School Finance Reform and the Distribution of Student Achievement[†]

By Julien Lafortune, Jesse Rothstein, and
Diane Whitmore Schanzenbach*

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. (JEL H75, I21, I22, I24, I28)

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).[1] 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)" (Hanushek 1997, 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

[1] There are also observational (Card and Krueger 1992a) and experimental (Krueger 1999; Dynarski, Hyman, and 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).
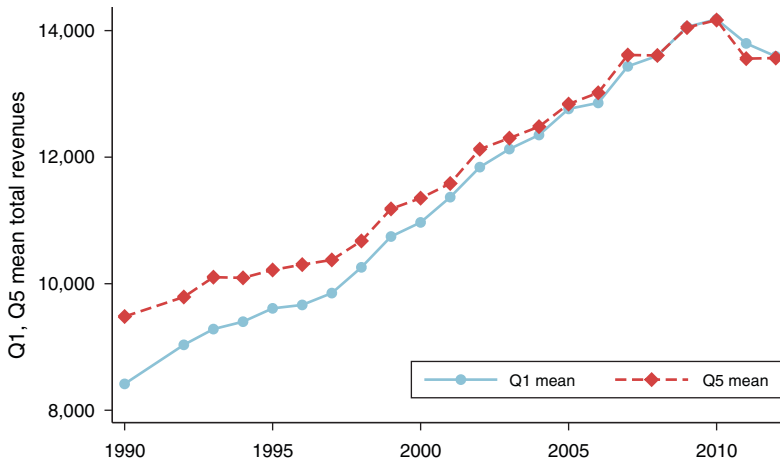
FIGURE 1. MEAN REVENUES PER PUPIL FOR HIGHEST AND LOWEST INCOME SCHOOL DISTRICTS, 1990–2012

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

district-level mean household income distribution).[2] Between 1990 and 2012, real per pupil revenues rose by roughly 30 percent in the highest income districts, and by over 50 percent in the lowest income districts. Thus, while low-income districts collected about 20 percent 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 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.[3]

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.

---

[2] 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.

[3] Studies of the implications of SFRs for school finance, mostly examining equity reforms, include Murray, Evans, and Schwab (1998); Card and Payne (2002); Hanushek and Lindseth (2009); Berry and Wysong (2012); Ladd and Goertz (2015); Sands (2015).
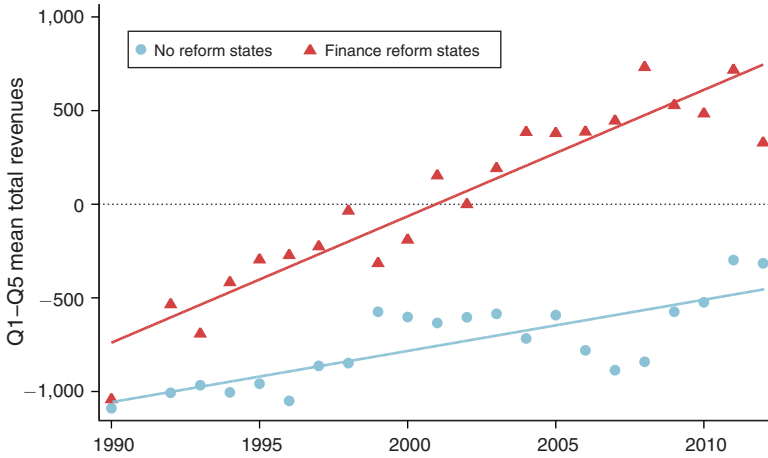
FIGURE 2. GAP IN REVENUES PER PUPIL BETWEEN LOWEST AND HIGHEST INCOME DISTRICTS, BY STATE FINANCE
REFORM STATUS, 1990–2012

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

What evidence there is derives from nonrepresentative 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 (Guryan 2001, Clark 2003, Hyman forthcoming).[4] 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 (Hanushek 1986, 2003, 2006; Card and Krueger 1992a; 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.

---

[4] Cascio, Gordon, and Reber (2013) and Cascio and Reber (2013) examine the introduction of federal Title I funding to low-income schools via the 1965 Elementary and Secondary Education Act.

