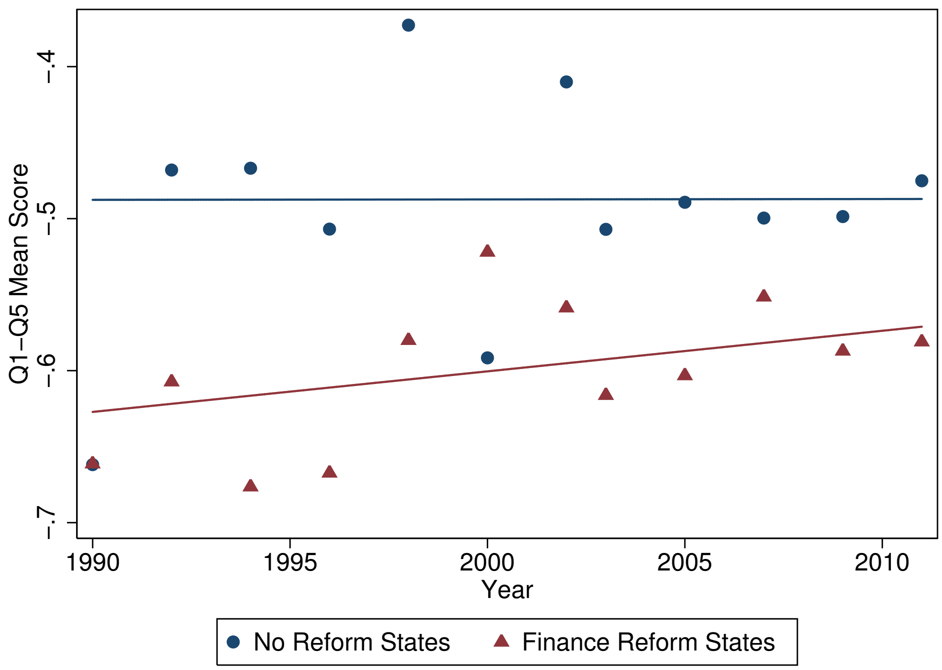
Appendix Figures

Figure A1: Geographic distribution of post-1989 school finance events



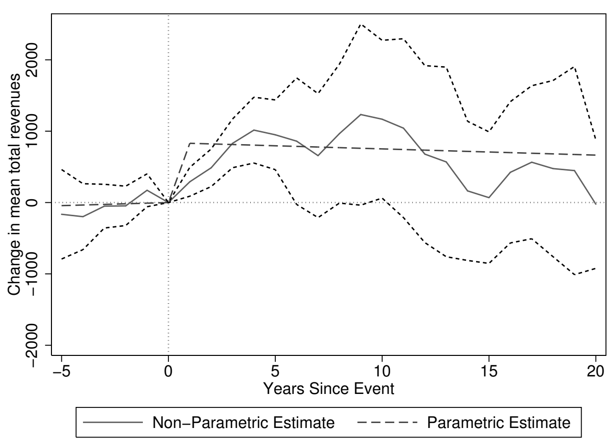
Notes: Map indicates states that had school finance reform events, as listed in Appendix Table A1, between 1990 and 2011.

Figure A2: Gap in average test scores between lowest and highest income districts, by state finance reform status, 1990-2011

