Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1287

Labor supply, learning time, and the efficiency of school spending: Evidence from school finance reforms

John Bodian Klopfer
The University of Hong Kong

Does school spending raise achievement? I show that effects, benchmarked by schools' daily value added, are one-tenth to one-third as large as spending growth. Using school finance reforms for identification, I show that schools did not raise quality measured by value added. Instead, schools raised quantity measured by time diaries of staff and student hours, more than spending, with most hours added after testing. Private time costs of reforms exceed public costs. Achievement effects are small due to fadeout and delayed measurement, suggesting long-run economic effects may be more informative.

VERSION: September 2025

Suggested citation: Bodian Klopfer, John . (2025). Labor supply, learning time, and the efficiency of school spending: Evidence from school finance reforms. (EdWorkingPaper: 25-1287). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/wjkb-k563

Labor supply, learning time, and the efficiency of school spending: evidence from school finance reforms*

John Bodian Klopfer†

September 12, 2025

Abstract

Does school spending raise achievement? I show that effects, benchmarked by schools' daily value added, are one-tenth to one-third as large as spending growth. Using school finance reforms for identification, I show that schools did not raise quality measured by value added. Instead, schools raised quantity measured by time diaries of staff and student hours, more than spending, with most hours added after testing. Private time costs of reforms exceed public costs. Achievement effects are small due to fadeout and delayed measurement, suggesting long-run economic effects may be more informative.

1 Introduction

Does money matter? How much? The evidence on school finance reforms and achievement growth remains disputed (Handel and Hanushek, 2023 and 2024; Jackson and Mackevicius, 2024) because nearly all studies report null or negative estimates alongside positive estimates (Downes, 1992; Downes and Figlio, 1997; Papke, 2000; Guryan, 2001; Card and Payne, 2002;

*I thank John List and Kevin Murphy for two lectures that led to this project, and Will Dobbie, Ilyana Kuziemko, and Christopher Neilson for advice and encouragement. This work was improved by discussions with Mark Aguiar, Henry Farber, Naijia Guo, Bo Honoré, Alan Krueger, David Laibson, Edward Lazear, Alexandre Mas, Philip Oreopoulos, Juan Pantano, Minseon Park, Harvey Rosen, Kenneth Shores, participants at the SOLE/EALE/AASLE 2025 World Labor Conference, the AEFP 2019 Annual Conference, the MPRC 2018 Time Use Conference, the APPAM 2017 Annual Conference, the NTA 2017 Annual Conference, and seminars at Princeton, Treasury, USNA, Haverford, and HKU. I am indebted to Cecilia Rouse and Eric Hanushek for access to the restricted NAEP; to Julien Lafortune, Jesse Rothstein, and Diane Schanzenbach for sharing their data and code; and to Anne Busacca-Ryan, Brianne Holland-Stergar, and Nora Niedzielski-Eichner for expert legal research on school finance reforms. This material is based upon work supported by the IR Section at Princeton, and by the NSF under Grant No. DCE-1148900. Any opinions, findings, and conclusions or recommendations expressed in this material are mine alone, and do not necessarily reflect the views of my colleagues or the NSF. *Truth is the daughter of time, not of authority.*—Sir Francis Bacon

†The University of Hong Kong. Correspondence: johnklopfer@gmail.com.

Clark, 2003; Chaudhary, 2009; Roy, 2011; Lafortune, Rothstein, and Schanzenbach, 2018; Brunner, Hyman, and Ju, 2020; Buerger, Lee, and Singleton, 2021; Dinerstein and Smith, 2021; Baron, Hyman, and Vasquez, 2024).[1] This pattern may owe to heterogeneous effects, as suggested by Handel and Hanushek (2023 and 2024); or it may owe to low power to detect true effects, as suggested by Jackson and Mackevicius (2024). Thus, whether money matters for achievement depends on two questions: Are effects small? And if so, why?

I begin by assessing achievement effects found in other studies of school finance reforms. Effect sizes are uncertain because achievement tests have no ratio scale, meaning changes cannot be compared to levels (Stevens, 1946). However, changes can be compared to changes: Klopfer (2025) observes that achievement effects in the National Assessment of Educational Progress (NAEP) can be benchmarked to schools' daily value added, measured using quasi-random test timing within the NAEP testing window.[2] Dividing achievement effects by the baseline rate of achievement growth gives them a ratio interpretation, in units of time. Dividing spending effects by the baseline rate of spending does the same. Dividing the achievement ratio by the spending ratio yields the elasticity of achievement to spending, a unitless measure of efficiency relative to baseline production.

I find that achievement effects are small. Benchmarking effects in studies using the NAEP (Clark, 2003; Lafortune, Rothstein, and Schanzenbach, 2018; Brunner, Hyman, and Ju, 2020; Buerger, Lee, and Singleton, 2021), in Section 5.1, I find that elasticities of achievement to spending typically range from one tenth to one third, and include negative estimates, across both advantaged and disadvantaged populations. I cannot reject an elasticity of zero for any estimate, and more often than not, reject unit elasticity.

These small achievement effects deepen a central puzzle in the school finance literature: if inframarginal spending produces achievement, then why doesn't marginal spending from school finance reforms produce proportional achievement? School operations are a black box, and there are many steps in the causal chain, inviting a range of hypotheses that have yet to be tested. School staff may not work harder (Sims, 2011b; Brunner, Hyman, and Ju, 2020).

[1]The focus of this paper is on achievement testing because it is the most common outcome, and is often used to predict and validate long-run outcomes. Other causal studies of school finance reforms have studied long-run outcomes including educational attainment (Hoxby, 2001; Hyman, 2017; Candelaria and Shores, 2019; Rothstein and Schanzenbach, 2022; Card et al., 2024), earnings (Jackson, Johnson, and Persico, 2016; Rothstein and Schanzenbach, 2022; Biasi, 2023; Card et al., 2024), crime (Baron, Hyman, and Vasquez, 2024), and civic participation (Asker, Brunner, and Ross, 2024).

[2]Cunha and Heckman (2007) criticize the school finance literature for failing to benchmark. Jackson and Mackevicius (2021) provisionally benchmark the literature in units of class size and teacher value added, but do not account for test scales, precision, or the comparability of interventions. Daily value added in the NAEP has advantages as a benchmark: it is measured per unit of time like spending, has validated conversions to state test scales (Reardon, Kalogrides, and Ho, 2021), is precise ($F = 64$), and measures average productivity from ongoing operations rather than marginal productivity from a specific intervention.

Parents may invest less as public schools invest more (Becker and Tomes, 1986; Todd and Wolpin, 2003; Glewwe and Kremer, 2006). Children may fail to match schools' investment (De Fraja, Oliveira, and Zanchi, 2010; Todd and Wolpin, 2018). Finally, measurement, both test design and study design, may not be suited to capture achievement effects.

In the remainder of this paper, I evaluate each step in the causal chain from spending to achievement. I quantify inputs to the education production function and their responses to school finance reforms, showing that there is no obvious broken link from spending to inputs, but that private costs may be much greater than previously appreciated. I then quantify probable achievement effects based on the input evidence, and probable measurement biases based on input timing, test timing, and common specification choices.

To identify the impact of reforms on funding, inputs, and achievement, I implement an event-study design identified by the timing of reforms in different states. The timing of reforms is arguably exogenous to other factors that could affect school, student, or parent inputs. Following the literature, I study two kinds of reforms: voluntary, and court-ordered. In court-ordered reforms, one plaintiff sets precedent for an entire state, the median case is filed nearly a decade before it results in a reform (separating the conditions at filing from those at reform), and courts necessarily override the intentions of the legislature. Voluntary reforms are determined by the marginal legislator, who in turn is elected according to a bundle of issues that might or might not include school finance. The exogeneity assumption passes falsification tests: the U.S. Department of Education's (ED) Common Core of Data (CCD), School District Finance Survey (SDFS) shows no evidence of pre-trends or inexplicable post-trends in enrollment or state, local, or federal funding.

To measure school quality, I use the NAEP and the Current Population Survey (CPS). I use the NAEP to estimate the cumulative achievement effects of reform exposures. Novel to this paper, I define school quality effects as changes in schools' daily value added, and interact the NAEP's quasi-random test timing (also used here to measure average daily value added for benchmarking) with school finance reform exposures, to estimate the causal effect of reforms on schools' daily value added. I use the CPS to measure the employment and qualifications of school staff, independent of employer reporting in ED's CCD.

To measure school quantity and household inputs, I use the American Time Use Survey (ATUS). Most inputs to education production involve time. Time diaries have not been used to study school finance reforms before, and make it possible to measure school staff labor supply, student learning time, parents' related time inputs, and crowd-out of other activities. Measuring school staff hours is crucial: validation studies show that time diaries are more accurate than stylized time use questions about working hours, as in the CPS, and I show that time diaries are more sensitive to changes in working hours. Measuring responses by children and parents is another new contribution to the literature.

I find that schools use reform funding to raise quantity, not quality. On average, each reform raises total spending by 4.3 percent, with the following effects: (1) *Achievement does not grow.* Cumulative gains on the NAEP are consistent with zero to small positive effects. School quality, measured by value added during the NAEP testing window, does not change. (2) *Schools pay similar workers more.* In the CPS sample, the elasticity of employment for workers in the K-12 industry ranges from -0.2 to 0.3. The average worker does not have more experience. However, school staff are paid 2 percent more. (3) *School staff and students spend more time at work and school.* K-12 staff report 116 more hours of work per year (6.8 percent), and children eligible for K-12 and aged 15-19 report 68 more hours of education activities per year (5.5 percent), per reform. Students spend more time in school after testing, in May, June, and July.[3] (4) *Time at school crowds out other activities.* Increased time in education activities is offset by reduced time in market and household production activities, but not in leisure or non-school investments in human capital. Parents of children eligible for K-12 and aged 5-19 do not allocate time differently after reforms.

Schools raise quantity because it is cost efficient (for schools) to raise quantity. Raising *quality* typically means hiring more or better staff, or incentivizing staff to exert more effort. Raising quality may or may not be cost efficient: see Hanushek, Kain, and Rivkin (2004), Rothstein (2015), and Biasi (2021) for ability; and Fryer (2013), Goodman and Turner (2013), and Bates and Johnston (2024) for effort. Raising *quantity* typically means adding hours for existing staff. The cost of added hours is so far unmeasured in the school finance literature, so we know little about the cost efficiency of raising quantity. Because of fixed costs in employment, and because the work schedules of school staff are constrained, marginal costs for hours may be less than average costs. I find that the marginal cost of added hours is less than the average cost: the elasticities of staff working hours to compensation expenditure and to weekly earnings are both greater than one. That is, hours grow more than expenditures. My causal estimates confirm observational estimates of the marginal rate of substitution between hours and earnings, for the broad labor market, in Lachowska et al. (2023).

Raising quantity imposes private costs on families. The cost of an hour put into school is the value of an hour taken out of another activity. Economists have taught at least since Mincer (1958), Schultz (1960), Becker (1965), and Ben-Porath (1967) that the time

[3] Surprisingly, studies of school finance reforms do not address learning time. Extended learning time is a major topic in federal, state, and local education policy, and a core feature of charter schools. It is often required by court orders and legislation: the plaintiffs in Butt v. State of California, 842 P.2d 1240 (Cal. 1992) sued to prevent their district from shortening the school year to address budget deficits, and the court forced the state to fund a longer school year; implementing a court order, the New York Education Budget and Reform Act of 2007 required additional spending on inputs including extended learning time.

costs of schooling are of first-order importance, yet time costs are entirely ignored in the literature on school finance. I find that private time costs outweigh public costs by a factor of four. This finding is new and important for the school finance debate, and also carries general lessons for program evaluation. Contrary to effects hypothesized in Becker and Tomes (1986), Todd and Wolpin (2003), and Glewwe and Kremer (2006), I find that parents are neither highly responsive, nor heavily involved in education-related childcare, which is less evident in the program evaluation literature relying on qualitative measures (Bonnesrønning, 2004; Houtenville and Conway, 2008; Datar and Mason, 2008; Pop-Eleches and Urquiola, 2013; Das et al., 2013; Gelber and Isen, 2013; Fredriksson, Öckert, and Oosterbeek, 2016; Kline and Walters, 2016; Araujo et al., 2016; Fu and Mehta, 2018; Attanasio et al., 2020; Bergman, 2021; and Greaves et al., 2023). Evaluations could give more attention to students, and measure a broader range of activities for parents and students.

Raising quantity could raise achievement. However, measurement is a serious challenge: inputs arrive after testing, so effects are measurable only at a one-year lag or cumulatively. Estimates for one-year fadeout range from 50 to 80 percent, and estimates for four-year fadeout are as high as 90 percent and not distinct from zero (Jacob, Lefgren, and Sims, 2010; Chetty et al., 2011; Chetty, Friedman, and Rockoff, 2014; Gilraine and Pope, 2021), so lagged or cumulative estimates may not be distinct from zero even when immediate estimates are. My cumulative estimate of the achievement effect of reforms is 0.000σ per year (95 percent CI: -0.004 to 0.005), consistent with estimates in Table 3 and with total fadeout. I construct a model-based estimate of the achievement effect of reforms, based on my estimates of daily value added in the NAEP and of student hours responses in the ATUS: gains on the NAEP of at least 0.51σ per year and a 5.5 percent increase in learning time (90 percent CI: 0.010 to 0.100) imply gains on the NAEP of 0.028σ per year of exposure (90 percent CI: 0.005 to 0.051), consistent with quasi-experimental and experimental evidence on learning time interventions. Immediate effects may be larger than annualized cumulative effects because of delayed measurement and fadeout. Evaluations relying on achievement tests should give more attention to input timing, test timing, and fadeout.

In summary, the use of reform funding to extend the school year is consistent with normal human capital production, and suggests that there could be proportionate effects on outcomes. Despite the low marginal cost of school staff hours supplied to schools, private costs are far larger than public costs and largely unacknowledged in research and education policy. Finally, achievement effects, if there are any, may be near zero due to the timing of inputs relative to measurement: evaluators should exercise caution when relying on achievement tests, and should place greater weight on long-run economic outcomes.

Contributions and Related Literature My paper contributes to the literature on school finance reforms, and also on school funding. Many other studies rely on achievement tests and face challenges reconciling positive effects with null or negative effects (van der Klaauw, 2008; Weinstein et al., 2009; Cellini, Ferreira, and Rothstein, 2010; Matsudaira, Hosek, and Walsh, 2012; Neilson and Zimmerman, 2014; Goncalves, 2015; Hong and Zimmer, 2016; Martorell, Stange, and McFarlin, 2016; Conlin and Thompson, 2017; Kogan, Lavertu, and Peskowitz, 2017; Carlson and Lavertu, 2018; Gigliotti and Sorensen, 2018; Kreisman and Steinberg, 2019; Shores and Steinberg, 2019; Abott et al., 2020; Jackson, Wigger, and Xiong, 2021; Goldstein and McGee, 2021; Baron, 2022; Brunner, Hoen, and Hyman, 2022; Lafortune and Schönholzer, 2022; Miller, 2022; and Biasi, Lafortune, and Schönholzer, 2025). My paper identifies the importance of input timing, test timing, and fadeout for estimates of cumulative effects. My findings may help to interpret and address heterogeneous, small achievement effects across studies (Handel and Hanushek, 2023 and 2024; Jackson and Mackevicius, 2024), and to reorient the school spending debate toward studies of long-run outcomes.

Second, my paper shows that the literature on school funding and the literature on learning time interventions are connected, through schools' choice of inputs when additional funds become available. There are many studies of learning time as a school input, but I am not aware of another study showing that marginal funds are allocated to learning time. Lavy (2020) is related, however: he says his paper is "not about what happens when schools are given more money and are allowed to spend it as they will," but the reform he studies changes schools' in-kind instructional budgets measured in staff hours, and schools allocate resources to learning time rather than class size reduction, training, support staff, or extracurriculars. I find convergent evidence: when schools *are* given more money, they allocate it to learning time. In this respect, both of our studies show that given more money, schools voluntarily follow the best practices in Dobbie and Fryer (2013) and Fryer (2014).

Because schools use marginal funds to extend learning time, the achievement effects of learning time predict the achievement effects of school funding. For example, Lavy (2020) measures some inputs before testing,[4] and mostly avoids fadeout: he compares achievement in schools in 2002-2003 before learning time was extended, and 2004-2005 after learning time was extended, so for achievement measured in 2004 there is no lagged exposure, and for achievement measured in 2005 there is only one lagged exposure. He estimates an effect of 0.04σ per 22-percent increase in weekly time, or about 0.01σ per 5.5 percent increase in weekly time, one-third of my benchmark; the shortfall could be attributed to early testing

[4] Angrist et al. (2019) report that the GEMS test also used in Lavy (2020) was administered from October to November in 2004-2006. Nonetheless, because schools added hours to each day instead of adding days to the end of the school year, up to one-third of inputs arrive before testing.

and fadeout, but to a less-extreme degree than usual in the school finance literature.