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.[5] 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 and smaller but still positive increases in high-income districts.[6] 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.[7] 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 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.

---

[5] 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.

[6] 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.

[7] STAR raised costs by about 30 percent in K–3, and raised early grade test scores by 0.17 SDs (Krueger 1999, 2003; Krueger and Whitmore 2001). 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.

I. School Finance Reforms[8]

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 1968; 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.[9] 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., Murray, Evans, and Schwab 1998; Card and Payne 2002; Hanushek and Lindseth 2009; Corcoran and Evans 2015).

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*;[10] 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.[11]

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

---

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

[9] The US Supreme Court held in 1973 that education is not a fundamental right under the US 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.

[10] 790 SW 2d 186. *Rose* was not the first adequacy ruling, but earlier rulings attracted less attention.

[11] 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.

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.

Online 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.[12] Thirty-nine (61 percent) 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.

Eighteen states had multiple events in our period. These were generally closely spaced: 60 percent 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 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:

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

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.

---

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

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 nonrandom, 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):

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

$\beta^{phasein}$ 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 nonparametric 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 nonparametric 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 nonparametric 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, Rivkin, and Taylor 1996 a, 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 percent 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.[13] A more negative slope corresponds to higher relative outcomes in low-income districts. For both 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.

A. 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:

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

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.[14] 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 nonzero 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

---

[13] 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 online Appendix for further detail.

[14] We restrict attention to t^* for which the estimated κ has the expected sign.

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 nonevents.

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.[15] In Monte Carlo simulations (see online 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 nonzero effect. Our final approach follows the prior literature, which generally emphasizes simple specifications 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.[16] 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 online 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 fourth and eighth grade assessments in math and reading in every odd-numbered year.[17] Table 1 shows the schedule. Tests are administered to around 100,000 students (more in later years) in each subject-grade-year. These consist

---

[15] Results are unchanged when data are reweighted to offset the overrepresentation of states with multiple events.

[16] Census data are available in 1989–1990 and 1991–1992, and annually since 1994–1995. We use samples from the Census Bureau's Annual Survey of Government Finances for 1992–1993 and 1993–1994.

[17] The NAEP also tests twelfth graders, but samples are smaller, and other subjects.

TABLE 1—NAEP TESTING YEARS

Year	Subjects and grades covered				Number of states	Number of students
	Math G4	Math G8	Reading G4	Reading G8		
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

Note: 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	
	Observations	Mean	SD	Q1	Q5
Enrollment	229,386	67,523	181,811	13,537	31,403
log(mean income, 1990)	223,334	10.53	0.2935	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
Noninstructional	229,386	5,419	2,221	5,178	5,949
NAEP scores	49,867	0.2559	0.4578	0.02925	0.5884

Note: 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.

of representative samples of about 3,500 students per state, spread across about 140 schools in 80 districts.

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 district finance and demographics data.[18]

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

---

[18] 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.
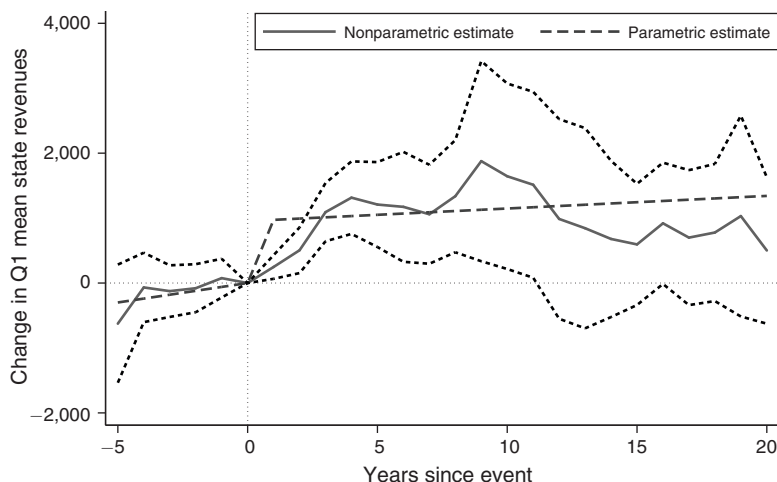
Notes: Figure displays coefficients from event study regressions. Dependent variables are mean state revenues in the lowest income quintile of districts, measured in 2013 dollars per pupil. Dashed lines show the three parameter parametric model (equation 2). Solid lines shows the nonparametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95 percent confidence intervals. Estimates for the parametric models are reported in Table 3, column 1, panel B. The *p*-value for the omnibus hypothesis test of zero pre-event effects in the nonparametric model 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.