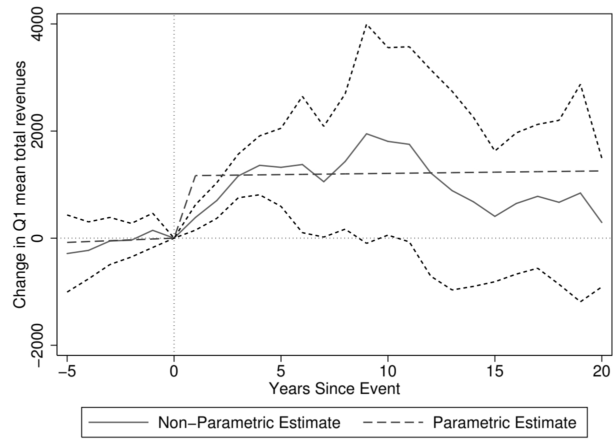


Notes: Lowest (Q1) and highest (Q5) income districts are defined as in Figure 1. NAEP observations in districts in each quintile are averaged, using NAEP sampling weights and separately for each grade and subject tested, and the Q1-Q5 difference is computed for each state. State-grade-subject Q1-Q5 differences are averaged separately for each group of states, weighting by the harmonic mean of the sum of the student weights in Q1 and Q5 districts. Lines show best linear fit to the time series.

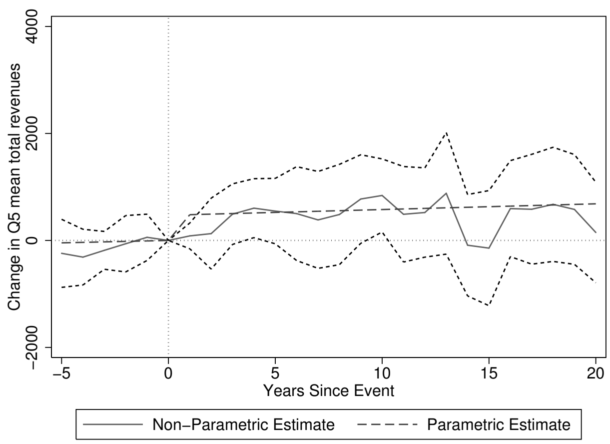
Figure A3: Event study estimates of effects of school finance reforms on mean total revenues in lowest and highest income districts