While my benchmark of daily value added in the NAEP (Klopfer, 2025) does not suffer from input timing and fadeout and is aligned with the tests and population in most studies of school finance reforms in the U.S., it measures inframarginal effects, as do other studies of random or quasi-random test timing (Sims, 2008; Fitzpatrick, Grissmer, and Hastedt, 2011; Hansen, 2011; Agüero and Beleche, 2013; Carlsson et al., 2015; Aucejo and Romano, 2016).

Many interventions show that the marginal achievement effects of extended learning time are comparable to my inframarginal benchmark; this suggests that output scales efficiently with input when measured correctly. Directly related to my finding that schools add learning time at the end of the year, several studies of extended school years (Frazier and Morrison, 1998; Sanz and Tena, 2023)[5] and summer school (Matsudaira, 2008; Pyne, Messner, and Dee, 2023) take care to test students after treatment, and find larger estimates, possibly due to test normalization within narrower populations. Other interventions add time evenly throughout the year: Bellei (2009) studies an extended-day program, and finds effects over two years of exposure about one-third of my benchmark; Kawaguchi (2016) and Thompson (2020) study school-week reductions, and Kawaguchi finds effects over two years of exposure about half of my benchmark, while Thompson finds effects controlling for lagged test scores that approach my benchmark, likely due to the lesser influence of fadeout. Subject-specific interventions involve same-year measurement: Taylor (2014), Cortes and Goodman (2014), Cortes, Goodman, and Nomi (2015) study double-dose math; and Andersen, Humlum, and Nandrup (2016) and Figlio, Holden, and Ozek (2018) study double-dose language. Testing after inputs, all find effects on subject-specific achievement that are comparable to the benchmark when scaling for intervention size. In the learning time literature, as in the school finance literature, input timing and fadeout are first-order important.

Third, my paper contributes to the literature on time use, for parents and for children. Time use has not been studied in the school finance literature, and is surprisingly little-studied in the learning time literature: Kawaguchi (2016) and Lavy (2020) study students' time on homework. Kawaguchi (2016), unique even in the time-use literature, goes further to show that time in school crowds out leisure.[6] I study a broader school finance policy that is not explicitly linked to students' schedules, and given that the student population is older, am able to show crowd-out of market and household production. I also quantify causal effects on time use by parents, in that sense linking the causal-but-qualitative program evaluation

[5] Pishke (2007), Parinduri (2014), and Fischer et al. (2020) also study quasi-experimental variation in term length, and look at employment and earnings: possibly due to system-wide compensation, Pischke finds no effect, while Parinduri and Fischer et al. find substantial effects.

[6] Conversely, Stinebrickner and Stinebrickner (2008) and Barwick, Chen, Fu, and Li (2024) show that students' leisure crowds out time in school.

literature following Todd and Wolpin (2003), with the quantitative-but-descriptive work of Guryan, Hurst, and Kearney (2008), Price (2008), and Ramey and Ramey (2010). I confirm qualitative observations that parents are not highly active or responsive in education-related childcare (Bergman, 2021), though parents are active in other kinds of childcare.

It is obvious that school staff time and students' time are complements in production, and others have observed that they are strategic complements (De Fraja, Oliveira, and Zanchi, 2010; Todd and Wolpin, 2018); I provide causal evidence that this is so, and while the sign of the effect may not be surprising, quantitative estimates are badly needed in order to show that weakly responsive children are not a bottleneck in production. My estimates also show that children are more responsive to schools' in-kind inputs than to cash incentives (Angrist and Lavy, 2009; Fryer, 2011), reinforcing efficiency arguments for spending on school staff.

2 Background: school finance reforms and the school spending debate

Every state constitution requires that the state legislature provide for publicly funded education (Parker, 2016), and requires equity in public education. For example, in New Jersey,

> "The Legislature shall provide for the maintenance and support of a thorough and efficient system of free public schools for the instruction of all the children in the State between the ages of five and eighteen years," (Art. VIII, Sec. IV, Par. 1 of the New Jersey State Constitution).

State legislatures provide half of funding for public school districts. Thus, state law is a major driver of public education outcomes.

State funding is vigorously debated among school districts, state legislatures, governors, and courts, touching on at least three core issues: first, taxation, including the allocation of tax burdens and grants; second, the local control of school districts, qualified by accountability measures tied to state funding and by incentives for or against local tax effort; third, the separation of powers between the legislature, the executive, and the judiciary.

The effect of spending is also debated, and is pervasive at almost every level. Cause and effect are not just an academic issue, and not just a policy issue, but also a threshold issue for the courts: to bring a case, plaintiffs must prove that they are injured by the defendant, and that the courts can remedy the injury. See for example *Morath v. The Texas Taxpayer and Student Fairness Coalition*, 490 S.W.3d 826, 847 (Tex. 2016).

In the past half-century, all states faced serious challenges to their school finance laws, and most passed school finance reforms.

Many of these reforms were voluntary, carried out through the ordinary political process. Many others were driven by litigation. No state avoided litigation: not even Hawaii, with a statewide school district. School districts and advocacy groups sue on the basis of two clauses found in every state constitution: equal protection clauses (in “equity” cases, where the state allegedly discriminates among districts in taxation or in the provision of public education, on the basis of property wealth or taxable income) and education clauses (in “adequacy” cases, where the state allegedly fails to provide public education up to the required standards). While some state courts rule that school funding can only be decided by the legislature, many overturn existing school finance laws and compel new legislation. In addition, state legislatures often settle cases, or pre-empt them, with legislation.

Economic research plays an important role in informing the school spending debate, legislation, and litigation. This paper's contribution is to show in detail how school districts spend funds on inputs after school finance reforms, how the chosen inputs might explain conflicting findings in the existing literature on school spending and achievement, and what researchers, school districts, legislatures, and courts can do to get more reliable evidence.

3 Empirical strategy

To identify the effects of school finance reforms on funding, spending, inputs, and achievement, I rely on event studies.

I select reform events, in the period from 1990 through 2013, from the listings in Jackson, Johnson, and Persico (2016) and Lafortune, Rothstein, and Schanzenbach (2018). I check that court orders are final, and confirm event timing. To ensure that reforms raised statewide funding, I check events for visual consistency with the time series of school district revenues. When there are multiple events in a single year, I code one event.

From 1990 through 2013, the combined list includes 86 reform events in 28 states, covering 59 percent of public school enrollment in the United States; these are the reforms used to confirm identification. In later sections, I use the subset of reforms from 2003 onward, which includes 29 reform events in 19 states, covering 48 percent of enrollment.

I measure funding and spending with the SDFS (NCES, 1990-2013). The SDFS collects fiscal variables and enrollment for the universe of public school districts in the United States, every fiscal year, based on their audited financial statements. To compare public school enrollment to private school enrollment, I complement the SDFS with ED's Private School Survey (PSS) (NCES, 1990-2012) records of enrollment for the universe of private schools.

3.1 Static event-study specification

This paper estimates the impacts of school finance reforms by looking at the change in funding in states that enact reforms, in the years subsequent to the reform, relative to states that did not enact reforms in the same years.

The static event-study specification is:

$$Y_{it} = \sum_{r \in R_s} \alpha \mathbb{I}_{trs} + A_s + B_t + \gamma X_{it} + \epsilon_{it} \tag{1}$$

where Y_{it} is a variable measured at the individual (or state, district, or household) level, A_s and B_t are state (or district) and year fixed effects, and X_{it} is a vector of controls. $\mathbb{I}_{trs}$ is an indicator that time t is after reform r in state s, and α is the average effect of a reform. The indicator is defined by $\mathbb{I}_{trs} = 1[t > t_{sr}]$, and is equal to zero in all periods in states that have not had a reform in the period from 1990 to present. Errors ϵ_{it} are clustered at the state level to account for state-level variation not captured in the model.[7]

I present state-level estimates for two reasons: first, reforms are state-level policies; and second, the key samples from the American Time Use Survey (ATUS) do not report geography at the county or school-district level.

I model all reforms at once (with the summation term) to avoid attributing the effects of omitted reforms to modeled reforms, to trends, or to error terms. For example, New York enacted several major reforms in quick succession: if I were to include one of these reforms, but not the others, the coefficient for that reform would be overstated because it would pick up the effects of other reforms. In addition, in an event study, failing to include earlier or later reforms would give rise to spurious pre- and post-trends. Including multiple reforms addresses a known source of misspecification error in previous studies, when reforms are not timed independently of each other within states.[8]

[7] In later sections, I present triple-difference specifications comparing the effects on a population of interest, such as K-12 staff, with the effects on a control population, such as other professionals:

$$Y_{it} = \sum_{r \in R_s} (\alpha g_i + \beta) \times \mathbb{I}_{trs} + D_{gs} + E_{gt} + (\gamma g_i + \delta) \times X_{it} + \epsilon_{it} \tag{2}$$

where $g_i = 1$ if the observation is in the group of interest and 0 if the observation is in the control group: β captures the baseline effect of a reform or of any confound associated with the reform, state and year fixed effects A_s and B_t are replaced with group-by-state and group-by-year fixed effects D_{gs} and E_{gt}, and control variables are also interacted with group membership, with the coefficient δ giving the baseline relationship to control variables for both groups.

[8] Including multiple reforms improves efficiency. It also addresses concerns about treatment-effect heterogeneity due to cumulative reforms, without relying on heterogeneity-robust estimators that are inefficient and may perform poorly in small samples (de Chaisemartin and D'Haultfouille, 2025, section 6.3): the Goodman-Bacon (2021) decomposition suggests that negative weights are not a major concern in this specification.

The identification assumption is that:

$$E[\epsilon_{it}| \sum_{r\in R_s} \mathbb{I}_{trs}, A_s, B_t, X_{it}] = 0$$

or in words, that reform timing is exogenous conditional on fixed effects, controls, and other reforms within the state.

There are strong procedural reasons to think reform timing is unrelated or weakly related to other determinants of funding, inputs, and outcomes. Reforms are large, abrupt events caused by small shifts in fundamentals. In legislation, marginal voters decide elections, voters vote for parties and candidates, not laws, and legislators bargain over unrelated issues. In litigation, case outcomes are unpredictable (Corcoran and Evans, 2007). Courts are bound by precedent, and reluctant to set precedent based only on the circumstances of an individual case. Precedent set by one plaintiff affects all school districts. Further, court orders are only loosely related to contemporary fiscal conditions: cases must be organized, funded, filed in trial courts, and fought, through years or decades of appeals to higher courts and remands to lower courts, until the state supreme court orders new legislation and compels the legislature and governor to comply. Court orders also override voter preferences, if those preferences were expressed in overturned school finance laws.

The assumption of exogenous timing is also falsifiable in district-level school finance and enrollment data. I run specification checks and falsification tests using a dynamic event-study estimator similar to the preferred static event-study estimator in Equation (1).

3.2 Dynamic event-study specification and falsification tests

The dynamic event-study specification is:

$$Y_{dt} = \sum_{r\in R_s} \sum_{k\in K} \alpha_k \mathbb{I}_{tkrs} + A_d + B_t + \gamma X_{dt} + \epsilon_{dt} \tag{3}$$

where Y_{dt} is a fiscal variable measured at the district (or county) level, A_d and B_t are district and year fixed effects, and X_{dt} is a vector of time-varying controls that may contain state time trends. $\mathbb{I}_{tkrs}$ is an indicator that time t is k periods after (or before, for negative values of k) reform r in state s, and α_k is the average effect, across all reforms, of a reform at lag k. The coefficient α_k restricts each reform, including multiple reforms within states, to have the same effect at a given time horizon before or after a reform event: this assumption reduces the scope for overfitting. The reform indicator is defined by $\mathbb{I}_{tkrs} = 1[t - t_{sr} = k]$;[9] I bin

[9] School finance data are indexed by the school fiscal year t, which in most states ends on June 30th of the coded year; reforms are indexed by the year t_{sr} in which a court decision was entered, or in which legis-

lags k at -5 and 5, so that $\mathbb{I}_{t-5rs} = 1[t - t_{sr} \leq -5]$ and $\mathbb{I}_{t5rs} = 1[t - t_{sr} \geq 5]$.[10] The reform indicator is equal to zero in all periods in states that have not had a reform in the period from 1990 to present. Finally, errors ϵ_{dt} are clustered at the state level.

Reforms have a significant effect on state revenues to school districts, with no evidence of pre-trends, or of post-trends more than four years after enactment (Figure 1). The absence of pre-trends suggests that modeling all reforms at once reduced specification error, and that the dates of reforms in the listing are coded correctly: measurement error in reform dates would introduce pre- and post- trends. The parallel trends assumption is not rejected.

Falsification tests fail to reject exogeneity with respect to fiscal conditions or the school population. Reforms are not related to past or future local revenue (Figure A1), as we would expect if the state acted to remedy local fiscal conditions, or if districts took advantage of state funding to reduce local tax effort (Hoxby, 2001). Reforms are not related to public or private enrollment before or after reforms (Figure A2 and A3), limiting the scope for selection into the state or the public school system; these results should not be surprising, given low baseline rates of interstate migration and school-switching once children are enrolled.

These findings address the primary concern for identification, that changes in unobservables might jointly determine reforms, funding, inputs, and outcomes.

3.3 Reforms, school revenue, and school spending

Table 1 reports the effects of reforms on fiscal variables from 2003 through 2013, matching the NAEP and ATUS. Reforms raised per-pupil spending by 611 dollars (2015 current) on average, or 4.3 percent; 338 dollars went to compensation, or 3.6 percent.

4 Data

To measure school quality, school quantity, and private inputs, I introduce new data sources and measures to the literature on school finance.

I show that the NAEP can be used to benchmark achievement effects against business-as-usual learning, making achievement effects quantitatively interpretable. I also show that the NAEP can be used to measure changes in daily value added, a new measure of quality.

lation was enacted. The timing of impacts, k, involves delays between court orders, enforcement, legislative enactment and appropriations, and the onset of a new fiscal year. Court injunctions or emergency legislation could affect fiscal variables when $k = 0$, but most reforms should not have effects until $k = 1$, or even $k = 2$. Some reforms, phasing funding in, may have additional effects at later lags.

[10] Binning lags ensures identification in settings, such as that of school finance reforms, where few units are untreated (Schmidheiny and Seigloch, 2023). Binning assumes that treatment effects are constant across binned lags (de Chaisemartin and D'Haultfouille, 2025, section 6.2.2).

Next, to open the black box of school, child, and parent inputs in education production, I rely on labor-force surveys that interview workers, children, and parents directly. The CPS reports employment and pay independently of employers, and the ATUS provides quantitative measures of work, schooling, and other activities, with two unique advantages: first, reports respond to policy changes; and second, they reveal the month of inputs.

4.1 The National Assessment of Educational Progress (NAEP)

In alternate years, the NAEP (NCES, 2003-2013) assesses a random sample of roughly 3000 students in roughly 50 schools in each state, in 4th and 8th grade math and reading.

The NAEP can be used to construct a direct measure of school quality, daily value added (Klopfer, 2025). The NAEP testing window spans January, February, and March: students are tested from the fourth week of the new year through the eleventh week. When test takers take the NAEP one day, week, or month later, their achievement grows, revealing the rate at which test takers learn tested content.[11] Because the learning rate (schools' daily value added) is measured for a fixed interval, it measures school quality independently of quantity (the length of the school year).

I work with a sample of 3,364,010 test takers in 80,310 school-by-grade-by-subject cells, in alternate years from 2003 through 2013. Because schools are frequently resampled across years, I am also able to construct an imbalanced panel of 2,336,500 test takers in 33,950 school-by-grade-by-subject cells, in which each cell is observed at least twice.

Exact testing dates are recorded starting in 2005, so the sample used to measure the learning rate runs from 2005 through 2013.

Daily value added is measured by the coefficient of test scores on the scaled testing date, and is given by β in the following specification:

$$A_{igst} = \alpha + \beta T_{gst} + \epsilon_{igst} \tag{4}$$

where A_{igst} is achievement for student i in grade g and school s in year t, and T_{gst} is the testing date. The effect of reforms on daily value added is estimated by fully interacting T_{gst} with the reform indicator in Equation (1).

Because test instruments vary across years, and because raw scores vary across subjects and grade levels, I standardize test scores within each year-by-grade-by-subject cell: first, I average across the "plausible values" reported by the NAEP for each individual to construct a

[11] Researchers have measured learning rates on a variety of other test instruments, for other populations and in different contexts: Sims (2008), Fitzpatrick, Grissmer, and Hastedt (2011), Hansen (2011), Agüero and Beleche (2013), Carlsson et al. (2015), and Aucejo and Romano (2016).

test score; then, I standardize by the national mean and standard deviation. Thus, estimates of the learning rate are in units of within-cell, national standard deviations.

The identification assumption is that:

$$E[\epsilon_{igst}|T_{gst}] = 0$$

or in other words that test timing is unconditionally exogenous. I condition on school and year fixed effects for precision.

Klopfer (2025) validates the identification assumption, showing that test timing is as good as random due to the test contractor's scheduling constraints, and confirming linearity of value added across weeks of the NAEP.

4.2 The Current Population Survey (CPS)

The CPS (Census, 1990-2015 via Flood et al., 2015) provides worker-reported data on employment in the K-12 education industry, and in every other industry.