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 nonparametric event study specification (3), while dotted lines show pointwise 95 percent 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 fourth 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 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
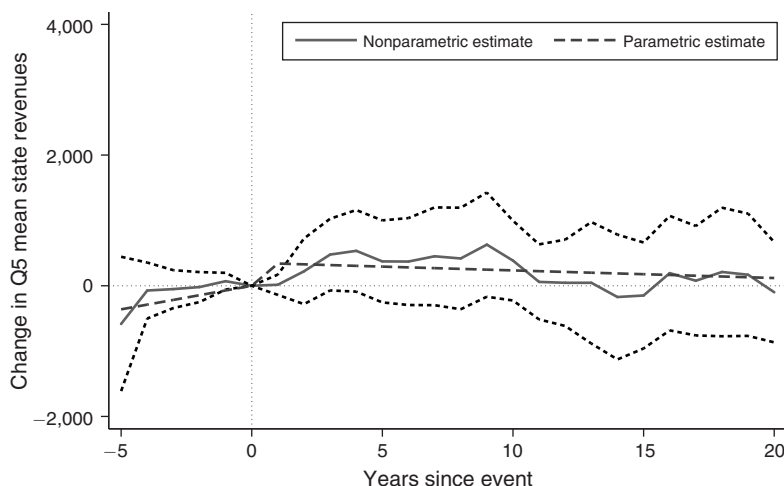
FIGURE 4. EVENT STUDY ESTIMATES OF EFFECTS OF SCHOOL FINANCE REFORMS ON MEAN STATE REVENUES IN HIGHEST INCOME DISTRICTS

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 nonparametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95 percent confidence intervals. Estimates for the parametric models are reported in Table 3, column 2, panel B. The *p*-value for the omnibus hypothesis test of zero pre-event effects in the nonparametric 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.

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 (online Appendix Figure A3) are also similar.

In additional analyses of state budgets (online 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.

TABLE 3—EVENT STUDY ESTIMATES OF EFFECTS OF SCHOOL FINANCE REFORMS ON REVENUES PER PUPIL

	Q1	Q5	All districts	Q1 − Q5 difference	Slope
Panel A. State revenue (1 parameter)					
Post event	1,225	527	912	711	−622
	(343)	(378)	(359)	(316)	(223)
Panel 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 × years elapsed	−40	−84	−61	42	−5
	(70)	(61)	(60)	(36)	(21)
Panel C. Total revenue (1 parameter)					
Post event	1,233	544	829	701	−424
	(370)	(277)	(302)	(309)	(304)
Panel 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 × years 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.

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.[19]

---

[19] When we estimate specifications similar to Card and Payne's (2002) closely related analysis of earlier SFRs (online Appendix Table A3), estimated SFR effects are slightly larger but imprecise, and well within the earlier
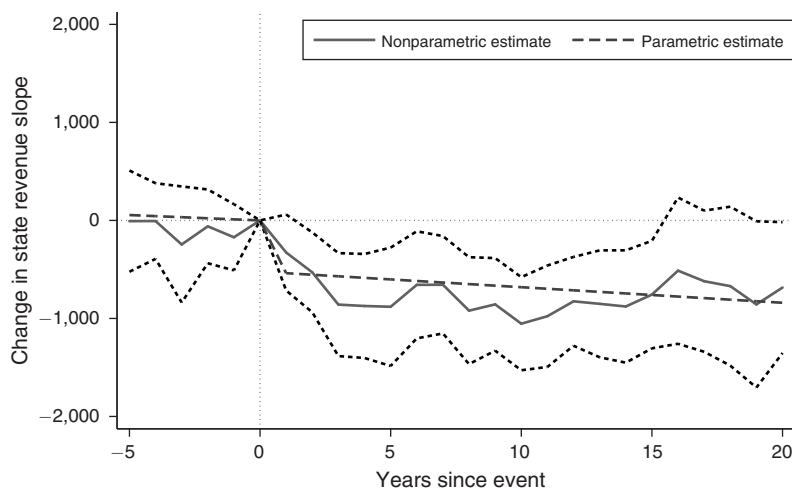
FIGURE 5. EVENT STUDY ESTIMATES OF EFFECTS OF SCHOOL FINANCE REFORMS ON PROGRESSIVITY OF STATE
REVENUES