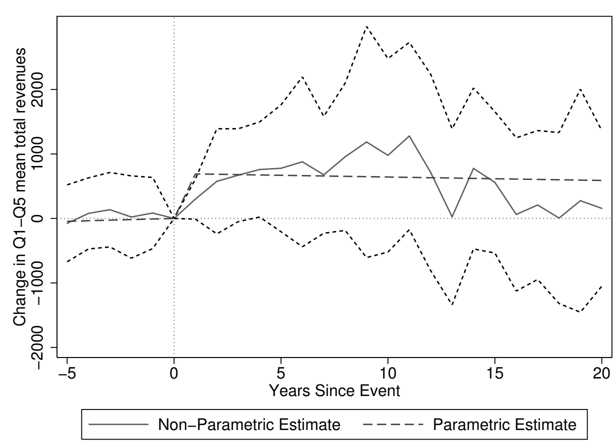
(a) Total revenue, mean



(b) Total revenue, Q1



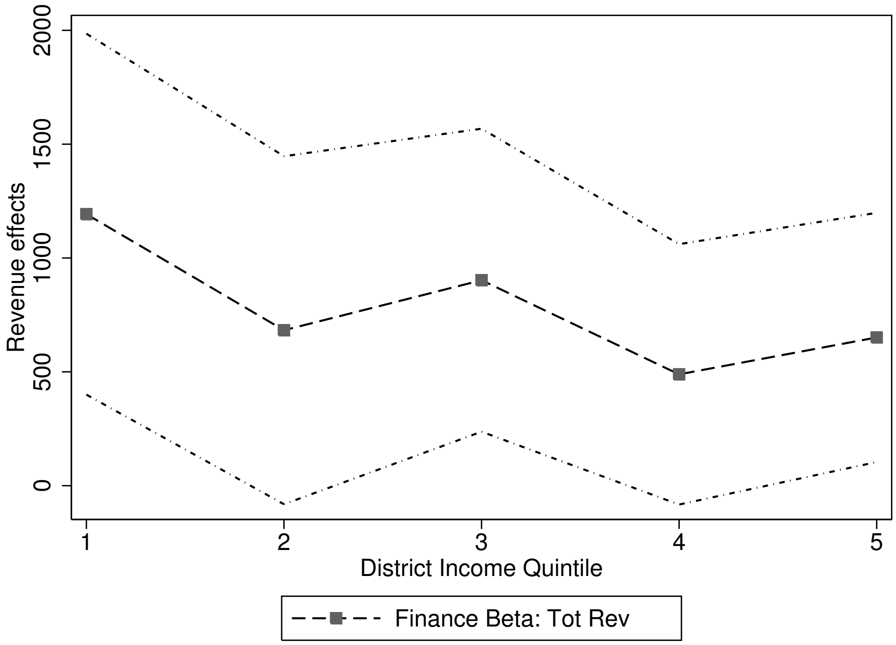
(c) Total revenue, Q5



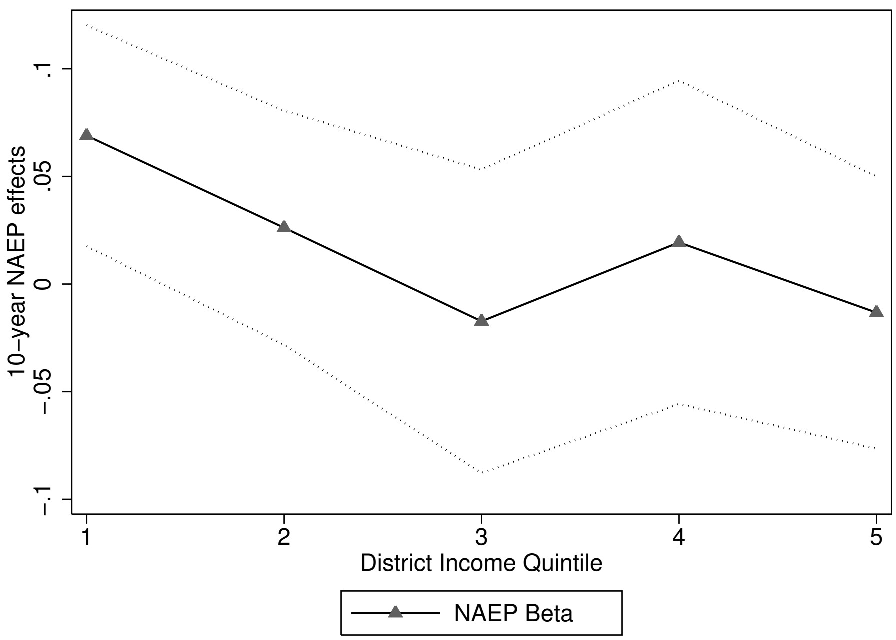
(d) Total revenue, Q1-Q5

Notes: Figure displays coefficients from event study regressions. Dependent variables are mean total revenues per pupil (panel A), mean total revenues per pupil in the lowest income quintile of districts (panel B), mean total revenues per pupil in the highest income quintile of districts (panel C), and the difference in mean total revenues per pupil between districts in the bottom and top income quintile in the state (panel D), all measured in 2013 dollars per pupil. Dashed lines show the three-parameter parametric model (equation 2). Solid lines shows the non-parametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 3, panel D, columns 1-4. p values for omnibus hypothesis tests of zero pre-event effects in the non-parametric model in panels A-D are 0.15, 0.40, 0.74, and 0.86, respectively; p-values for zero post-event effects are <0.001 in all panels. In the parametric model, the p-values for the hypothesis that the pre-event trend is zero are 0.79, 0.68, 0.78, and 0.72; for the test that the post-event jump and change in trend is zero they are 0.01, <0.001, 0.22, and 0.01.

Figure A4: Event study estimates for total revenues and test scores by district income group



(a) Total revenue



(b) NAEP

Notes: Figure shows event study estimates from one-parameter parametric models for mean revenues and mean test scores in each quintile. Estimates for quintiles 1 and 5 are shown in Table 3, panel C, columns 1-2, and Table 5, columns 3 and 4. 95% confidence intervals shown by dotted lines. Standard errors are clustered at the state level.