I work with a sample of 42,095,421 respondents in the monthly CPS from 1990 through 2015 (school fiscal years ending in 1990 through 2016): among these respondents, 48 percent (20,097,241) are employed, and among the employed, 6 percent (1,181,249) are employed in the K-12 education industry. I work primarily with observations from 2003 onwards, but use the full sample for comparison with other results in the literature based on the CCD.

The CPS has advantages relative to the CCD, which is the only other large, national, regularly reported dataset on employment for school staff. First, the CPS is reported by workers rather than by employers, reducing error at the school or district level and providing an independent check on results based on the employer-reported CCD. Definitions in the CPS (including industry, occupation, and employment) are consistent across years and employers. Second, unlike the CCD, the CPS measures experience.

Unfortunately, the CPS proxies or imputes two-thirds of hours responses. I complement the CPS with the ATUS to measure labor supply on the hours margin.

4.3 The American Time Use Survey (ATUS)

The ATUS (BLS, 2003-2015 via Hofferth, Flood, and Sobek, 2015) measures time allocation using time diary methods. Time diaries are quantitative and complete. Each respondent is asked about his or her activities over 24 hours, starting at 4:00am on the preceding day; in addition to coding the stopping and starting times of detailed activities, the ATUS diary captures the location of activities, and the identities of others who were present.

I work with a subset of the 170,842 time diaries collected in the ATUS from 2003 through 2015 (school fiscal years ending in 2003 through 2016). The ATUS is a stratified random subsample of CPS outgoing rotation groups, and consequently, each respondent reports one diary day, yielding repeated cross-sections. I scale each respondent's report to a full year or month by multiplying daily time allocation by 365.25 or by 365.25/12. Although time allocation could vary across days, the regimented nature of activities like work and schooling ensures that diary days are informative about non-diary days.

Because the ATUS is a representative sample of the United States population, it captures school staff (workers who report employment in the K-12 education industry; N=7,267), high school children (aged 15 through 19 and eligible for high school: either enrolled in high school or not completed high school; N=9,150), and parents of school children (own household children aged 5 through 19 and eligible for school, and parents themselves are not enrolled in any level of education, or employed or reporting previous employment in the education industry or occupations; N=43,409). ATUS samples are often an order of magnitude larger than those available in other quantitative surveys used in the literature.[12]

The ATUS methodically codes activities related to child investment and crowd-out. For school staff, I focus on work activities. For children, I focus on education activities (class, homework, in-school extra-curriculars, and administrative activities), and related travel. For parents, I focus on time with children; time spent on childcare (activities initiated by or because of children) including homework help, medical care, playing games, talking, and providing supervision; and related travel. For both children and parents, I define market and household production as the aggregate of work and work related activities, household activities (chores including cleaning, food preparation, maintenance, yard work, and animal care), consumer purchases, and related travel. The ATUS Activity Lexicon, 2003-2015 (BLS, 2016a) gives examples. My measures of childcare and household production are similar to those in Guryan, Hurst, and Kearney (2008), differing only for infrequent activities.

Time diaries are the most accurate survey method to quantify time allocation across activities (Juster and Stafford, 1991; Hamermesh, Frazis, and Stewart, 2005). Time diaries benefit from both from their recent recall period, and the specificity of starting and stopping times. Recall is almost identical to direct observation for the same individuals over the same period, and is surprisingly accurate even for short-duration activities like walking, searching for lost items, and telephone calls (Robinson, 1985). Since the ATUS interview takes only 15-20 minutes (BLS, 2016b), and is administered to experienced CPS respondents, almost

[12]The largest samples of children found in the literature (Todd and Wolpin, 2018 and Barwick, Chen, Fu, and Li, 2024) are purpose-built and are 50-65 percent as large as the ATUS; only Kawaguchi (2016) studies a larger sample of children, from the Japan Time Use Survey. I am not aware of any other study using time diaries to measure school staff labor supply or parent input adjustment.

all diaries are completed and judged accurate by interviewers and supervisors.

The ATUS time diaries are highly responsive to real changes in work activities. The CPS is not responsive, to a degree that may shock economists who rely on CPS actual and usual hours reports. Consider the seasonality of work in K-12 education: Figure 2 shows that for workers employed in the K-12 education industry, ATUS diary hours on weekdays decline by 60 percent from October to July. CPS "actual hours of work last week" decline by only 18 percent, and CPS "usual hours of work" do not decline at all.

Measurement error for hours in the CPS and other labor force surveys is especially problematic because it is not classical measurement error. Instead, measurement error is negatively correlated with true deviations from the mean, and will bias the effects of reforms on hours towards zero. Measurement error is even worse in the American Community Survey (ACS), due to unsupervised response (Baum-Snow and Neal, 2009).

These facts emphasize the value of using the ATUS time diaries, in which measurement error is minimized and is plausibly orthogonal to the ground truth, to measure the effects of reforms on labor supply, learning time, and other time allocation.

5 Results

Schools use reform funding to raise quantity, not quality. I begin by presenting evidence that school spending does not raise achievement, and effects in the literature are generally small. Next, I show that schools' daily value added, my measure of school quality, does not change, and that schools do not use reform funding to expand staffing. Then, I present evidence that schools instead extend the school year using reform funding. Finally, I present evidence on the opportunity cost of schooling. These findings are discussed further in Section 6.

5.1 Reforms and achievement

School finance reforms do not raise statewide achievement on the NAEP. In the period from 2003 through 2013, I find null effects for three different independent variables: cumulative reforms, cumulative years the school has been exposed to reforms, and cumulative years students have been exposed to reforms given the number of grades completed (Table 2). The confidence interval for years of exposure excludes effects larger than 0.005σ, or one percent of average annual value added (Table 2, Column 2 and 4).

Null estimates are not unusual in the literature on school finance reforms. My estimates are indistinguishable from the similar estimate in Lafortune, Rothstein, and Schanzenbach (2018), based on state-level variation, of 0.004σ per year (95 percent CI: -0.002 to 0.010); all studies relying on the NAEP report at least one null estimate (Table 3, Column 2).

Null estimates are common because true effects are small. I use schools' daily value added (Klopfer, 2025) to benchmark achievement effects in studies relying on the NAEP (Table 3). Specifically, I calculate the elasticity of achievement to spending: see Appendix A.3 for a formal derivation. The elasticity depends on the percentage change in achievement, calculated by dividing the annual effect of reform exposure on NAEP scores by the average annual NAEP value added. All elasticities from the literature are one third or smaller, some are one tenth or smaller, and most statistical tests reject unit elasticity.

The central puzzle of the school finance literature is why such large increases in spending lead to such small increases in achievement. The remainder of Section 5 investigates the role of inputs in the puzzle, while Section 6 investigates the role of measurement; Appendix A.3 and A.4 provide a formal framework for the investigation.

5.2 Reforms and school quality

Schools do not use resources to improve quality. This can be demonstrated with measures of daily value added and of inputs.

5.2.1 School quality

Schools' daily value added does not rise after reforms. In the period from 2005 through 2013, the effect of reforms on the rate of learning was a precise zero, with a confidence interval on the order of half a percent of the mean rate of learning across schools (Table 2, Column 4).

Together with the fact that reforms do not raise achievement, this novel evidence that reforms do not change daily value added suggests either that inputs were not adjusted, that inputs were adjusted in ways that are not effective, or that inputs were adjusted only outside the testing window: tests are taken in January, February, and March.

5.2.2 School staff employment

Schools do not hire more staff with funding from reforms. In the period from 2003 through 2015, the proportion of employed workers reporting employment as K-12 staff falls by 0.5 percent, per reform (Table 4, Column 1; implied elasticities to total and compensation expenditure of -0.12 and -0.14 respectively). This small negative estimate is highly robust. The only positive, significant estimate, an increase of 1.2 percent (Table A1, Column 8; implied elasticities to total and compensation expenditure of 0.28 and 0.33 respectively) comes from a specification including CPS data and reforms from 1990 onwards. Thus, it appears that earlier reforms had stronger effects on staffing. Regardless, in no case are the changes in employment large enough to match rising expenditures.

Estimates from other data sources (typically the SDFS and SDUS) also imply low employment elasticities: Clark (2003), -0.10; Chaudhary (2009), -0.12 to 0.55; Sims (2011b), 0.40; Jackson, Johnson, and Persico (2016), 0.18 to 0.22; and Buerger, Lee, and Singleton, 0.20 to 0.29. Some are higher: Hyman (2017), 0.77; Lafortune, Rothstein, and Schanzenbach (2018), 0.64; and Brunner, Hyman, and Ju (2020), 0.67 to 1.22.

5.2.3 School staff experience

School staff are not more experienced than before reforms. In the period from 2003 through 2015, the mean age of K-12 staff fell by 0.06 years, per reform (Table 5, Column 1; implied elasticities of experience[13] to total and compensation expenditure of -0.06 and -0.08 respectively). This small, insignificant estimate is robust to the inclusion of controls, sample restrictions, and triple-difference comparison groups (Tables A2 and A3).

5.2.4 School staff earnings

Schools raise salaries after reforms. In the period from 2003 through 2015, the weekly earnings of K-12 staff rise by 2 percent, per reform (Table 6, Column 1; implied elasticities to total and compensation expenditure of 0.48 and 0.56 respectively). This positive, significant estimate is robust to the inclusion of controls, sample restrictions, and triple-difference comparison groups (Tables A4 and A5), with one exception: the smallest earning effect is for the period from 1990 onwards. Thus, it appears that earlier reforms gave more emphasis to salaries and less to staffing, compared to the reforms we study from 2003 through 2013.

Estimates from other data sources also imply large salary elasticities: Chaudhary (2009), -0.11 to 0.80; Sims (2011b), 0.50; Jackson, Johnson, and Persico (2016), 0.41; and Buerger, Lee, and Singleton, 0.30 to 0.57. Some are lower: Hyman (2017), 0.19; Lafortune, Rothstein, and Schanzenbach (2018), -0.07; and Brunner, Hyman, and Ju (2020), 0.04 to 0.28. Larger employment elasticities correspond to smaller salary elasticities. Estimates that imply the smallest elasticities infer salaries from the SDFS and SDUS, dividing total salary payments by employment. This underscores the value of an independent estimate.

As I will argue in the next section, staff are being paid in reforms to increase their labor supply on the intensive margin (that is, salaries aren't just providing rents).

[13] I assume age 23 at labor market entry. In unreported results, I find that school staff are not more likely to have education beyond a bachelor's degree.

5.3 Reforms and school quantity

Instead of raising quality, schools choose to improve quantity (extended learning time, ELT, in the form of a longer school calendar).

I show the effects of reforms on school quantity using two independent samples of respondents from the ATUS: school staff and students. I present findings on aggregate hours from both, and on the timing of added hours across the school calendar from the ATUS student sample.[14] Finding school quantity effects in two independent samples lends confidence beyond nominal levels of statistical significance.

5.3.1 School staff hours

School staff work more hours after reforms. Their labor supply rationalizes salary increases paid out by schools.

School staff report 116 more hours of work per year (6.8 percent more) in the ATUS diaries, per reform (preferred estimate reported in Table 7, Column 1, Panel A). The elasticities of work hours to total spending, compensation spending, and weekly earnings are 1.6, 1.9, and 3.4 respectively.

School staff hours estimates are robust. Estimates range from 82 to 108 hours of work per reform when controls for education, gender, and marital status, for employment and full time status, and for state trends are included; from 124 to 164 hours when either reforms are allowed to have effects a year earlier, or the first, second, and third reforms in each state are entered separately (Table 7). Estimates range from 68 to 163 hours when the sample is altered by winsorizing the sampling weights, restricting responses to the CPS sampling frame that ran from August 2005 through April 2014, adding respondents who were enrolled in school alongside their employment in the K-12 industry, dropping respondents who would not have been included in the sample based on their responses in their last CPS interview, dropping responses in the summer, on weekends, and that were deemed low quality by interviewers (Table A6). In all cases, reform effects on hours are larger than the preferred estimate for high school children over the same period. Finally, effects are robust to triple-difference specifications that use other sectors and industries, or a re-weighted sample of workers with matched propensity scores, as control groups: K-12 staff hours grow more when compared to other government and non-profit workers, somewhat less relative to other workers in the professional services industries and the subset of education and health services workers, and about half as much when compared a group of workers matched and re-weighted for their

[14] The school calendar analysis relies on a consistent sample definition across months: the ATUS school staff sample is not suitable for this analysis because it is defined by employment and industry, which staff do not consistently report over the summer vacation.

estimated propensities to work in the K-12 education industry; the lowest estimate is 48 hours per reform (Table A7, Column 8).

Confirming that the results are not peculiar to ATUS diary reports, estimates based on usual hours reported in the ATUS are attenuated by two thirds due to measurement error, but remain significant (Tables 7, A6, and A7, Panel B).[15]

5.3.2 Student hours

Reforms extend learning time for students, by adding weeks to the end of the school year. The effect of reforms on student education activities is smaller than the effect of reforms on school staff hours, because staff work longer than students do, to prepare lessons, grade assignments, and participate in service and professional development.

Children eligible for K-12 schooling and aged 15-19 report 68 more hours of education activities per year (5.5 percent more) in the ATUS diaries, per reform (preferred estimate reported in Table 8, Column 1, Panel A). The elasticities of education hours to total and compensation expenditure are 1.3 and 1.5 respectively; and the costs are 9 and 5 dollars per hour respectively (combining estimates from Table 8 and Table 1).

More than two thirds of the additional time is spent in classes (Table 8, Panel B); estimates are not significant but are proportionate to those for all education activities. Effects on homework are proportionate and significant (Table 8, Panel C).

Like the estimates for staff, student hours estimates are highly robust. They are similar or larger when controls for employment and full or part time status, the presence and education of parents, and state trends are included, when reforms are allowed to take effect a year earlier, and when the first, second, and third reforms in each state are entered separately (Table 8); and when the sample is altered by winsorizing the sampling weights, restricting responses to the CPS sampling frame that ran from August 2005 through April 2014, and dropping older children, children who do not report being enrolled in high school (who may be on holiday), summers, weekends, and interviews deemed by the interviewer to be low quality (Table A8). Students also report 11 more hours of travel related to education per year, per reform (preferred estimate reported in Table 8, Column 1, Panel D); this is an independent check on the results for education activities.

More than half of student hours are added in May, June, and July (Figure 3), at the end of the school year, with the largest effects in June.[16] Hours are added for class and

[15]Because "usual hours worked" and "actual hours worked last week" are poorly measured in the CPS, reforms did not affect CPS-reported hours. Because the ATUS is a subsample of the CPS, the CPS finding can be attributed to the survey instrument.

[16]Jackson, Johnson, and Persico (2016) provide the only related estimate of school calendar effects, reporting that a 1-log-point increase in spending would extend a 179-day school calendar by 13.56 days. This

homework in their usual proportion (Figures A4 and A5). Travel for school provides further confirmation that time is added at the end of the school year (Figure A6).

Schools may prioritize the end of the school year because of legal constraints: many states and school districts set an earliest date that schools may open. Five of the ten most populous states have a statewide earliest start date (Florida, Michigan, New York, North Carolina, and Texas). Most states do not set a latest date that schools must close. Thus, schools can readily add days at the end of the year, but not at the beginning.

5.4 The extended school year and opportunity costs

Time at school and homework necessarily crowd out other activities. But which? Do students forego leisure, market and household production, or other human capital investment? Do parents compensate for differences in the education their children are given, or for differences in their children's contributions to the household?

5.4.1 Time reallocated by students

Students reduce market and home production activities in exact proportion to the increase in education activities, after reforms. No notable effects were found for other investments, for leisure, or for any other category of activities coded in the ATUS.

Children eligible for K-12 schooling and aged 15-19 report 71 fewer hours of market and household production per year (9 percent fewer) in the ATUS diaries, per reform (preferred estimate reported in Table 8, Column 1, Panel E).

This estimate, like those for education activities, is highly robust. It is similar when including controls for employment and full or part time status, the presence and education of parents, and state trends, when reforms are allowed to take effect a year earlier, and when the first, second, and third reforms in each state are entered separately (Table 8); and when the sample is altered by winsorizing the sampling weights, restricting responses to the CPS sampling frame that ran from August 2005 through April 2014, and dropping older children, children who do not report being enrolled in high school (who may be on holiday), summers, weekends, and low quality interviews (Table A8).

5.4.2 Time reallocated by parents

Parents do not adjust time spent on children to compensate for school finance reforms, in either the long run or the short run. Parents do not reinforce the effects of school finance

implies that, for our average reform raising spending by 4.3 percent, schools would add only half a school day, or a 0.33 percent increase, an order of magnitude smaller than what I find in the ATUS.

reforms, either. Parents, instead, are neutral.

Parents report 8 more hours of time with their own K-12 eligible household children in the ATUS diaries, per year, per reform, within a narrow confidence interval that includes zero (preferred estimate reported in Table 9, Column 1, Panel A). The elasticities of time with children to total and compensation expenditure are 0.12 and 0.14 respectively (combining estimates from Table 9 and Table 1). Similarly indicative of small effects on parental behavior, parents report 6 more hours of childcare, 4 fewer hours of education-related childcare per year, and 1 more hour of related travel, per reform (Table 9, Column 1, Panel B-D). While the implied elasticity of education-related childcare to reforms is large, the absolute change in time allocation is not important for education production: parental contributions of education-related childcare amount to 50 hours per year, or approximately 4 percent of the time their children spend on education annually.