Notes: Figure displays coefficients from event study regressions. Dependent variable is the slope of state per pupil revenues (in 2013 dollars) 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 show the nonparametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95 percent confidence intervals. Estimates for the parametric models are reported in Table 3, column 5, panel B. The *p*-value for the omnibus hypothesis tests of zero pre-event effects in the nonparametric 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.

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.[20] Finally, we see large effects on noninstructional expenditures, particularly capital outlays.

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.[21]

---

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.

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

[21] Neilson and Zimmerman (2014) find that school reconstruction causes increases in student achievement. Cellini, Ferrerra, and Rothstein (2010) and Martorell, Stange, and McFarlin (2016) fail to find significant effects, but each study is under-powered to detect effects of plausible magnitude.

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
Panel A. 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)
Panel B. 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	63,321	−30	170	508	−247
		(1,016)	(1,052)	(932)	(1,127)
Pupil teacher ratio	15.50	−0.59	−0.65	0.03	0.20
		(0.19)	(0.19)	(0.20)	(0.17)
Noninstructional 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 in columns 3–5 are those used in Table 3, panels A and C. See notes to Table 3.

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 triple-difference specifications that use the achievement gap between low- and high-income districts as the dependent variable.
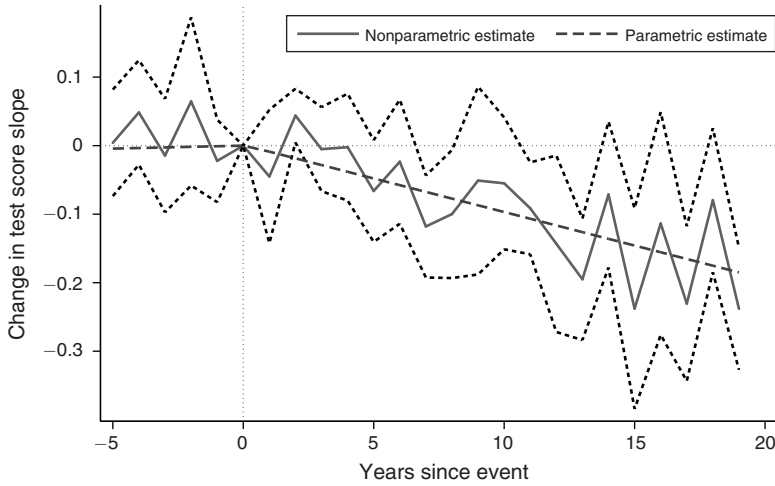
FIGURE 6. EVENT STUDY ESTIMATES OF EFFECTS OF SCHOOL FINANCE REFORMS ON PROGRESSIVITY OF TEST SCORES

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 show the nonparametric model (equation 3) with the event year (indicated as 0) as the excluded category; dotted lines represent 95 percent confidence intervals. Both event study regressions include state and subject-grade-year fixed effects. Estimates for the parametric models are reported in Table 5, column 1. The *p*-value for the omnibus hypothesis test of zero pre-event effects in the nonparametric 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 is 0.02. Standard errors are clustered at the state level.