Appendix Tables

Appendix Table A1
Complete Event List, 1990-2011

State	Year	Event	Lafortune, Rothstein & Schanzenbach (2016)	Jackson, Johnson & Persico (2016)	Corcoran & Evans (2015)
Alabama	1993	Alabama Coalition for Equity (ACE) v. Hunt; Harper v. Hunt		X	
Alaska	**1999**	Kasayulie v. State of Alaska	Court	X	
Arizona	1994	Roosevelt v. Bishop	Court	X	
	1997	Hull v. Albrecht	Court	X	
	1998	Hull v. Albrecht	Court	X	
	2007	Flores v. Arizona		X	
Arkansas	1994	Lake View v. Arkansas	Court	X	
	1995	Approved Equitable School Finance Plan (Acts 917, 916, and 1194)	Bill		n/a
	2002	Lake View v. Huckabee	Court	X	X
	2005	Lake View v. Huckabee	Court	X	X
	2007	Various acts resulting from Master's Report findings	Bill		n/a
California	1998	Leroy F. Greene School Facilities Act of 1998	Bill		n/a
	2004	Senate Bill 6, Senate Bill 550, Assembly Bill 1550, Assembly Bill 2727, and Assembly Bill 3001	Bill		n/a
Colorado	**2000**	Bill 181; Various Other Acts	Bill		n/a
Connecticut	1995	Sheff v. O'Neill		X	
	2010	Coalition for justice in Education Funding, Inc. v. Rell		X	n/a
Idaho	**1993**	Idaho Schools for Equal Educational Opportunity v. Evans (ISEEO)	Court		
	1994	Senate Bill 1560	Bill		n/a
	1998	Idaho Schools for Equal Educational Opportunity v. State (ISEEO III)		X	
	2005	Idaho Schools for Equal Educational Opportunity v. Evans (ISEEO V)	Court	X	
Indiana	**2011**	HB 1001 (Pl229)	Bill		n/a

(continued)

Appendix Table A1 (continued)

State	Year	Event	LRS (2016)	JJP (2016)	CE (2015)
Kansas	1992	The School District Finance and Quality Performance Act	Bill		n/a
	2005	Montoy v. State; Montoy v. State funding increases	Both	X	X
Kentucky	(1989)	Rose v. Council for Better Education, Inc.	Court	X	X
	1990	Kentucky Education Reform Act (HB 940)	Bill		n/a
Maryland	1996	Bradford v. Maryland State Board of Education	Court		
	2002	Bridge to Excellence in Public Schools Act (BTE) (Senate Bill 856)	Bill		n/a
	2005	Bradford v. Maryland State Board of Education		X	(upheld)
Massachusets	**1993**	McDuffy v. Secretary of the Executive Office of Education; Massachusetts Education Reform Act	Both	X	X
Michigan	1997	Durant v. State of Michigan		X	
Missouri	**1993**	Committee for Educational Equality v. State of Missouri; Outstanding Schools Act (S.B. 380)	Both	X	
	2005	Senate Bill 287	Bill		n/a
Montana	1993	House Bill 667	Bill	X	
	2005	Columbia Falls Elementary School v. State	Court	X	X
	2007	M.C.A. § 20-9-309	Bill		n/a
	2008	Montana Quality Education Coalition v. Montana		X	n/a
New Hampshire	1993	Claremont New Hampshire v. Gregg	Court	X	
	1997	Claremont School District v. Governor	Court	X	X
	1998	Opinion of the Justices--School Financing (Claremont III)			X
	1999	Claremont v. Governor (Claremont III); RSA chapter 193-E	Both	X	X
	2000	Opinion of the Justices--School Financing (Claremont VI)			X
	2002	Claremont School District v. Governor	Court	X	X
	2006	Londonderry School District v. New Hampshire		X	
	2008	SB 539	Bill		n/a

(continued)

Appendix Table A1 (continued)