These null results are very robust: out of 64 estimates presented in Tables 9 and A9, Panel A-D, only two reach conventional levels of significance, and the majority are close in sign and magnitude to the preferred estimates in Column 1 of each table. The tables vary the controls, measures of reforms, and samples used to assess parental behavior. It is hard to attribute the results to measurement error: validation studies show that time diaries closely track direct observation, and this paper reports significant effects of reforms on diary measures of time use by school staff and children.

In addition to null effects on time with children, reforms did not affect time spent by parents on market work or household production (Tables 9 and A9, Panel E). While parents did not adjust work hours, parents spent modestly less time at work with children present.

6 Discussion: school quantity versus school quality

The new findings presented in this paper speak to the efficiency of school spending, to the private costs of reforms, and to our ability to measure the impacts of reforms. When schools improve quantity, and not quality, conventional estimates likely understate both the true costs of reforms and the true achievement effects.

6.1 The efficiency of school spending

Schools raise quantity because it is cost efficient (for schools, although perhaps not for families) to raise quantity.

Raising quantity means adding hours for existing staff. It may be possible to scale hours faster than costs. That is, marginal costs for hours may be less than average costs. Lachowska et al. (2023) observe job transitions to estimate that the typical worker would accept half of the going wage for marginal hours. School staff, whose schedules are unusually constrained, may be more willing to accept additional hours at less than the going wage.

I find that the marginal cost of adding hours for existing staff is less than the average cost: the elasticities of staff working hours to total spending, compensation spending, and weekly earnings are all greater than one. I estimate elasticities of 1.6, 1.9, and 3.4, respectively. My estimates, using school finance reforms as a labor demand shifter, contribute the first causal evidence of marginal wages lower than average wages.

6.2 The private costs of reforms

Raising quantity imposes private costs on students and families. Time for school is taken out of time for other activities including leisure, market work, household production, or non-school investments in human capital. Since parents contribute to these activities, they may also be affected by learning time interventions and hence by reforms.

The cost of an hour put into school is the value of an hour taken out of another activity. This cost is fundamental in education production, and economists have taken it seriously at least since the work of Mincer (1958), Schultz (1960, 1961), and Ben-Porath (1967). Becker (1965) wrote that, "Most economists have now fully grasped the importance of foregone earnings in the educational process and, more generally, in all investments in human capital, and criticize educationalists and others for neglecting them," yet, in the present literatures on school finance and on learning time (on the expansion of school days and school years) the opportunity cost of students' time remains neglected.

The private cost of additional instruction is a factor of four larger than the public cost. Since all people face the same time budget, we may assume the value of time for students is similar to the value of time for staff. Since students outnumber staff seven to one, and respond to reforms at about two-thirds of the magnitude, each hour of work by staff (the public cost) is matched by 4-5 hours from students (the private cost). Thus, for a given gain in achievement, an intervention that raises quantity is more costly than an intervention that raises quality. Cost-benefit analysis must account for private cost accordingly.

How private costs are allocated matters. The complete time diary data used in this paper lead to two novel conclusions about family behavior:

First, children's market and home production is an important margin of adjustment. This is the first study of crowd-out for work-eligible children. Students allocate time costs from the extended school year to production (apparently home production), rather than to leisure or non-school investments, so the ultimate cost is lower consumption. This is the first evidence that children reallocate time from production to education activities as schools adjust inputs.[17] In Japan, Kawaguchi (2016) finds complementary evidence that a school-week reduction increases work-ineligible minor children's leisure activities.

Second, there is no crowd-out of parents' activities, because parents do not adjust school-related activities after reforms. This finding is surprising and deserves an explanation. Becker and Tomes (1986), Todd and Wolpin (2003), and Glewwe and Kremer (2006) have all argued that parents may reduce school-related activities following public investment in education. However, I find that parents are already largely crowded out of substitutable activities before reforms, consistent with qualitative evidence in Bergman (2021). This may be explained in our American setting by the long and deliberate replacement of parental inputs with public inputs over the past half-century, documented by Flyer and Rosen (1997).

6.3 The understated benefits of reforms for achievement

Raising the quantity of instruction could raise test scores, if inputs and tests were appropriately timed, but they are not: NAEP tests are administered in January, February, and March; hours of instruction are added in May, June, and July. If hours were added evenly, the NAEP would fail to measure almost half of the input until the following year; and as it is, the NAEP measures almost none of the input until the following year.

In this section, I discuss the problems of delayed measurement inherent in the NAEP and in cumulative-exposure estimates, and their contribution to reported achievement gains. First, I consider the immediate effect of reforms on achievement, based on this paper's estimates of additional learning time and of daily value added. Next, I consider the (annualized) cumulative effect of reforms, and find that it is much smaller. Finally, I show that my estimates are reconciled by consensus estimates of fadeout.

I conclude by discussing the implications for adult outcomes, and for the hierarchy of evidence on school finance reforms.

6.3.1 Reform effects: immediate

The rate of learning I estimate from the timing of NAEP tests, 0.51σ per school year (Table 2, lower bound of the 95 percent CI), is not subject to fadeout, because measurement is immediate. In Section 5, I estimate a 5.5 percent increase in the length of the school year (90 percent CI: 0.010 to 0.100), per reform.

[17] I could not find other empirical work on school inputs and students' market and household production. However, my findings relate to the literature on conditional cash transfers, schooling, and child labor: see for example Sviatschi (2022), Edmonds and Schady (2012), and Ravallion and Wodon (2001).

Taken together, my estimates imply immediate effects on the NAEP of 0.028σ per year (90 percent CI: 0.005 to 0.051), per reform.[18]

Given that daily value added is measured in the middle of the year, and hours of instruction are added at the end, this is an extrapolation.

However, extrapolation could be justified. One day is like another. Test score gains are linear in the NAEP testing window,[19] and are not capped: there are no ceiling effects even for students at the 95th percentile (Klopfer, 2025). Added days outside the window are allocated to the usual mix of travel, class, and homework. This suggests the textbook-standard argument that, by replicating usual inputs, schools should replicate usual outputs (Koopmans, 1957); for a formal treatment of the production function, see Appendix A.3.

6.3.2 Reform effects: cumulative

My estimate of the cumulative effect of reforms on achievement gains on the NAEP is 0.000σ per year of exposure (95 percent CI: -0.004 to 0.005). See Section 5.

The discrepancy between the estimates is significant. In a Welch test of the null hypothesis that immediate effects are less than or equal to cumulative effects, if the estimates are uncorrelated, then the p-value is 0.05; if they are correlated, then the p-value is smaller.

6.3.3 Fadeout predicts small cumulative effects

Delayed measurement and consensus estimates of fadeout reconcile immediate and cumulative effects of reforms.

Why do delayed measurement and fadeout matter? The cumulative effect of reforms on achievement is estimated by regressing test scores on years of exposure, so that measured effects are an average of immediate effects multiplied by fadeout at each lag.

Fadeout at long lags is substantial. Chetty, Friedman, and Rockoff (2014) find that effects fade out to 50, 33, 25, and 20 percent over four years; Gilraine and Pope (2021) find 34, 21, 18, and 16 percent; and Jacob, Lefgren, and Sims (2010) find as little as 20 and 16 percent over two years. Chetty et al. (2011) find that kindergarten effects fade out to less than 10 percent by 4th grade, with estimates through 8th grade indistinguishable from zero.

[18] My model-based achievement effect is conservative relative to direct quasi-experimental estimates from term length by Frazier and Morrison (1998) and Sanz and Tena (2023); and from summer school attendance by Matsudaira (2008) and Pyne, Messner, and Dee (2023). My smaller estimate may relate to my choice of model inputs, and to their within-sample test-score standardization.

[19] For a window covering the remainder of the school year from March through June, Fitzpatrick, Grissmer, and Hastedt (2011) test for and cannot reject linearity in kindergarten and 1st-grade test score gains. They also find that time outside of school does not raise test scores, confirming zero counterfactual gains.

How does fadeout add up? For a student exposed from kindergarten through grade 4, and tested in grade 4 before inputs are delivered, these figures imply that the cumulative effect could range between $(0+50+33+25+20)/5 = 26$ percent and $(0+20+16+16+10)/5 = 12$ percent, or less, of the immediate effect; for a formal treatment of the bias in estimates, see Appendix A.4. Fadeout over eight years is stronger.

Fadeout predicts small effects. Multiplying an immediate effect of 0.028σ by the estimated fadeout over four years yields a prediction of 0.007 to 0.003σ, within the confidence interval for the estimated cumulative effect. Fadeout to zero is possible.

Fadeout may explain the pattern of null or negative cumulative achievement estimates found alongside positive estimates in nearly all studies of school finance reforms.

6.3.4 The hierarchy of evidence

Fadeout presents a serious problem for the literature on school finance reforms, and on school finance in general. What can be trusted?

Achievement effects from multi-year exposures are most problematic. Even when same-year inputs are captured, past-year inputs subject to fadeout dominate cumulative estimates.

Achievement effects from same-year exposures are preferable, but researchers should report input and test timing, which determine measurability.[20]

Model-based estimates, building on inputs, input-output relationships, and clearly-stated assumptions, should help when input and test timing interfere with measurement. While this approach is unusual in the school finance literature, model-based estimates avoid known bias from missing inputs and fadeout. By combining a broader range of evidence, they also avoid cherry-picking and improve precision (Athey et al., 2025). Structural labor economists and macroeconomists rely heavily on model-based estimates in the human capital literature, to weigh the magnitudes in cost-input-output relationships.

In the long run, it would be better to focus on economic outcomes, which are not subject to fadeout. It should be emphasized more in the school finance literature that achievement effects fade out, but re-emerge in economic outcomes (Currie and Thomas, 1995; and Garces, Thomas, and Currie, 2002; Deming, 2009; Chetty et al., 2011; Fredriksson, Öckert, and Oosterbeek, 2013; Chetty, Friedman, and Rockoff, 2014).

[20]Only three states (Kentucky, North Carolina, and South Carolina) test in the last 1-3 weeks of the school year, and apart from Florida, all others test starting in April or earlier. Only one study attempts to measure the same-year effects of school finance reforms: Dinerstein and Smith (2021), who residualize current-grade test scores on previous-grade test scores. If inputs are timed before tests, their approach can be justified under the assumptions justifying Equation (10) in Todd and Wolpin (2003): that the test-score impact of a given input does not vary with age, and that the impact fades out at a fixed rate for each year since input. The authors find effects 2-6 times larger than annual effects from multi-year exposures.

This hierarchy is relevant for the interpretation of past evidence: fadeout and re-emergence resolve inconsistencies among studies of school finance reforms. Effects re-emerge in studies of Michigan's Proposal A of 1994: achievement effects are mixed (Papke, 2000, 2005; Chaudhary, 2009; Roy, 2011; Baron, Hyman, and Vasquez, 2024), while educational attainment (Hyman, 2017) and crime (Baron, Hyman, and Vasquez, 2024) effects are large. Nationally, reforms increase attainment (Candelaria and Shores, 2019; Rothstein and Schanzenbach, 2022), raise earnings (Jackson, Johnson, and Persico, 2016; Rothstein and Schanzenbach, 2022; Biasi, 2023), and raise civic participation (Asker, Brunner, and Ross, 2024).

However, more evidence is needed before making conclusions about school finance reforms specifically. The evidence on multi-year achievement effects could be given less weight due to the issues with fadeout and measurement highlighted here, but while that would give more weight to other evidence, only one study examines same-year achievement effects, and only a handful of studies focus on long-run economic outcomes. This study resolves the puzzle of weak multi-year achievement effects, not the overall spending debate.

7 Conclusion

This paper contributes three novel observations to the school finance literature. First, given funding, schools spend on extended learning time at low marginal cost; this resolves long-standing questions about the use of reform funding to raise teacher salaries. Second, the private costs of extended learning time, and hence of school finance reforms, are greater than the public costs. Finally, the achievement effects of reforms are small, not due to any broken link along the chain from spending to human capital production, but instead due to an unfortunate combination of delayed inputs, early testing, and cumulative measurement.

The implications are straightforward: policymakers should give more attention to private costs; researchers should worry less about the nominal accounting and achievement effects of reforms, giving much more attention to real inputs and real outputs.

8 References

Abott, Carolyn, Vladimir Kogan, Stéphane Lavertu, and Zachary Peskowitz. 2020. "School district operational spending and student outcomes: Evidence from tax elections in seven states," *Journal of Public Economics* 183:104142

Agüero, Jorge and Trinidad Beleche. 2013. "Test-Mex: Estimating the effects of school year length on student performance in Mexico," *Journal of Development Economics* 103, pp. 353-361

Anderson, Simon, Maria Humlum, and Anne Nandrup. 2016. "Increasing instruction time in school does increase learning," *Proceedings of the National Academy of Science* 113:27, pp. 7481-7484

Angrist, Joshua, and Victor Lavy. 2009. "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial," *American Economic Review* 99:4, pp. 1384-1414

Angrist, Joshua, Victor Lavy, Jetson Leder-Luis, and Adi Shany. 2019. "Maimonides' Rule Redux," *American Economic Review: Insights* 1:3, pp. 309-324

Araujo, M. Caridad, Pedro Carneiro, Yyannú Cruz-Aguayo, and Norbert Schady. 2016. "Teacher Quality and Learning Outcomes in Kindergarten," *Quarterly Journal of Economics* 131:3, pp. 1415-1453

Asker, Erdal, Eric Brunner, and Steve Ross. 2024. "The impact of school spending on civic engagement: Evidence from school finance reforms," *Journal of Urban Economics* 143:103688

Attanasio, Orazio, Sarah Cattan, Emla Fitzsimmons, Costas Meghir, and Marta Rubio-Codina. 2020. "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia," *American Economic Review* 110:1, pp. 48-85

Athey, Susan, Raj Chetty, Guido Imbens, and Hyunseung Kang. "The Surrogate Index: Combining Short-Term Proxies to Estimate Long-Term Treatment Effects More Rapidly and Precisely," NBER Working Paper No. 26463

Aucejo, Esteban and Teresa Foy Romano. 2016. "Assessing the effect of school days and absences on test score performance," *Economics of Education Review* 55, pp. 70-87

Baron, E. Jason. 2022. "School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin," *American Economic Journal: Economic Policy* 14:1, pp. 1-39

Baron, E. Jason, Joshua Hyman, and Brittany Vasquez. 2024. "Public School Funding, School Quality, and Adult Crime," *Review of Economics and Statistics*

Barwick, Panle Jia, Siyu Chen, Chao Fu, and Teng Li. 2024. "Digital Distractions with Peer Influence: The Impact of Mobile App Usage on Academic and Labor Market Outcomes," NBER Working Paper No. 33054

Bates, Michael and Andrew Johnston. 2024. "Do Pensions Enhance Worker Effort and Selection? Evidence from Public Schools," EdWorkingPaper: 24-957

Baum-Snow, Nathaniel and Derek Neal. 2009. "Mismeasurement of usual hours worked in the census and ACS," *Economics Letters* 102:1, pp. 39-41

Becker, Gary. 1965. "A Theory of the Allocation of Time," *Economic Journal* 75:299, pp. 493-517

Becker, Gary and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics* 4:3 (Part 2), pp. S1-S39

Bellei, Cristián. 2009. "Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile," *Economics of Education Review* 28:5, pp. 629-640

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75:4 Part I, pp. 352-365

Bergman, Peter. 2021. "Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment," *Journal of Political Economy* 129:1, pp. 286-322

Biasi, Barbara. 2021. "The Labor Market for Teachers under Different Pay Schemes," *American Economic Journal: Economic Policy* 13:3, pp. 63-102

Biasi, Barbara. 2023. "School Finance Equalization Increases Intergenerational Mobility," *Journal of Labor Economics* 41:1, pp. 1-38

Biasi, Barbara, Julian Lafortune, and David Schönholzer. 2025. "What Works and for Whom? Effectiveness and Efficiency of School Capital Investments Across the United States," *Quarterly Journal of Economics* 140:3, pp. 2329-2379

Bonnesrønning, Hans. 2004. "The determinants of parental effort in education production: do parents respond to changes in class size?" *Economics of Education Review* 23, pp. 1-9

Brunner, Eric, Joshua Hyman, and Andrew Ju. 2020. "School Finance Reforms, Teachers' Unions, and the Allocation of School Resources," *Review of Economics and Statistics* 102:3, pp. 473-489

Brunner, Eric, Ben Hoen, and Joshua Hyman. 2022. "School district revenue shocks, resource allocations, and student achievement: Evidence from the universe of U.S. wind energy installations," *Journal of Public Economics* 206:104586

Buerger, Christian, Seung Hyeong Lee, and John Singleton. 2021. "Test-Based Accountability and the Effectiveness of School Finance Reforms," *AEA Papers and Proceedings* 111, pp. 455-459

Bureau of Labor Statistics. 2016a. "American Time Use Survey Activity Lexicon 2003-2015"

Bureau of Labor Statistics. 2016b. "American Time Use Survey User's Guide: Understanding ATUS 2003 to 2015"