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 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 nonparametric 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 nonparametric 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)

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 × years 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	1,498	1,498	1,509	1,506	1,504	1,504
p, total event effect = 0	0.02	0.01	0.02	0.69	0.04	0.07
State fixed effects	X	X	X	X	X	X
Subject-grade-year fixed effects	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.

fits the data nearly as well). The downward trend continues through the end of our sample.[22]

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. 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 fourth and eighth graders, respectively, will have attended school solely in the post-event period.[23] This may reflect the use of some additional funds for durable

---

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

[23] We have estimated separate nonparametric models for fourth and eighth grade scores. Both sets of effects grow roughly linearly through the end of our panels. See Lafortune, Rothstein, and Schanzenbach (2016), online Appendix Figure 4.
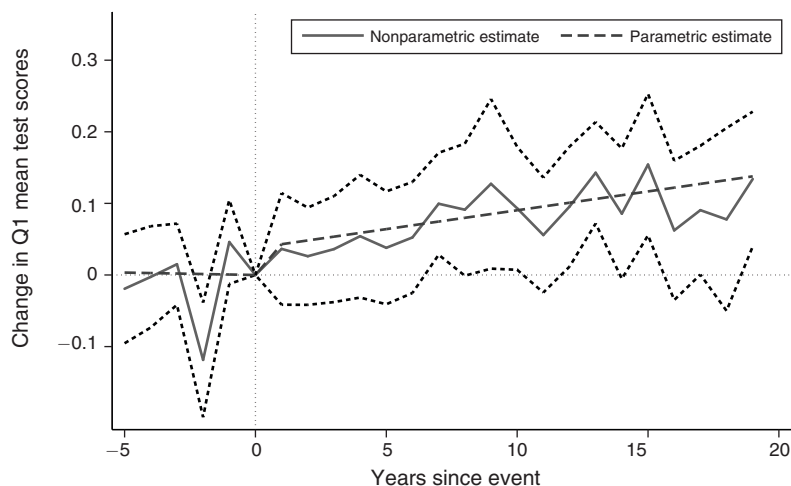
FIGURE 7. EVENT STUDY ESTIMATES OF EFFECTS OF SCHOOL FINANCE REFORMS ON MEAN TEST SCORES IN LOWEST
INCOME SCHOOL DISTRICTS

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 show the nonparametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95 percent 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 nonparametric 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 is 0.01. Standard errors are clustered at the state level.

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.[24] 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 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. Online 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

---

[24] 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.
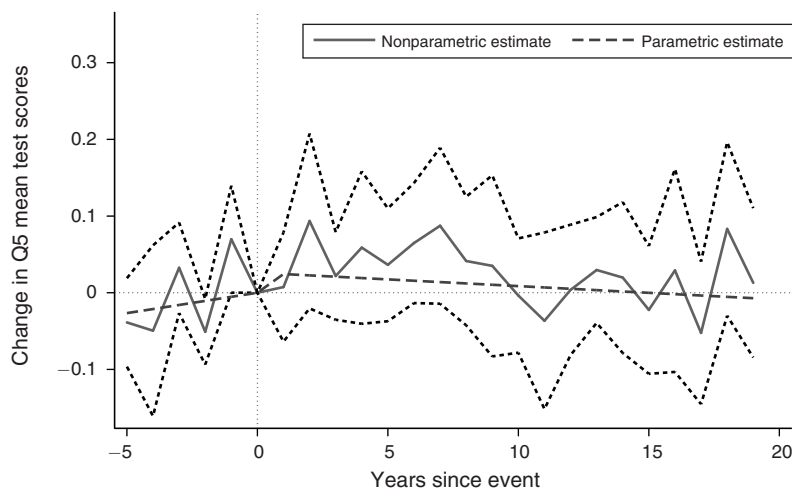
FIGURE 8. EVENT STUDY ESTIMATES OF EFFECTS OF SCHOOL FINANCE REFORMS ON MEAN TEST SCORES IN HIGHEST INCOME SCHOOL DISTRICTS

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 show the nonparametric model (equation 3), with the event year (indicated as 0) as the excluded category; dotted lines represent 95 percent 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 nonparametric 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, is 0.25. 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)
Panel A. 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
Panel B. 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 3, columns 4 and 5, panels A and C, and Table 5, 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.

the idea that funding is more productive in low-income districts, but equal ratios cannot be ruled out.