State	Year	Event	LRS (2016)	JJP (2016)	CE (2015)
New Jersey	1990	The Quality Education Act; Abbot v. Burke	Both	X	X
	1991	Abbott v. Burke		X	
	1994	Abbott v. Burke	Court	X	X
	1996	Comprehensive Educational Improvement and Financing Act of 1996	Bill		n/a
	1997	Special Master's Report; Abbott v. Burke	Bill		X
	1998	Abbott v. Burke	Court		X
	2000	Abbott v. Burke	Court		
	2008	The School Funding Reform Act of 2008	Bill		n/a
New Mexico	1998	Zuni School District v. State		X	
	1999	Zuni School District v. State	Court		
	2001	Deficiencies Corrections Program; Public School Capital Outlay Act	Bill		n/a
New York	2003	Campaign for Fiscal Equity, Inc. v. State	Court	X	X
	2006	Campaign for Fiscal Equity, Inc. v. State	Court	X	
	2007	Education Budget and Reform Act	Bill		n/a
North Carolina	**1997**	Leandro v. State	Court	X	
	2004	Hoke County Board of Education v. State	Court	X	X
North Dakota	**2007**	SB 2200	Bill		n/a
Ohio	**1997**	DeRolph v. Ohio	Court	X	X
	2000	DeRolph v. Ohio; Increased school funding (see 93 Ohio St.3d 309)	Both	X	X
	2001	DeRolph v. Ohio			X
	2002	DeRolph v. Ohio	Court	X	X
Oregon	2009	Pendleton School District 16R v. State		X	n/a
South Carolina	2005	Abbeville County School District v. State		X	

(continued)

Appendix Table A1 (continued)

State	Year	Event	LRS (2016)	JJP (2016)	CE (2015)
Tennessee	1992	The Education Improvement Act	Bill		n/a
	1993	Tennessee Small School Systems v. McWherter	Court	X	X
	1995	Tennessee Small School Systems v. McWherter	Court	X	X
	2002	Tennessee Small School Systems v. McWherter	Court	X	X
Texas	1991	Edgewood Independent School District v. Kirby	Court	X	X
	1992	Carrolton-Farmers Branch ISD v. Edgewood Independent School District	Court	X	X
	1993	Senate Bill 7	Bill		n/a
	2004	West Orange-Cove ISD v. Nelson		X	
	2005	West Orange-Cove Consolidated ISD v. Neeley			X
Vermont	1997	Brigham v. State	Court	X	X
	2003	Revisions to Act 68; H.480	Bill		n/a
Washington	1991	Seattle II		X	
	2007	Federal Way School District v. State		X	
	2010	McCleary v. State	Court	n/a	n/a
West Virginia	**1995**	Tomblin v. Gainer	Court	X	
Wyoming	1995	Campbell County School District v. State	Court	X	X
	1997	The Wyoming Comprehensive Assessment System; The Education Resource Block Grant Model	Bill		n/a
	2001	Campbell II; Recalibration of the MAP model	Bill	X	n/a

Notes: Table lists all events included in any of the Lafortune-Rothstein-Schanzenbach (2016); Jackson-Johnson-Persico (2016); or Corcoran-Evans (2015) event lists, from 1990 onward. Xs indicate events that appear in the relevant event list; n/a indicates events that were out of scope for the relevant list, either because they were too recent or because it included only court cases and not legislative events. In Lafortune et al. column, events are classified as "court," "bill," or "both"; rows without an entry are not included in our event database but are included in one of the comparison samples. Bold years indicate the single event per state selected by our algorithm (see text). Appendix D discusses discrepancies between Lafortune et al. and Jackson et al. lists.