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

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, pp. 49-82

Card, David, Leah Clark, Ciprian Domnisoru, and Lowell Taylor. 2024. "School Equalization in the Shadow of Jim Crow: Causes and Consequences of Resource Disparity in Mississippi circa 1940," CES Working Paper 24-45

Carlson, Deven and Stéphane Lavertu. 2018. "School Improvement Grants in Ohio: Effects on Student Achievement and School Administration," *Educational Evaluation and Policy Analysis* 40:3, pp. 287-315

Carlsson, Magnus, Gordon Dahl, Björn Öckert, and Dan-Olof Rooth. 2015. "The Effect of Schooling on Cognitive Skills," *Review of Economics and Statistics* 97:3, pp. 533-547

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

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

Chaudhary, Latika. 2009. "Education inputs, student performance and school finance reform in Michigan," *Economics of Education Review* 28, pp. 90-98

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

Clark, Melissa. 2003. "Education Reform, Redistribution, and Student Achievement: Evidence From the Kentucky Education Reform Act," working paper

Conlin, Michael and Paul Thompson. 2017. "Impacts of new school facility construction: An analysis of a state-financed capital subsidy program in Ohio," *Economics of Education Review* 59, pp. 13-28

Corcoran, Sean and William Evans. 2007. "Equity, Adequacy, and the Evolving State Role in Education Finance," in *Handbook of Research in Education Finance and Policy*, Eds. Helen F. Ladd and Edward B. Fiske. New York, NY: Routledge, pp. 332-356

Cortes, Kalena and Joshua Goodman. 2014. "Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement," *American Economic Review: Papers & Proceedings* 104:5, pp. 400-405

Cortes, Kalena, Joshua Goodman, and Takako Nomi. 2015. "Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra," *Journal of Human Resources* 50:1, pp. 108-158

Cunha, Flavio and James Heckman. 2007. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources* 43:4, pp. 738-782

Currie, Janet and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85:3, pp. 341-364

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Scores," *American Economic Journal: Applied Economics* 5:2, pp. 29-57

Datar, Ashlesha and Bryce Mason. 2008. "Do reductions in class size 'crowd out' parental investment in education?" *Economics of Education Review* 27, pp. 712-723

de Chaisemartin, Clément and Xavier D'Haultfouille. 2025. "Credible Answers to Hard Questions: Differences-in-Differences for Natural Experiments," working paper

De Fraja, Gianni, Tania Oliveira, and Luisa Zanchi. 2010. "Must Try Harder: Evaluating the Role of Effort in Educational Attainment," *Review of Economics and Statistics* 92:3, pp. pp. 577-597

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics* 1:3, pp. 111-134

Dinerstein, Michael and Troy Smith. 2021. "Quantifying the Supply Response of Private Schools to Public Policies," *American Economic Review* 111:10, pp. 3376-3417

Dobbie, Will and Roland Fryer. 2013. "Getting Beneath the Veil of Effective Schools: Evidence from New York City," *American Economic Journal: Applied Economics* 5:4, pp. 28-60

Downes, Thomas. 1992. "Evaluating the Impact of School Finance Reform on the Provision of Public Education: the California Case," *National Tax Journal* 45:4, pp. 405-419

Downes, Thomas and David Figlio. 1997. "School Finance Reforms, Tax Limits, and Student Performance: Do Reforms Level-Up or Dumb Down?" IRP Discussion Paper No. 1142-97

Edmonds, Eric and Norbert Schady. 2012. "Poverty Alleviation and Child Labor," *American Economic Journal: Economic Policy* 4:4, pp. 100-124

Figlio, David, Kristian Holden, and Umut Ozek. 2018. "Do Students Benefit from Longer School Days? Regression Discontinuity Evidence from Florida's Additional Hour of Literacy Instruction," CALDER Working Paper No. 201-0818-1

Fischer, Martin, Martin Karlsson, Therese Nilsson, and Nina Schwarz. 2020. "The Long-Term Effects of Long Terms–Compulsory Schooling Reforms in Sweden," *Journal of the European Economic Association* 18:6, pp. 2776-2823

Fitzpatrick, Maria, David Grissmer, and Sarah Hastedt. 2011. "What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment," *Economics of Education Review* 30:2, pp. 269-279

Flood, Sarah M., Miriam King, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0 [dataset]. Minneapolis, MN: University of Minnesota, 2015. https://doi.org/10.18128/D030.V4.0 (accessed October 8, 2016)

Flyer, Frederick and Sherwin Rosen. 1997. "The New Economics of Teachers and Education," *Journal of Labor Economics* 15:1, pt. 2, pp. S104-S139

Frazier, Julie and Frederick Morrison. 1998. "The Influence of Extended-Year Schooling on Growth of Achievement and Perceived Competence in Early Elementary School," *Child Development* 69:2, pp. 495-517

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2013. "Long-Term Effects of Class Size," *Quarterly Journal of Economics* 128:1, pp. 249-285

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2016. "Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule," *Journal of Human Resources* 51:4, pp. 832-868

Fryer, Roland. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials," *Quarterly Journal of Economics* 126, pp. 1755-1798

Fryer, Roland. 2013. "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools," *Journal of Labor Economics* 31:2, pp. 373-427

Fryer, Roland. 2014. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments," *Quarterly Journal of Economics* 129:3, pp. 1355-1407

Fu, Chao and Nirav Mehta. 2018. "Ability Tracking, School and Parental Effort, and Student Achievement: A Structural Model and Estimation," *Journal of Labor Economics* 36:4, pp. 923-979

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start," *American Economic Review* 92:4, pp. 999-1012

Gelber, Alexander and Adam Isen. 2013. "Children's schooling and parents' behavior: Evidence from the Head Start Impact Study," *Journal of Public Economics* 101, pp. 25-38

Gigliotti, Philip and Lucy Sorensen. 2018. "Educational resources and student achievement: Evidence from the Save T Harmless provision in New York State," *Economics of Education Review* 66, pp. 167-182

Gilraine, Michael and Nolan Pope. 2021. "Making Teaching Last: Long-Run Value-Added," NBER Working Paper No. 29555

Glewwe, Paul and Michael Kremer. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries," *Handbook of the Economics of Education* Vol. 2, Eds. Eric Hanushek and Finis Welch. Amsterdam: North-Holland, pp. 945-1017

Goldstein, Jessica and Josh McGee. 2021. "Did Spending Cuts During the Great Recession Really Cause Student Outcomes to Decline?" EdWorkingPaper 20-303

Goncalves, Felipe. 2015. "The Effects of School Construction on Student and District Outcomes: Evidence from a State-Funded Program in Ohio," working paper

Goodman, Sarena and Lesley Turner. 2013. "The Design of Teacher Incentive Pay and Educational Outcomes: Evidence from the New York City Bonus Program," *Journal of Labor Economics* 31:2, pp. 409-420

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing," *Journal of Econometrics* 225:2, pp. 254-277

Greaves, Ellen, Ifthikhar Hussain, Birgitta Rabe, and Imran Rasul. 2023. "Parental Responses to Information about School Quality: Evidence from Linked Survey and Administrative Data," *Economic Journal* 133, pp. 2334-2402

Guryan, Jonathan. 2001. "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts," NBER Working Paper No. 8269

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. "Parental Education and Parental Time with Children," *Journal of Economic Perspectives* 22:3, pp. 23-46

Hamermesh, Daniel, Harley Frazis, and Jay Stewart. 2005. "Data Watch: The American Time Use Survey," *Journal of Economic Perspectives* 19:1, pp. 221-232

Hansen, Benjamin. 2011. "School Year Length and Student Performance: Quasi-Experimental Evidence," working paper

Handel, Danielle Victoria and Eric Hanushek. 2023. "US School Finance: Resources and outcomes," *Handbook of the Economics of Education* Vol. 7, pp. 143-226

Handel, Danielle Victoria and Eric Hanushek. 2024. "Contexts of Convenience: Generalizing from Published Evaluations of School Finance Policies," *Evaluation Review* 48:3, pp. 461-494

Hanushek, Eric, John Kain, and Steven Rivkin. 2004. "Why Public Schools Lose Teachers," *Journal of Human Resources* 39:2, pp. 326-354

Hofferth, Sandra L., Sarah M. Flood, and Matthew Sobek. American Time Use Survey Data Extract Builder: Version 2.5 [dataset]. College Park, MD: University of Maryland and Minneapolis, MN: University of Minnesota, 2015. https://doi.org/10.18128/D060.V2.5 (accessed August 25, 2016)

Hong, Kai and Ron Zimmer. 2016. "Does Investing in School Capital Infrastructure Improve Student Achievement?" *Economics of Education Review* 53, pp. 143-158

Houtenville, Andrew J. and Karen S. Conway. 2008. "Parental Effort, School Resources, and Student Achievement," *Journal of Human Resources* 43:2, pp. 437-453

Hoxby, Caroline. 2001. "All School Finance Equalizations Are Not Created Equal," *Quarterly Journal of Economics* 116:4, pp. 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, pp. 256-280

Jackson, C. Kirabo, Rucker 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, pp. 157-218

Jackson, C. Kirabo, Cora Wigger, and Heyu Xiong. 2021. "Do School Spending Cuts Matter? Evidence from the Great Recession," *American Economic Journal: Economic Policy* 13:2, pp. 304-335

Jackson, C. Kirabo and Claire Mackevicius. 2024. "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States," *American Economic Journal: Applied Economics* 16:1, pp. 412-446

Jacob, Brian, Lars Lefgren, and David Sims. 2010. "The Persistence of Teacher-Induced Learning Gains," *Journal of Human Resources* 45:4, pp. 915-943

Juster, F. Thomas and Frank Stafford. 1991. "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement," *Journal of Economic Literature* 29:2, pp. 471-522

Kane, Thomas and Douglas Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," NBER Working Paper No. 14607

Kawaguchi, Daiji. 2016. "Fewer school days, more inequality," *Journal of the Japanese and International Economies* 39, pp. 35-52

Kline, Patrick and Christopher Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start," *Quarterly Journal of Economics* 131:4, pp. 1795-1848

Klopfer, John. 2025. "Learning Time and Achievement: Evidence from a Nationwide Natural Experiment"

Kogan, Vladimir, Stéphane Lavertu, and Zachary Peskowitz. 2017. "Direct democracy and administrative disruption," *Journal of Public Administration Research and Theory* 27:3, pp. 381-399

Koopmans, Tjalling. 1957. *Three Essays on the State of Economic Science*, New York, NY: McGraw-Hill

Kreisman, Daniel and Matthew Steinberg. 2019. "The effect of increased funding on student achievement: Evidence from Texas's small district adjustment," *Journal of Public Economics* 176, pp. 118-141

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A. Woodbury. 2023. "Work Hours Mismatch," NBER Working Paper No. 31205

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

Lafortune, Julien and David Schönholzer, 2022. "The Impact of School Facility Investments on Students and Homeowners: Evidence from Los Angeles," *American Economic Journal: Applied Economics* 14:3, pp. 254-289

Lavy, Victor. 2020. "Expanding School Resources and Increasing Time on Task: Effects on Students' Academic and Noncognitive Outcomes," *Journal of the European Economic Association* 18:1, pp. 232-265

Martorell, Paco, Kevin Stange, and Isaac McFarlin, Jr. 2016. "Investing in schools: capital spending, facility conditions, and student achievement," *Journal of Public Economics* 140, pp. 13-29

Matsudaira, Jordan. 2008. "Mandatory summer school and student achievement," *Journal of Econometrics* 142:2, pp. 829-850

Matsudaira, Jordan, Adrienne Hosek, and Elias Walsh. 2012. "An integrated assessment of the effects of Title I on school behavior, resources, and student achievement," *Economics of Education Review* 31, pp. 1-14

Miller, Corbin. 2022. "The Effect of Education Spending on Student Achievement: Evidence from Property Values and School Finance Rules," working paper

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy* 66:4, pp. 281-302

National Center for Education Statistics (NCES). 1990-2013. School District Finance Survey (SDFS) (F-33): Public-Use Data Files [dataset]. Washington, DC: US Department of Education, Institute of Education Sciences

National Center for Education Statistics (NCES). 2003-2012. Private School Universe Survey (PSS): Public-Use Data Files [dataset]. Washington, DC: US Department of Education, Institute of Education Sciences

National Center for Education Statistics (NCES). 2003-2013. National Assessment of Educational Progress (NAEP): Restricted-Use Data Files [dataset]. Washington, DC: US Department of Education, Institute of Education Sciences

Neilson, Christopher and Seth Zimmerman. 2014. "The effect of school construction on test scores, school enrollment, and home prices," *Journal of Public Economics* 120, pp. 18-31

Papke, Leslie. 2000. "Final Report: Michigan Applied Public Policy Research Project on K-12 School Finance," mimeo, Michigan State University, East Lansing, MI

Papke, Leslie. 2005. "The effects of spending on test pass rates: evidence from Michigan," *Journal of Public Economics* 89:5-6, pp. 821-839

Parinduri, Rasyad. 2014. "Do children spend too much time in schools? Evidence from a longer school year in Indonesia," *Economics of Education Review* 41, pp. 89-104

Parker, Emily. 2016. "Constitutional obligations for public education," Education Commission of the States Report

Pischke, Jörn-Steffen. 2007. "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years," *Economic Journal* 117:523, pp. 1216-1242

Pop-Eleches, Christian and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses," *American Economic Review* 103:4, pp. 1289-1324

Price, Joseph. 2008. "Parent-Child Quality Time: Does Birth Order Matter?" *Journal of Human Resources* 43:1, pp. 240-265

Pyne, James, Erica Messner, and Thomas Dee. "The Dynamic Effects of a Summer Learning Program on Behavioral Engagement in School," *Education Finance and Policy* 18:1, pp. 127-155

Ramey, Garey and Valerie Ramey. 2010. "The Rug Rat Race," *Brookings Papers on Economic Activity* Spring 2010, pp. 129-199

Ravallion, Martin and Quentin Wodon. 2001. "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy," *Economic Journal* 110:462, pp. 158-175

Reardon, Sean, Demetra Kalogrides, and Andrew Ho. 2021. "Validation Methods for Aggregate-Level Test Scale Linking: A Case Study Mapping School District Test Score Distributions to a Common Scale," *Journal of Educational and Behavioral Statistics* 46:2, pp. 138-161

Robinson, John. 1985. "The Validity and Reliability of Diaries versus Alternative Time Use Measures," in *Time, Goods, and Well-Being*, Eds. F. Thomas Juster and Frank P. Stafford. Ann Arbor, MI: Institute for Social Research, pp. 33-62

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," *Quarterly Journal of Economics* 125:1, pp. 175-214

Rothstein, Jesse. 2015. "Teacher Quality Policy When Supply Matters," *American Economic Review* 105:1, pp. 100-130

Rothstein, Jesse and Diane Schanzenbach. 2022. "Does Money Still Matter? Attainment and Earnings Effects of Post-1990 School Finance Reforms," *Journal of Labor Economics* 40:S1, pp. S141-S178

Roy, Joydeep. 2011. "Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence from Michigan," *Education Finance and Policy* 6:2, pp. 137-167

Sanz, Ismael and J.D. Tena. 2023. "Do 2 weeks of instruction time matter? Using a natural experiment to estimate the effect of a calendar change on students' performance," *Kyklos* 76, pp. 778-808

Schmidheiny, Kurt and Sebastian Seigloch. 2023. "On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization," *Journal of Applied Econometrics* 38:5, pp. 695-713

Schultz, Theodore. 1960. "Capital Formation by Education," *Journal of Political Economy* 68:6, pp. 571-583

Schultz, Theodore. 1961. "Investment in Human Capital," *American Economic Review* 51:1, pp. 1-17

Shores, Kenneth and Matthew Steinberg. 2019. "Schooling During the Great Recession: Patterns of School Spending and Student Achievement Using Population Data," *AERA Open* 5:3, pp. 1-29

Sims, David. 2008. "Strategic responses to school accountability measures: It's all in the timing," *Economics of Education Review* 27, pp. 58-68

Sims, David. 2011a. "Lifting all boats? Finance litigation, education resources, and student needs in the post-*Rose* era," *Education Finance and Policy* 6:4, pp. 455-485

Sims, David. 2011b. "Suing for your supper? Resource allocation, teacher compensation and finance lawsuits," *Economics of Education Review* 30, pp. 1034-1044

Stevens, Stanley Smith. 1946. "On the Theory of Scales of Measurement," *Science* 103:2684, pp. 677-680

Stinebrickner, Ralph and Todd Stinebrickner. 2008. "The Causal Effect of Studying on Academic Performance," *B.E. Journal of Economic Analysis & Policy* 8:1, Article 14

Sviatschi, Maria Micaela. 2022. "Making a Narco: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths," *Econometrica* 90:4, pp. 1835-1878

Taylor, Eric. 2014. "Spending more of the school day in math class: Evidence from a regression discontinuity in middle school," *Journal of Public Economics* 117, pp. 162-181

Thompson, Paul. 2020. "Is four less than five? Effects of four-day school weeks on student achievement in Oregon," *Journal of Public Economics* 192, 104308

Todd, Petra and Kenneth Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal* 113:485, pp. F3-F33