A. *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 nonwhite or free- or reduced-price lunch eligible share (online 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 online 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 (Card and Krueger 1992b; Krueger and Whitmore 2002). 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 row 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.

Panel A shows impacts on the standard deviation or interquartile range of achievement within states, while panels B and C 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.

TABLE 8—EVENT STUDY ESTIMATES FOR MEAN NAEP SCORES BY SUBGROUP

	Post event $\times$ years elapsed	
Overall mean	0.004	(0.003)
Spread of distribution		
Standard deviation	-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 twenty-fifth and seventy-fifth 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.

Online Appendix Tables A5 and A6 resolve the discrepancy. While nonwhite, 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 nonwhite and low-scoring students, and one-third of low-income students, live in first-quintile districts, while about 10 percent of each live in fifth-quintile districts (online 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 online Appendix Table A6, indicate that finance events raise relative per pupil revenues in the average black student's school district by only $195 (SE 164), decrease relative per pupil revenues in the average low-income student's district by $33 (SE 219), and raise relative per-pupil revenues in the average low-scoring student's district by $193 (SE 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 prekindergarten programs—that may have contributed to the achievement effects, though our impression is that for the typical reform the main change was in funding.[25] 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 eighth 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 10 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.

---

[25] 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 percent 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.[26]

This ratio is not wholly robust. Our quintile analysis shows larger revenue effects, implying a benefit-cost ratio below one, while Jackson, Johnson, and Persico (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 Jackson, Johnson, and Persico (2016). School finance reforms are blunt tools, and some critics (Hoxby 2001; Hanushek 2006) 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, Jushan. 1997. "Estimation of a Change Point in Multiple Regression Models." *Review of Economics and Statistics* 79 (4): 551–63.

Berry, Christopher, and Charles Wysong. 2012. "Making Courts Matter: Politics and the Implementation of State Supreme Court Decisions." *University of Chicago Law Review* 79 (1): 1–29.

---

[26] The earnings effects of increases in eighth 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.

Burtless, Gary. 1996. *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, DC: Brookings Institution Press.

Card, David, and Alan B. Krueger. 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, David, and Alan B. Krueger. 1992b. "School Quality and Black-White Relative Earnings:A Direct Assessment." *Quarterly Journal of Economics* 107 (1): 151–200.

Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics* 123 (1): 177–218.

Card, David, and A. Abigail Payne. 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, Elizabeth U., Nora Gordon, and Sarah Reber. 2013. "Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South." *American Economic Journal: Economic Policy* 5 (3): 126–59.

Cascio, Elizabeth U., and Sarah Reber. 2013. "The Poverty Gap in School Spending Following the Introduction of Title I." *American Economic Review* 103 (3): 423–27.

Cellini, Stephanie R., Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125 (1): 215–61.

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

Clark, Melissa A. 2003. "Education Reform, Redistribution, and Student Achievement: Evidence From the Kentucky Education Reform Act." https://www.mathematica-mpr.com/-/media/publications/pdfs/education/edreform_wp.pdf.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity*. Department of Health, Education, and Welfare. Washington, DC, July.

Coons, John E., William H. Clune, III, and Stephen D. Sugarman. 1970. *Private Wealth and Public Education*. Cambridge, MA: Belknap Press.

Corcoran, Sean P., and William N. Evans. 2015. "Equity, Adequacy, and the Evolving State Role in Education Finance." In *Handbook of Research in Education Finance and Policy*, 2nd ed., edited by Helen F. Ladd and Margaret E. Goertz, 353–75. New York: Routledge.

Downes, Thomas A., and Leanna Stiefel. 2015. "Measuring Equity and Adequacy in School Finance." In *Handbook of Research in Education Finance and Policy*, 2nd ed., edited by Helen F. Ladd and Margaret E. Goertz, 244–59. New York: Routledge.

Duncombe, William, Phuong Nguyen-Hoang, and John Yinger. 2015. "Measurement of Cost Differentials." In *Handbook of Research in Education Finance and Policy*, 2nd ed., edited by Helen F. Ladd and Margaret E. Goertz, 260–78. New York: Routledge.