Table A2: Event studies for state budgets

	Per capita	Per pupil
Tax revenues:		
Total revenues	235	2,736
	(258)	(2,044)
Expenditures:		
General expenditures	290**	2,536*
	(138)	(1,505)
Education expenditures	114	1,029
	(70)	(643)
General expenditures (less education)	176**	1,508
	(90)	(977)
Health + welfare expenditures	73	514
	(49)	(457)
General expenditures (less education, health, welfare)	103	993
	(77)	(700)

Notes: Table shows estimates from the one-parameter event study specification (equation (1)) for state budgetary aggregates. State and year fixed effects are included. Standard errors are clustered at the state level.

Table A3: Comparison to Card-Payne

	State revenues (per capita)			Total revenues (per capita)		
	1997-1992 (CP)	1990-2012 (LRS)		1997-1992 (CP)	1990-2012 (LRS)	
	Long diff	Long diff	Event study	Long diff	Long diff	Event study
Court Ruling:						
Upheld	-0.81			0.20		
	(0.67)			(0.52)		
Unconstitutional	-1.89***			-1.10**		
	(0.62)			(0.48)		
Selected Events:						
Post Event		-2.06	-2.25**		-2.44	-1.61
		(2.24)	(0.89)		(4.73)	(2.38)

Notes: This table shows results using slopes from a regression of per capita state or total funding on district mean household income (note: district mean income here is in *levels*, not logs). Columns 1 and 4 are from table 4 of Card and Payne (2002) and show the long difference from 1977-1992 in the level-level slope coefficient. In columns 2 and 5, we replicate the Card and Payne specification using data from 1990 and 2012. Columns 3 and 6 show estimated effects from the one parameter event study specification (equation (1)) where level-level per capita slope coefficients are the dependent variables. Standard errors are clustered at the state level.

Table A4: Event study for log income, race, free lunch

(a) Income gradients

	Log mean income		Minority share		Free lunch share	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Event * Yrs Elapsed	-0.0010	0.0008	0.0022	0.0017	0.0058	0.0086
	(0.0029)	(0.0040)	(0.0013)	(0.0016)	(0.0065)	(0.0072)
Trend		-0.0026		0.0009		-0.0019
		(0.0042)		(0.0007)		(0.0031)
Post Event		0.0193		-0.0041		-0.0266
		(0.0368)		(0.0056)		(0.0289)
Observations	147	147	1045	1045	957	957
p(post-event=post-event*trend=0)	0.72	0.87	0.11	0.56	0.37	0.49
State FEs	X	X	X	X	X	X
Yr FEs	X	X	X	X	X	X

(b) Q1-Q5 difference

	Log mean income		Minority share		Free lunch share	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Event * Yrs Elapsed	-0.0017	-0.0008	-0.0014	-0.0019	-0.0034*	-0.0049*
	(0.0029)	(0.0035)	(0.0018)	(0.0020)	(0.0020)	(0.0025)
Trend		-0.0004		0.0004		0.0016
		(0.0035)		(0.0016)		(0.0022)
Post Event		-0.0073		0.0009		-0.0017
		(0.0290)		(0.0088)		(0.0149)
Observations	145	145	1045	1045	962	962
p(post-event=post-event*trend=0)	0.55	0.95	0.41	0.64	0.09	0.11
State FEs	X	X	X	X	X	X
Yr FEs	X	X	X	X	X	X

Notes: Table presents event study specifications where the dependent variable is the slope of the indicated demographic characteristic with respect to the district's 1990 log mean household income (panel A) or the gap between the average for districts in the bottom and top quintiles of the 1990 income distribution (panel B). Minority share and free lunch share are available annually from the Common Core of Data (though missing in some states and some years); log mean income is available from the Census in 1990 and 2000 and from the American Community Survey in 2007-11 (coded as 2011). Standard errors are clustered at the state level.