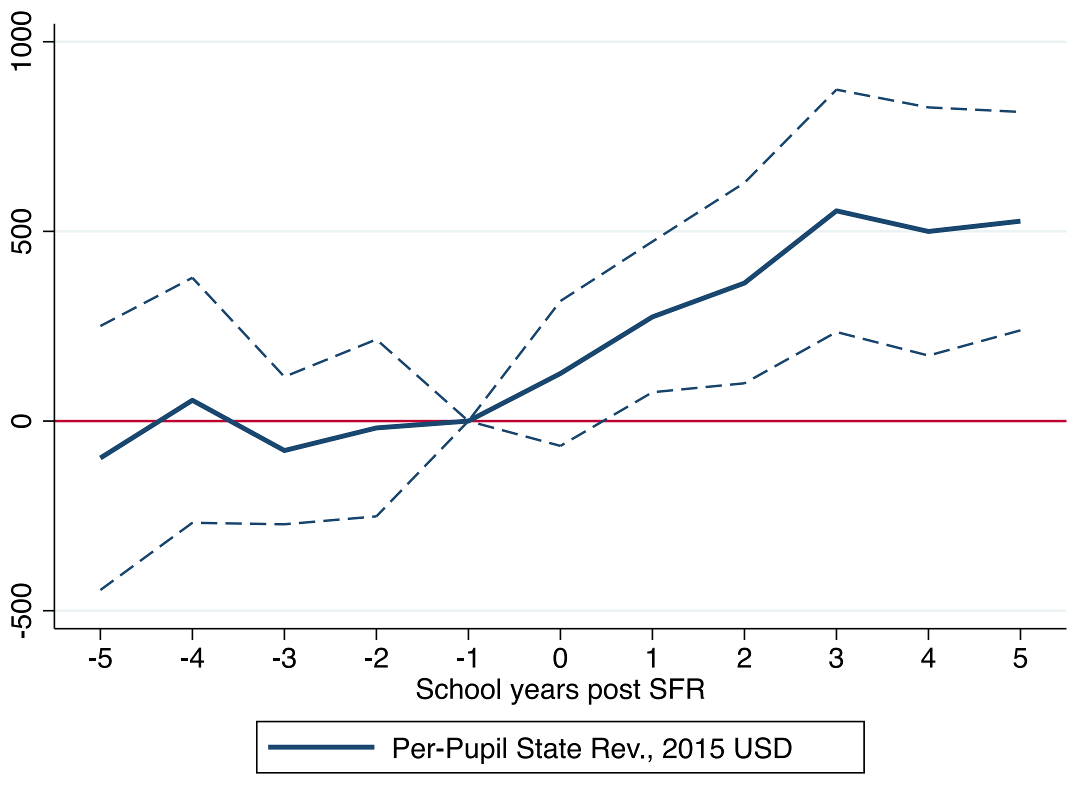
Todd, Petra and Kenneth Wolpin. 2018. "Accounting for Mathematics Performance of High School Students in Mexico: Estimating a Coordination Game in the Classroom," *Journal of Political Economy* 126:6, pp. 2608-2650

van der Klaauw, Wilbert. 2008. "Breaking the link between poverty and low student achievement: An evaluation of Title I," *Journal of Econometrics* 142, pp. 731-756

Weinstein, Meryle, Leanna Stiefel, Amy Ellen Schwartz, and Luis Chalico. 2009. "Does Title I Increase Spending and Improve Performance? Evidence from New York City," IESP Working Paper 09-09

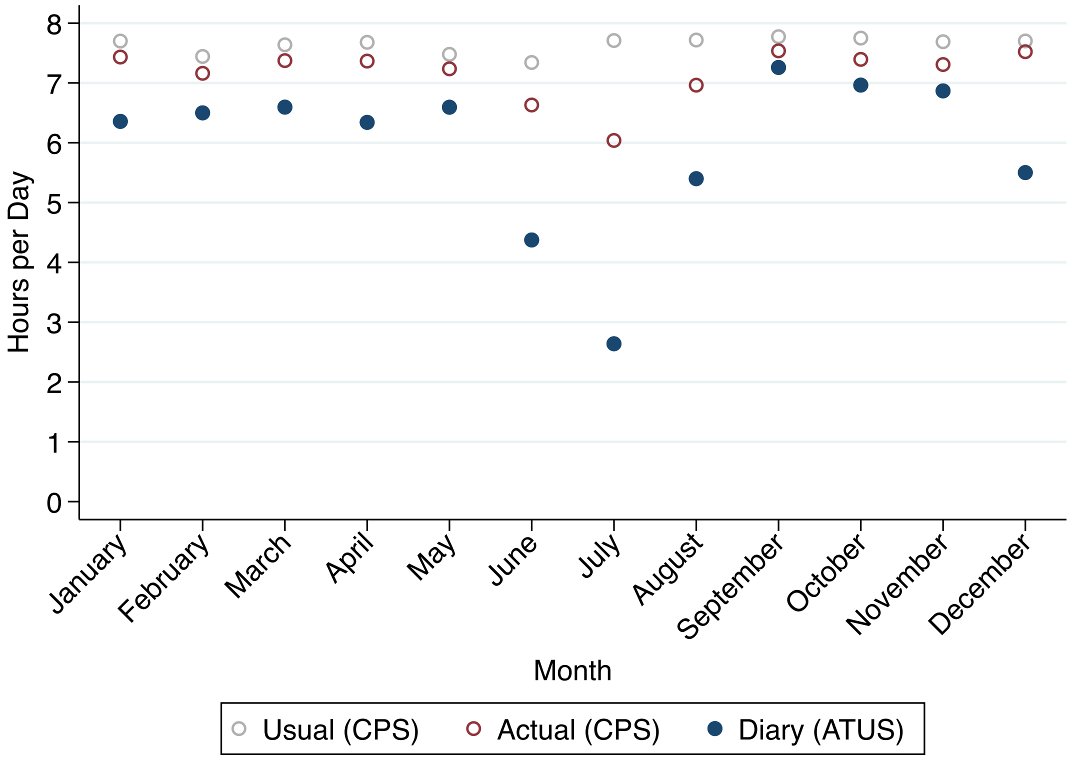
9 Figures

Figure 1: SFR and State Revenues to School Districts



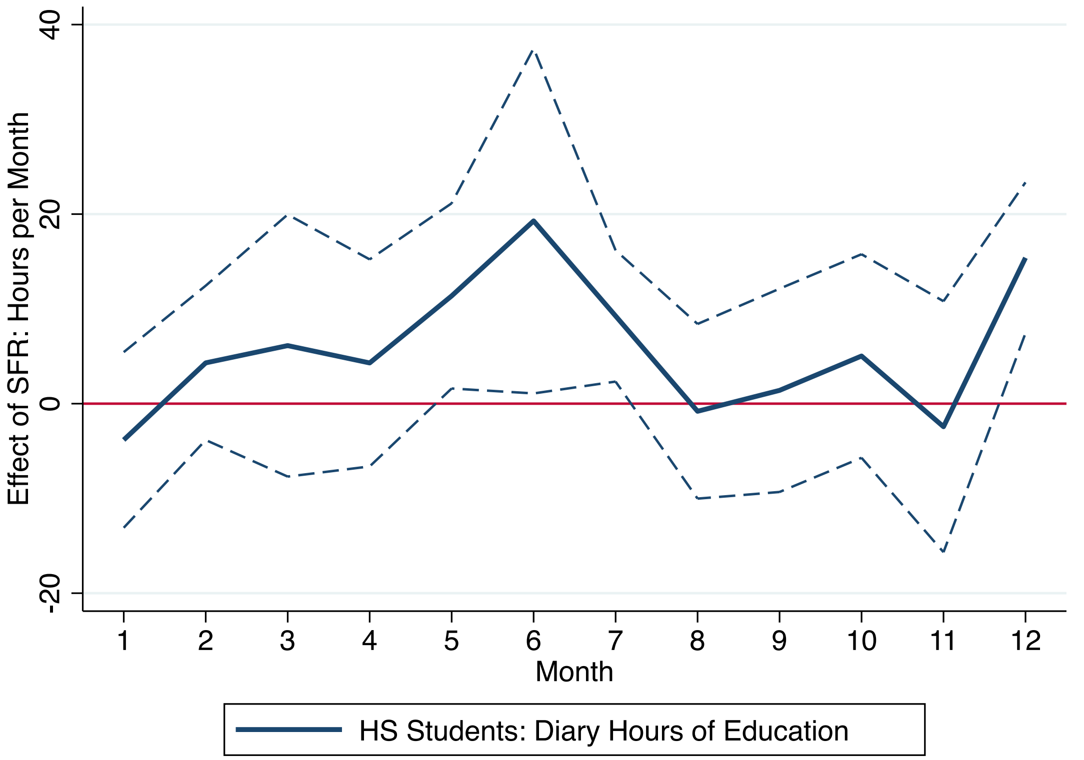
Data: school finance reforms (independent variable) from 1990 through 2013 (N=86 events, in 28 states). SDFS fiscal variables (dependent variables) annually from 1990 through 2013 (N=311193 observations, from 15087 local education authorities in 50 states and Washington, D.C.; observations from 1991 are omitted because of problems with the source data).
Outcomes: divided by fall enrollment (V33) and by CPI-U (base: school fiscal year July 2014-June 2015) to represent real per-pupil values.
Regression: reform coefficients are estimated using Equation (3), a state-level dynamic event study that includes district fixed effects, year fixed effects, and state trends. Year -1 is the last pre-reform year. Standard errors are clustered by state, and 95 percent confidence intervals are displayed with dotted lines.

Figure 2: Accuracy of the ATUS



Data: merged ATUS weekday diaries (N=3776) and CPS ORG (usual N=7652, actual N=7171) from 2003 through 2015, reporting employment in K-12 industry. Data include employed workers on school holidays, leave, and sick days.
Outcomes: diary (ATUS): activity code 05 (work and work-related activities). Usual (CPS) and Actual (CPS): hours of work as reported in CPS interview, divided by 5; variable hours omitted.
Regression: month coefficients are estimated from a regression on month indicators, with no controls. Confidence intervals are small and omitted for legibility.

Figure 3: Children: SFR and Diary Hours of Education, by Month



Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents eligible for K-12 schooling (N=9150, from 50 states and Washington, D.C.): aged 15 through 19, and either reporting not completed 12th grade or enrolled in high school. Date controls: month, day, and holiday. Grade controls: education, age.
Outcomes: multiplied by 365.25/12. Education: activity code 06.
Regression: reform coefficients are estimated using Equation (1), interacting the reform indicator with month indicators. 95-percent CI are shown by dotted lines, clustered by state.

10 Tables

Table 1: SFR and School Resources

	(1) State Rev.	(2) Federal Rev.	(3) Local Rev.	(4) Total Rev.	(5) Total Exp.	(6) Current Exp.	(7) Comp. Exp.
SFR	355***	-19	203	539**	611***	396*	338*
	(122)	(23)	(149)	(209)	(176)	(212)	(199)
Mean of D.V.	6461	1199	6427	14087	14063	11832	9329
β/Mean	0.055	-0.016	0.032	0.038	0.043	0.033	0.036
SFR	2003-	2003-	2003-	2003-	2003-	2003-	2003-
State FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
States	51	51	51	51	51	51	51
Districts	13761	13761	13761	13761	13761	13761	13761
Observations	144999	144999	144999	144999	144999	144999	144999
R^2	0.029	0.101	0.048	0.063	0.036	0.089	0.104

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). SDFS fiscal variables annually from 2003 through 2013.
Outcomes: divided by fall enrollment (V33) and by CPI-U (base: school fiscal year July 2014-June 2015) to represent real per-pupil values.
Regression: reform coefficients are estimated using Equation (1). Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2: SFR, School Days, their Interaction, and Achievement in SD

	(1)	(2)	(3)	(4)
SFR	0.003	0.000	-0.003	-0.001
	[-0.021,0.027]	[-0.004,0.005]	[-0.006,0.001]	[-0.046,0.045]
Week Days Comp. / 195				0.577
				[0.505,0.649]
Week Days Comp. / 195				0.000
x SFR				[-0.003,0.004]
SFR Variable	Cumul.	Cumul. Years	Cumul. Min(Years,Grade)	Cumul.
Year FE	Y	Y	Y	Y
School FE	Y	Y	Y	Y
N	2937470	2937470	2937470	2463870
States	51	51	51	51
R^2	0.24	0.24	0.24	0.24

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). NAEP from 2003 through 2013, grade 4 and 8, math and reading. Column 4 restricted to 2005 through 2013. School fixed effects restrict estimation to schools seen in more than one year.
Outcomes: mean of plausible NAEP-scale scores, standardized by mean and standard deviation within year, grade, and subject.
Regression: reform coefficients in Column 1-3 are estimated using Equation (1). Column 1: Cumulative SFR are as defined in the specification. Column 2: Cumulative years are the number of years post-reform, added up across all reforms, and are comparable to the post-event indicator multiplied by a trend term commonly used in studies with only one reform event per state. Column 3: Cumulative grades are equivalent to cumulative years, capped at up to 4 or 8 years of exposure per reform, depending on the grade at measurement. Column 4 jointly estimates the main reform and test-date effects from Equations (1) and (4), and their interaction. The test-date variable is divided by 195 to account for all of the week days, including holidays, in a nine-month school year: the coefficients measure value added per school-year weekday.
95-percent CI in brackets, standard errors clustered by state

Table 3: Education Inputs, Outputs, and Elasticities

Paper	Samp.	Input Defn.	Input Source	(1) ΔInput/ Input	Output Source	(2) ΔOutput/ Year	(3) ΔOutput/ Output	(4) Elasticity of Out- to Input
C03		Curr.	Tab 2, 3	0.203	Tab 9	0.010	0.020	0.099
		Exp.		[0.018, 0.388]		[-0.005, 0.025]	[-0.010, 0.049]	[-0.817, 1.015]
LRS18	All	Tot.	Tab 4	0.078	Tab 8	0.004	0.008	0.100
		Exp.		[0.028, 0.128]		[-0.002, 0.010]	[-0.004, 0.020]	[-0.539, 0.740]
LRS18	Q1	Tot.	Tab 2, 4	0.127	Tab 5	0.007	0.014	0.108
		Exp.		[0.059, 0.195]		[0.001, 0.013]	[0.002, 0.025]	[-0.425, 0.641]
LRS18	Q5	Tot.	Tab 2, 3	0.046	Tab 5	-0.001	-0.002	-0.043
		Rev.		[-0.001, 0.092]		[-0.005, 0.007]	[-0.010, 0.014]	[-1.061, 0.976]
BHJ20	All	Tot.	Tab 1;	0.046	Tab 7	0.007	0.014	0.299
		Rev.	Fig A3	[0.023, 0.069]		[0.003, 0.011]	[0.006, 0.022]	[-0.204, 0.802]
BHJ20	T1	Tot.	Tab 1;	0.069	Tab 7	0.009	0.018	0.256
		Rev.	Fig A3	[0.046, 0.092]		[0.005, 0.013]	[0.010, 0.025]	[-0.078, 0.591]
BHJ20	T3	Tot.	Tab 1;	0.023	Tab 7	0.004	0.008	0.342
		Rev.	Fig A3	[0.000, 0.046]		[-0.000, 0.008]	[-0.000, 0.016]	[-0.665, 1.348]
BLS21	Q1	Tot.	Tab A7	0.070	Tab 1	0.012	0.024	0.336
	Ac.	Exp.		[0.013, 0.093]		[0.000, 0.024]	[0.000, 0.047]	[-0.359, 1.031]
BLS21	Q1	Tot.	Tab A7	0.088	Tab 1	0.006	0.012	0.128
	No-Ac.	Exp.		[0.018, 0.158]		[0.000, 0.012]	[0.000, 0.024]	[-0.611, 0.867]
BLS21	Q5				Tab A3	0.008	0.016	
	Ac.					[-0.004, 0.020]	[-0.008, 0.039]	
BLS21	Q5				Tab A3	-0.003	-0.006	
	No-Ac.					[-0.009, 0.003]	[-0.018, 0.006]	
This		Tot.	Tab 1	0.043	Tab 2	0.000	0.000	0.000
		Exp.		[0.018, 0.068]		[-0.004, 0.005]	[-0.008, 0.010]	[-0.581, 0.581]
This		Tch.	Tab 6	0.020	Tab 2	0.000	0.000	0.000
		Pay		[0.012, 0.029]		[-0.004, 0.005]	[-0.008, 0.010]	[-0.425, 0.425]
This		Tch.	Tab 7	0.068	Tab 2	0.000	0.000	0.000
		Hrs.		[-0.007, 0.144]		[-0.004, 0.005]	[-0.008, 0.010]	[-1.110, 1.110]
This		Stu.	Tab 8	0.055	Tab 2	0.000	0.000	0.000
		Hrs.		[0.001, 0.110]		[-0.004, 0.005]	[-0.008, 0.010]	[-0.991, 0.991]

Definitions: Input is defined as reported. Output is defined as achievement in SD of the NAEP. ΔInput/Input, the percentage increase in input is calculated by dividing the treatment effect per year of reform exposure by mean input per year. ΔSD/Year is the treatment effect in SD per year of exposure; I divide C03's DDD estimate for 8th grade NAEP scores by 8, for comparability with her DDD estimate for spending. ΔSD/SD, the percentage increase in output measured in SD of the NAEP, is calculated by dividing ΔSD/Year of exposure by SD/Year mean growth, from Table 2, Column 4; conservatively, I use the lower end of the 95-percent confidence interval, 0.51 SD per year. Column 4 presents elasticities of achievement in SD to input.
References: C03: Clark (2003). LRS18: Lafortune, Rothstein, Schanzenbach (2018). BHJ20: Brunner, Hyman, Ju (2020). BLS21: Buerger, Lee, Singleton (2021).
Approximate 95-percent CI in brackets: I assume no uncertainty from numerators in Column 1 and 3, multiply standard errors by 2, and approximate the standard error of the elasticity in Column 4 by the Taylor expansion

$$\mathrm{Var}(x/y) \approx \frac{\mathrm{Var}(x)}{\mathrm{E}(y)^2} + \frac{\mathrm{E}(x)\mathrm{Var}(y)^2}{\mathrm{E}(y)^4} - \frac{2\mathrm{E}(x)\mathrm{Cov}(x,y)}{\mathrm{E}(y)^3} \tag{5}$$

assuming conservatively that $\mathrm{Cov}(x, y) = 0$.