Dynarski, Susan, Joshua M. Hyman, and Diane W. Schanzenbach. 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, Ann E., and Sheila E. Murray. 2004. "A Decade of Reform: The Impact of School Reform in Kentucky." In *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, edited by John Yinger, 195–213. Cambridge: MIT Press.

Guryan, Jonathan. 2001. "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts." National Bureau of Economic Research (NBER) Working Paper 8269.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24 (3): 1141–77.

Hanushek, Eric A. 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis* 19 (2): 141–64.

Hanushek, Eric A. 2003. "The Failure of Input-Based Schooling Policies." *Economic Journal* 113 (485): F64–98.

Hanushek, Eric A. 2006. "School Resources." In *Handbook of the Economics of Education*, Vol. 2, edited by E. Hanushek and F. Welch, 865–908. Amsterdam: North-Holland.

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

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996a. "Aggregation and the Estimated Effects of School Resources." *Review of Economics and Statistics* 78 (4): 611–27.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996b. "The Identification of School Resource Effects." *Education Economics* 4 (2): 105–25.

Horowitz, Harold W. 1966. "Unseparate but Unequal: The Emerging Fourteenth Amendment Issue in Public School Education." *UCLA Law Review* 13: 1147–72.

Hoxby, Caroline M. 2001. "All School Finance Equalizations Are Not Created Equal." *Quarterly Journal of Economics* 116 (4): 1189–1231.

Hyman, Joshua. 2017. "Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment." *American Economic Journal: Economic Policy* 9 (4): 256–80.

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

Kirp, David L. 1968. "The Poor, the Schools, and Equal Protection." *Harvard Educational Review* 38 (4): 635–68.

Koski, William S., and Jesse Hahnel. 2015. "The Past, Present and Possible Futures of Educational Finance Reform Litigation." In *Handbook of Research in Education Finance and Policy*, 2nd ed., edited by Helen F. Ladd and Margaret E. Goertz, 41–59. New York: Routledge.

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114 (2): 497–532.

Krueger, Alan B. 2003. "Economic Considerations and Class Size." *Economic Journal* 113 (485): F34–63.

Krueger, Alan B., and Diane M. Whitmore. 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." *Economic Journal* 111 (468): 1–28.

Krueger, Alan B., and Diane M. Whitmore. 2002. "Would Smaller Classes Help Close the Black-White Achievement Gap?" In *Bridging the Achievement Gap*, edited by John E. Chubb and Tom Loveless, 14–41. Washington, DC: Brookings Institution Press.

Ladd, Helen F., and Margaret E. Goertz, ed. 2015. *Handbook of Research in Education Finance and Policy.* 2nd ed. New York: Routledge.

Lafortune, Julien, Jesse Rothstein, and Diane W. Schanzenbach. 2016. "School Finance Reform and the Distribution of Student Achievement." National Bureau of Economic Research (NBER) Working Paper 22011.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2018. "School Finance Reform and the Distribution of Student Achievement: Dataset." *American Economic Journal: Applied Economics.* https://doi.org/10.1257/app.20160567.

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

Murray, Sheila E., William N. Evans, and Robert M. Schwab. 1998. "Education-Finance Reform and the Distribution of Education Resources." *American Economic Review* 88 (4): 789–812.

Nielson, Christopher A., and Seth D. Zimmerman. 2014. "The effect of school construction on test scores, school enrollment, and home prices." *Journal of Public Economics* 120: 18–31.

Sands, Melissa. 2015. "From the Statehouse to the Schoolhouse: The Distributive Politics of Education Funding." https://www.dropbox.com/s/rjx7745ttopwmdq/education_sands.pdf?dl=0.

Sims, David P. 2011a. "Lifting All Boats? Finance Litigation, Education Resources, and Student Needs in the Post-'Rose' Era." *Education Finance and Policy* 6 (4): 455–85.

Sims, David P. 2011b. "Suing for Your Supper? Resource Allocation, Teacher Compensation and Finance Lawsuits." *Economics of Education Review* 30 (5): 1034–44.

Wise, Arthur E. 1968. *Rich Schools, Poor Schools: The Promise of Equal Educational Opportunity.* Chicago: University of Chicago Press.