Table A5: Stratification of race, FRL, & achievement, by quintile

	Q1	Q2	Q3	Q4	Q5
Black	0.24	0.24	0.24	0.17	0.11
Black/Hispanic	0.25	0.23	0.24	0.18	0.11
White	0.20	0.20	0.18	0.20	0.22
Free/reduced-price lunch	0.32	0.23	0.20	0.15	0.10
25th pctl or below (NAEP)	0.27	0.21	0.22	0.17	0.13
75th pctl or above (NAEP)	0.14	0.15	0.17	0.22	0.32

Note: Table shows fraction of students of various groups in districts in various quintiles of the state's district income distribution. Each row sums to 1. Racial and free lunch shares are computed using CCD district-level data for the year 1994. The distribution of high- and low-achieving students is based on the 2003 NAEP data, which is the first year of comprehensive data for all grades and subjects.

Table A6: Event studies for district-mean resource gaps by race, FRL, & achievement

	Black/White		Free Lunch		25th/75th Pctl (NAEP)	
	St. Rev	Tot. Rev	St. Rev	Tot. Rev	St. Rev	Tot. Rev
Post Event	197	195	2	23	143	193*
	(160)	(164)	(185)	(195)	(141)	(101)
Observations	1047	1047	938	938	1509	1509
State FEs	X	X	X	X	X	X
Yr FEs	X	X	X	X		
Sub-gr-yr FEs					X	X

Note: In columns 1 and 2, the dependent variable in event study specifications is the average per-pupil revenue in the district attended by the average black student, less that in the district attended by the average white student in the same state. In columns 3 and 4, analogous revenue gaps are constructed for free/reduced-price lunch and non-free/reduced-price lunch students. In columns 5 and 6, analogous revenue gaps are constructed for students scoring at or below the 25th percentile in the NAEP, and students scoring at or above the 75th percentile in the NAEP. The *Post Event* coefficient shows the estimated event effect from parametric event study model without controlling for prior trends. State and year fixed effects are included in columns 1-4. State and grade-subject-year fixed effects are included in columns 5 and 6. Standard errors are clustered at the state level.

Table A7: Impacts of student sorting on student achievement results

	Q1-Q5 difference	Slope
Baseline Estimates	0.008**	-0.010***
	(0.004)	(0.003)
Decomposition 1: Common covariates		
Predicted score	0.003	-0.003
	(0.004)	(0.004)
Residual score	0.005**	-0.007**
	(0.002)	(0.003)
Decomposition 2: Richer covariates		
Predicted score	0.004	-0.004
	(0.004)	(0.003)
Residual score	0.004*	-0.006***
	(0.002)	(0.002)

Notes: First row repeats estimates from Table 6, columns 2 and 5. In subsequent rows, dependent variables are modified. We estimate student-level regressions of NAEP scores on student demographic characteristics, with year fixed effects, then compute predicted and residual test scores. We compute separate slopes with respect to district income and quintile gaps for the predicted and residual test scores, and estimate separate event study regressions for each. In decomposition 1, student demographic characteristics are race/ethnicity and gender, along with school means (in the NAEP sample) of each. Decomposition 2 adds indicators for students whose parent is a college graduate and for free or reduced-price lunch receipt, along with indicators for NAEP samples where these variables are unavailable and school means of each. Standard errors are clustered at the state level.

Table A8: Multiple events robustness: Monte Carlo simulations

	First event	All events (stacked)
DGP 1: Constant event effect		
Post coefficient	0.789	0.577
DGP 2: Only one event		
Post coefficient	0.788	0.577

Notes: Table reports estimates of average post-event "jump" coefficient from Monte Carlo simulations using the empirical distribution of event dates, in which some states had multiple school finance reform events. Column 1 shows estimates from event study models estimated using only the first event in a state. Column 2 shows estimates using all events in a state, stacking panels and adding a joint state-panel copy fixed effect (see table 7 column 2). In both columns, estimates are from parametric event study models with a single coefficient (equation 1). Row 1 shows estimates from a simulated DGP where every event in a state has a constant effect. Row 2 shows estimates from a DGP where only one event (randomly chosen within state) has an effect. In both DGPs the total event effect over all events within a state is equal to 1. All DGPs include i.i.d. error terms and are simulated 5000 times.