Table 4: Workers: SFR and K-12 Staff Share: robustness to controls and reforms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	-0.00031	-0.00017	-0.00034	-0.00020	-0.00033	-0.00093	0.00056	-0.0011
	(0.00051)	(0.00053)	(0.00052)	(0.00077)	(0.00064)	(0.00096)	(0.00070)	(0.00069)
Mean of D.V.	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.062
β/Mean	-0.005	-0.003	-0.005	-0.003	-0.005	-0.015	0.009	-0.017
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul. One-Year Lag	First	Second	Third
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State Trends				Y				
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Demographic Controls		Y						
Employment Controls	Y	Y			Y	Y	Y	Y
States	51	51	51	51	51	51	51	51
Respondents	8893716	8893716	8893716	8893716	8893716	8893716	8893716	8893716
R^2	0.010	0.087	0.001	0.001	0.010	0.010	0.010	0.010

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS from 2003 through 2015, sample of respondents reporting employment, not enrolled. Date controls: month. Demographic controls: education, gender, and age. Employment controls: employment status (present or absent).
Outcomes: K-12 staff: 1990 industry code 842 (elementary and secondary schools).
Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Cumulative SFR are as defined in the specification; lagged SFR may have effects one year earlier than the standard variable; and the first, second, and third SFR are zero if that state had fewer SFR in the period from 2003 through 2013. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 5: K-12 Staff: SFR and Age: robustness to controls and reforms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	-0.060	-0.12	-0.058	-0.31**	0.031	-0.41**	0.16	0.30**
	(0.075)	(0.077)	(0.075)	(0.12)	(0.086)	(0.18)	(0.15)	(0.13)
Mean of D.V.	45	45	45	45	45	45	45	45
β/Mean	-0.001	-0.003	-0.001	-0.007	0.001	-0.009	0.004	0.007
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul. One-Year Lag	First	Second	Third
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State Trends				Y				
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Demographic Controls		Y						
Employment Controls	Y	Y			Y	Y	Y	Y
States	51	51	51	51	51	51	51	51
Respondents	564861	564861	564861	564861	564861	564861	564861	564861
R^2	0.005	0.045	0.004	0.005	0.005	0.005	0.005	0.005

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month. Demographic controls: education and gender. Employment controls: employment status (present or absent).

Outcomes: age in years.

Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Cumulative SFR are as defined in the specification; lagged SFR may have effects one year earlier than the standard variable; and the first, second, and third SFR are zero if that state had fewer SFR in the period from 2003 through 2013. Standard errors in parentheses, clustered by state.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 6: K-12 Staff: SFR and Weekly Earnings: robustness to controls and reforms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	18.4***	17.5***	18.3***	8.41	16.0***	5.72	39.2***	53.9***
	(3.84)	(2.71)	(3.81)	(5.47)	(5.81)	(7.05)	(6.51)	(3.75)
Mean of D.V.	909	909	909	909	909	909	909	909
β/Mean	0.020	0.019	0.020	0.009	0.018	0.006	0.043	0.059
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul. One-Year Lag	First	Second	Third
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State Trends				Y				
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Demographic Controls		Y						
Employment Controls	Y	Y			Y	Y	Y	Y
States	51	51	51	51	51	51	51	51
Respondents	141983	141983	141983	141983	141983	141983	141983	141983
R^2	0.033	0.323	0.033	0.033	0.033	0.033	0.033	0.033

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS ORG (months 4 and 8) from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month. Demographic controls: education, gender, and age. Employment controls: employment status (present or absent).

Outcomes: divided by CPI-U (base: school fiscal year July 2014-June 2015) to represent real values; recoded to zero if not employed as a wage or salary worker or otherwise not in earnings universe. 8.9 percent of values are missing in 1993. Weekly earnings before deductions are as reported in the CPS ORG: higher of direct report, or weekly hours multiplied by hourly wage rate.

Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Cumulative SFR are as defined in the specification; lagged SFR may have effects one year earlier than the standard variable; and the first, second, and third SFR are zero if that state had fewer SFR in the period from 2003 through 2013. Standard errors in parentheses, clustered by state.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 7: K-12 Staff: SFR and Diary Hours of Work: robustness to controls and reforms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Diary Hours of Work								
SFR	115.6*	107.5*	95.4	82.0	124.3*	162.0*	163.7*	131.6
	(63.8)	(61.9)	(81.7)	(76.3)	(71.2)	(95.1)	(87.4)	(163.5)
Mean of D.V.	1688	1688	1688	1688	1688	1688	1688	1688
β/Mean	0.068	0.064	0.057	0.049	0.074	0.096	0.097	0.078
Panel B: Usual Hours of Work, for Comparison								
SFR	36.6**	32.1*	31.4	77.5*	24.6	-2.99	85.2***	109.6***
	(17.9)	(18.6)	(27.7)	(38.8)	(20.7)	(31.4)	(26.6)	(19.9)
Mean of D.V.	2107	2107	2107	2107	2107	2107	2107	2107
β/Mean	0.017	0.015	0.015	0.037	0.012	-0.001	0.040	0.052
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul. One-Year Lag	First	Second	Third
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State Trends				Y				
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Demographic Controls		Y						
Employment Controls	Y	Y			Y	Y	Y	Y
States	51	51	51	51	51	51	51	51
Respondents (Diary)	6891	6891	6891	6891	6891	6891	6891	6891
Respondents (Usual)	6604	6604	6604	6604	6604	6604	6604	6604

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month, day, and holiday. Demographic controls: education, gender, and marital status. Employment controls: employment status (present or absent) and full or part time status.
Outcomes: multiplied by 365.25 to represent annual hours. Diary hours of work: activity code 0501. Usual hours of work: as reported in ATUS interview, divided by 5; variable hours recoded to missing.
Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Cumulative SFR are as defined in the specification; lagged SFR may have effects one year earlier than the standard variable; and the first, second, and third SFR are zero if that state had fewer SFR in the period from 2003 through 2013. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 8: Children: SFR and Diary Hours: robustness to controls and reforms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Education								
SFR	68.0**	72.8**	77.7**	63.8	93.1***	85.5	58.6	186.0***
	(33.4)	(33.0)	(36.1)	(60.6)	(33.8)	(56.0)	(71.6)	(34.4)
Mean of D.V.	1231	1231	1231	1231	1231	1231	1231	1231
β/Mean	0.055	0.059	0.063	0.052	0.076	0.069	0.048	0.151
Panel B: Class								
SFR	44.3	47.8	48.8	53.5	72.1**	54.8	13.0	179.4***
	(34.1)	(34.2)	(34.8)	(51.4)	(28.7)	(52.8)	(73.0)	(25.5)
Mean of D.V.	956	956	957	956	956	956	956	956
β/Mean	0.046	0.050	0.051	0.056	0.075	0.057	0.014	0.188
Panel C: Homework								
SFR	13.8*	14.8*	17.8*	6.99	16.0	26.3*	21.1	-9.63
	(8.18)	(8.17)	(9.47)	(17.0)	(12.7)	(15.3)	(15.4)	(11.4)
Mean of D.V.	240	240	240	240	240	240	240	240
β/Mean	0.057	0.061	0.074	0.029	0.066	0.110	0.088	-0.040
Panel D: Travel Related to Education								
SFR	10.9**	11.3**	11.1**	11.6	10.8**	12.9*	7.35	36.7***
	(5.01)	(5.02)	(4.93)	(8.18)	(4.53)	(6.87)	(9.93)	(3.66)
Mean of D.V.	74	74	74	74	74	74	74	74
β/Mean	0.148	0.153	0.151	0.158	0.147	0.175	0.100	0.499
Panel E: Market and Household Production								
SFR	-70.5*	-78.8**	-79.4**	-16.2	-100.4***	-60.5	-86.1	-220.1***
	(40.9)	(33.1)	(33.5)	(62.3)	(37.1)	(70.5)	(86.1)	(54.6)
Mean of D.V.	748	748	748	748	748	748	748	748
β/Mean	-0.094	-0.105	-0.106	-0.022	-0.134	-0.081	-0.115	-0.294
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul. One-Year Lag	First	Second	Third
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State Trends				Y				
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Grade Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls		Y	Y					
Parent Controls			Y					
States	51	51	51	51	51	51	51	51
Respondents	9150	9150	9142	9150	9150	9150	9150	9150

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents eligible for K-12 schooling: aged 15 through 19, and either reporting not completed 12th grade or enrolled in high school. Date controls: month, day, and holiday. Grade controls: education, age. Employment controls: employment status (present or absent) and full or part time status. Parent controls: own parent in household, parents' average years of education.
Outcomes: multiplied by 365.25 to represent annual hours. Education: activity code 06. Class: activity code 0601. Homework: activity code 0603. Travel related to education: related travel code 1806. Market and household production: activity codes 02 (household activities), 05 (work and work-related activities), 07 (consumer purchases), and related travel activity codes 1802, 1805, and 1807.
Regression: reform coefficients are estimated using Equation (1). Preferred specification in Column 1. Cumulative SFR are as defined in the specification; lagged SFR may have effects one year earlier than the standard variable; and the first, second, and third SFR are zero if that state had fewer SFR in the period from 2003 through 2013. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 9: K-12 Parents: SFR and Diary Hours: robustness to controls and reforms

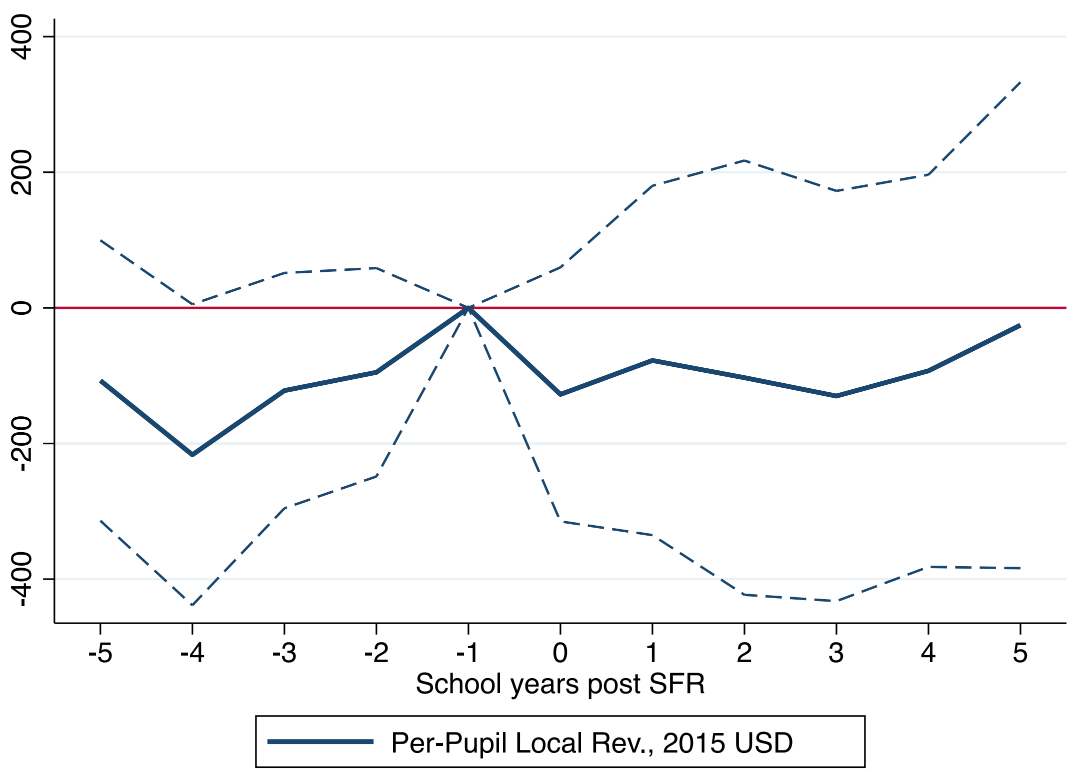
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Activities with Own Household Children								
SFR	8.49	3.16	14.0	21.6	10.3	-19.2	26.0	68.8**
	(21.1)	(18.8)	(18.6)	(23.0)	(27.2)	(29.7)	(50.7)	(26.3)
Mean of D.V.	1684	1684	1684	1684	1684	1684	1684	1684
β/Mean	0.005	0.002	0.008	0.013	0.006	-0.011	0.015	0.041
Panel B: Childcare with Own Household Children								
SFR	5.89	3.29	5.89	4.81	15.0**	-7.86	21.4	25.1
	(8.78)	(9.00)	(8.25)	(13.4)	(6.51)	(11.5)	(17.7)	(22.2)
Mean of D.V.	319	319	319	319	319	319	319	319
β/Mean	0.018	0.010	0.018	0.015	0.047	-0.025	0.067	0.078
Panel C: Childcare Related to Education with Own Household Children								
SFR	-4.17	-4.30	-3.83	-4.33	-3.50	-5.69	-8.08	-0.67
	(2.93)	(2.94)	(2.81)	(3.19)	(3.37)	(4.64)	(5.91)	(9.02)
Mean of D.V.	50	50	50	50	50	50	50	50
β/Mean	-0.084	-0.087	-0.077	-0.087	-0.070	-0.114	-0.162	-0.013
Panel D: Travel Related to Childcare Related to Education with Own Household Children								
SFR	0.89	0.81	1.03	0.62	0.27	2.80	-0.45	0.22
	(1.07)	(1.12)	(1.11)	(1.41)	(1.15)	(2.22)	(3.51)	(2.27)
Mean of D.V.	35	35	35	35	35	35	35	35
β/Mean	0.026	0.023	0.030	0.018	0.008	0.081	-0.013	0.006
Panel E: Market and Household Production								
SFR	9.69	13.4	-10.9	32.5	9.68	-5.55	48.2**	-6.84
	(14.4)	(14.4)	(10.6)	(23.1)	(18.5)	(34.3)	(22.6)	(14.9)
Mean of D.V.	2693	2693	2693	2693	2693	2693	2693	2693
β/Mean	0.004	0.005	-0.004	0.012	0.004	-0.002	0.018	-0.003
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul. One-Year Lag	First	Second	Third
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State Trends				Y				
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Demographic Controls		Y	Y					
Employment Controls			Y					
States	51	51	51	51	51	51	51	51
Respondents	47170	47170	47170	47170	47170	47170	47170	47170

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents not working in or enrolled in K-12 schools, with own household children eligible for K-12 schooling: aged 5 through 19, reporting enrolled in grade school or not completed 12th grade. Date controls: month, day, and holiday. Demographic controls: gender, 10-year age group, education, number of own household children, average age of own household children among those eligible for K-12 schooling. Employment controls: employment status (present or absent), full or part time status, and enrollment status. Parent controls: own parent in household, parents' average years of education.
Outcomes: multiplied by 365.25 to represent annual hours. Childcare: activity codes 0301 (caring for and helping household children), 0302 (activities related to —'s education), and 0303 (activities related to —'s health). Childcare related to education: activity code 0302. Travel related to childcare related to education: related travel activity code 180302. Market and household production: activity codes 02 (household activities), 05 (work and work-related activities), 07 (consumer purchases), and related travel activity codes 1802, 1805, and 1807.
Regression: reform coefficients are estimated using Equation (1). Preferred specification in Column 1. Cumulative SFR are as defined in the specification; lagged SFR may have effects one year earlier than the standard variable; and the first, second, and third SFR are zero if that state had fewer SFR in the period from 2003 through 2013. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

A Online Appendix

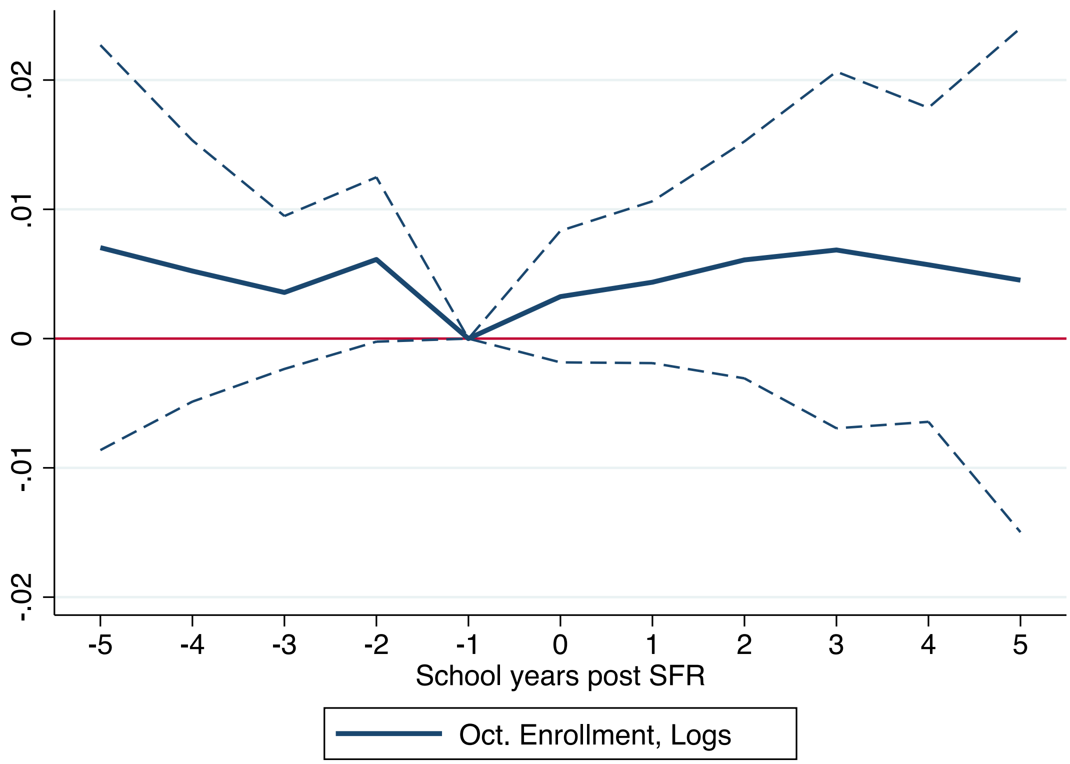
A.1 Figures

Figure A1: Reforms Do Not Crowd Out Local Revenues



Data: school finance reforms (independent variable) from 1990 through 2013 (N=86 events, in 28 states). SDFS fiscal variables (dependent variables) annually from 1990 through 2013 (N=311193 observations, from 15087 local education authorities in 50 states and Washington, D.C.; observations from 1991 are omitted because of problems with the source data).
Outcomes: divided by fall enrollment (V33) and by CPI-U (base: school fiscal year July 2014-June 2015) to represent real per-pupil values.
Regression: reform coefficients are estimated using Equation (3), a state-level dynamic event study that includes district fixed effects, year fixed effects, and state trends. Year -1 is the last pre-reform year. Standard errors are clustered by state, and 95 percent confidence intervals are displayed with dotted lines.

Figure A2: Estimates of Reform Effects Are Not Confounded by Enrollment

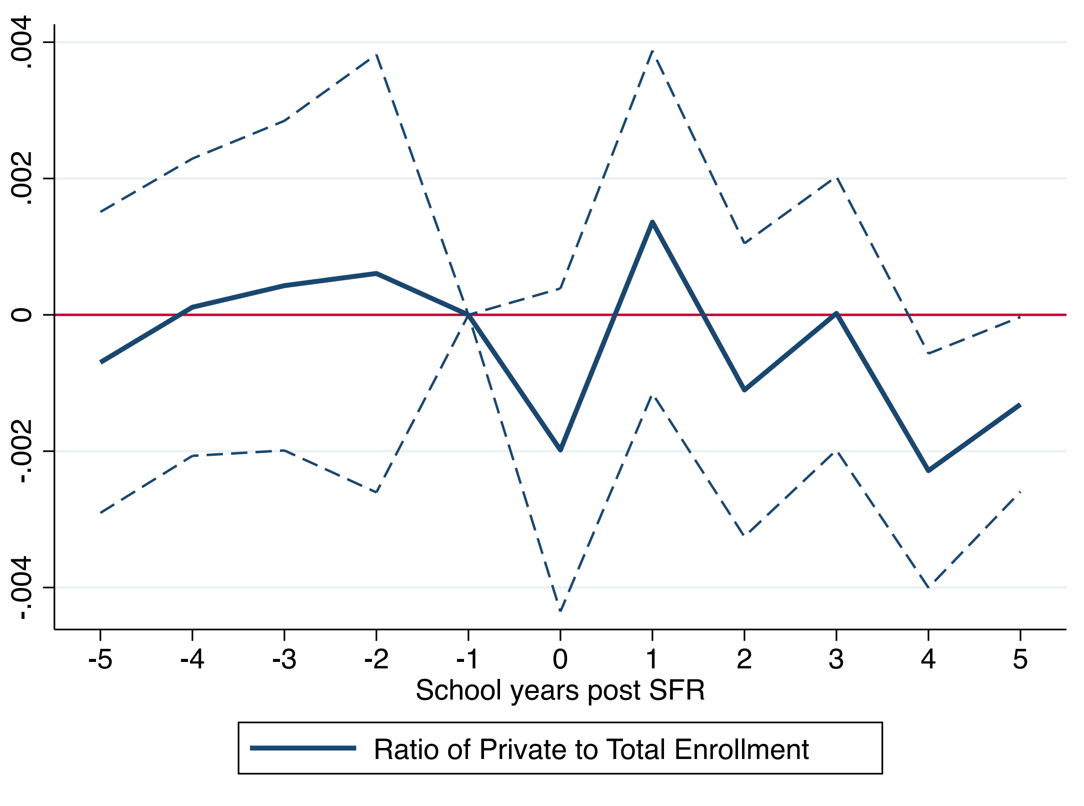


Data: school finance reforms (independent variable) from 1990 through 2013 (N=86 events, in 28 states). SDFS fiscal variables (dependent variables) annually from 1990 through 2013 (N=311193 observations, from 15087 local education authorities in 50 states and Washington, D.C.; observations from 1991 are omitted because of problems with the data extract).

Outcomes: fall enrollment (V33), in logarithms.

Regression: reform coefficients are estimated using Equation (3), a state-level dynamic event study that includes district fixed effects, year fixed effects, and state trends. Year -1 is the last pre-reform year. Standard errors are clustered by state, and 95 percent confidence intervals are displayed with dotted lines.

Figure A3: Estimates of Reform Effects Are Not Confounded by Sector Shifting



Data: school finance reforms (independent variable) from 1990 through 2013 (N=86 events, in 28 states). SDUS and PSS (dependent variables) in even years from 1990 through 2012 are aggregated up to the county level before the ratio of private school to total enrollment is taken (N=25635 observations, from 2532 counties in 50 states and Washington, D.C.).

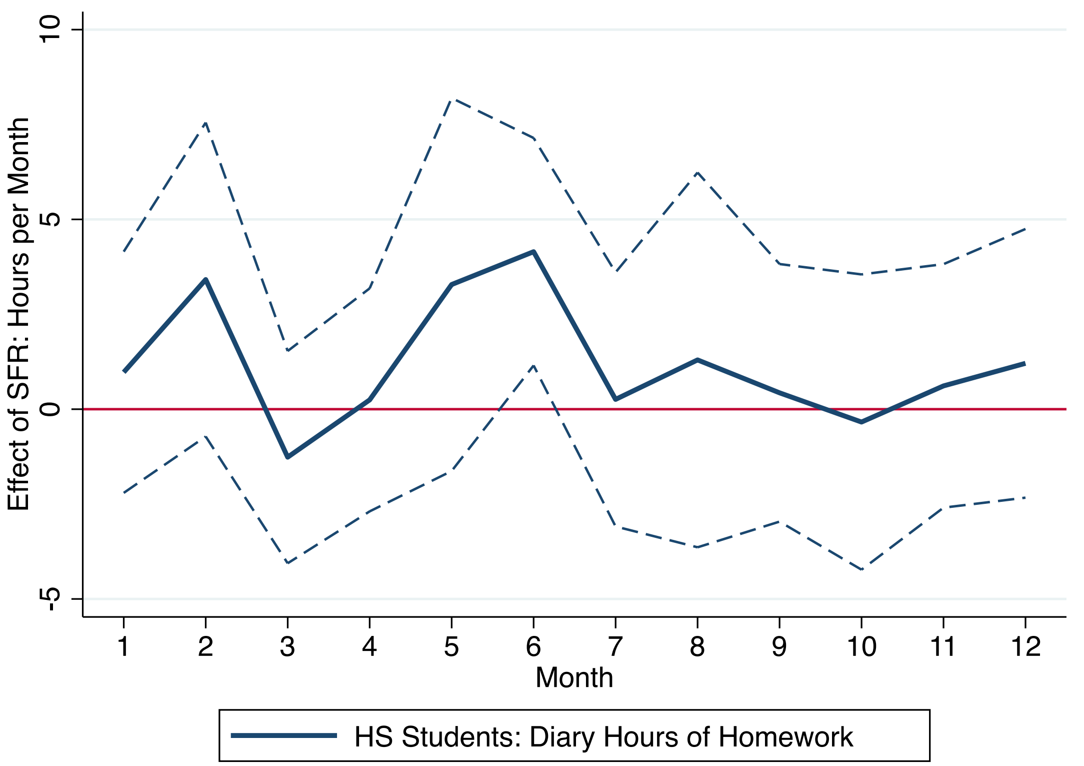
Outcomes: private enrollment (NUMSTUDS), divided by sum of private enrollment and public fall enrollment (MEMBER).

Regression: reform coefficients are estimated using Equation (3), a state-level dynamic event study that includes county fixed effects, year fixed effects, and state trends. Year -1 is the last pre-reform year. Standard errors are clustered by state, and 95 percent confidence intervals are displayed with dotted lines.

Figure A4: Children: SFR and Diary Hours of Class, by Month

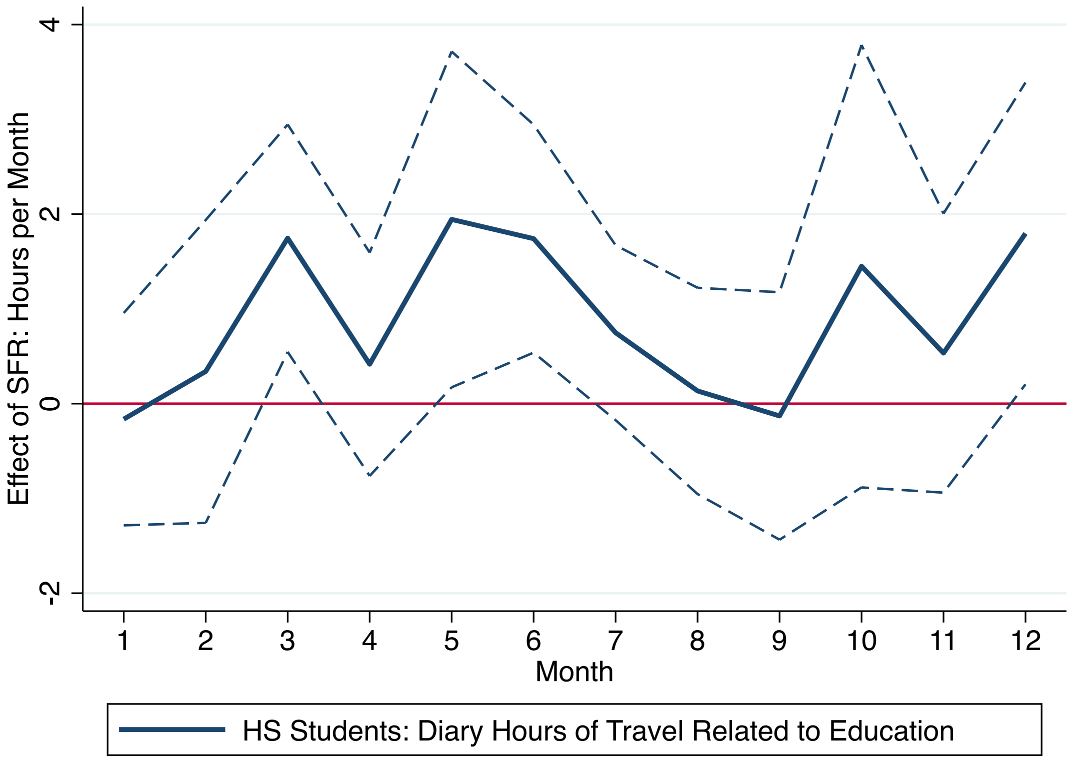
Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents eligible for K-12 schooling (N=9150, from 50 states and Washington, D.C.): aged 15 through 19, and either reporting not completed 12th grade or enrolled in high school. Date controls: month, day, and holiday. Grade controls: education, age.
Outcomes: multiplied by 365.25/12. Class: activity code 0601.
Regression: reform coefficients are estimated using Equation (1), interacting the reform indicator with month indicators. 95-percent CI are shown by dotted lines, clustered by state.

Figure A5: Children: SFR and Diary Hours of Homework, by Month



Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents eligible for K-12 schooling (N=9150, from 50 states and Washington, D.C.): aged 15 through 19, and either reporting not completed 12th grade or enrolled in high school. Date controls: month, day, and holiday. Grade controls: education, age.
Outcomes: multiplied by 365.25/12. Homework: activity code 0603.
Regression: reform coefficients are estimated using Equation (1), interacting the reform indicator with month indicators. 95-percent CI are shown by dotted lines, clustered by state.

Figure A6: Children: SFR and Diary Hours of Travel Related to Education, by Month



Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents eligible for K-12 schooling (N=9150, from 50 states and Washington, D.C.): aged 15 through 19, and either reporting not completed 12th grade or enrolled in high school. Date controls: month, day, and holiday. Grade controls: education, age. **Outcomes:** multiplied by 365.25/12. Travel Related to Education: activity code 1806. **Regression:** reform coefficients are estimated using Equation (1), interacting the reform indicator with month indicators. 95-percent CI are shown by dotted lines, clustered by state.

A.2 Tables

Table A1: Workers: SFR and K-12 Staff Share: robustness to samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	-0.00031	-0.00026	-0.0014	-0.00021	-0.00027	-0.00022	0.00026	0.00073**
	(0.00051)	(0.00052)	(0.00098)	(0.00051)	(0.00049)	(0.00056)	(0.00057)	(0.00032)
Mean of D.V.	0.062	0.062	0.062	0.057	0.060	0.064	0.059	0.059
β/Mean	-0.005	-0.004	-0.022	-0.004	-0.005	-0.003	0.004	0.012
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample Restriction		Trim Weights p5-p95	Frame Aug 05 Apr 14	Include Ever Emp.	Include Enrolled Resp.	Drop Jun-Aug	Data 1990-2015	SFR 1990-2013
States	51	51	51	51	51	51	51	51
Respondents	8893716	8893716	6388852	9637227	9360735	6614130	18047794	18047794
R^2	0.010	0.010	0.010	0.014	0.009	0.001	0.010	0.010

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS from 2003 through 2015, sample of respondents reporting employment, not enrolled. Date controls: month. Employment controls: employment status (present or absent). Samples: Column 2 winsorizes the CPS weights at the 5th and 95th percentiles of their distribution, Column 3 restricts the sample to the longest available CPS frame (PSUs) from August 2005 through April 2014, Column 4 adds workers ever employed, but never enrolled, Column 5 adds workers who reported being enrolled in school or university, Column 6 drops responses in the summer vacation months, Column 7 includes CPS from 1990 through 2015, and Column 8 includes CPS from 1990 through 2015 and reforms from 1990 through 2013 (N=86 events, in 28 states).
Outcomes: K-12 staff: 1990 industry code 842 (elementary and secondary schools).
Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A2: K-12 Staff: SFR and Age: robustness to samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	-0.060	-0.056	0.047	-0.043	-0.053	-0.065	-0.061	-0.047
	(0.075)	(0.078)	(0.11)	(0.076)	(0.077)	(0.091)	(0.20)	(0.093)
Mean of D.V.	45	45	45	45	44	45	44	44
β/Mean	-0.001	-0.001	0.001	-0.001	-0.001	-0.001	-0.001	-0.001
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample Restriction		Trim Weights p5-p95	Frame Aug 05 Apr 14	Include Ever Employed	Include Enrolled Resp.	Drop Jun-Aug	Data 1990-2015	SFR 1990-2013
States	51	51	51	51	51	51	51	51
Respondents	564861	564861	407600	567886	577964	435485	1090453	1090453
R^2	0.005	0.005	0.005	0.013	0.004	0.004	0.010	0.010

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month. Employment controls: employment status (present or absent). Samples: Column 2 winsorizes the CPS weights at the 5th and 95th percentiles of their distribution, Column 3 restricts the sample to the longest available CPS frame (PSUs) from August 2005 through April 2014, Column 4 adds workers ever employed, but never enrolled, Column 5 adds workers who reported being enrolled in school or university, Column 6 drops responses in the summer vacation months, Column 7 includes CPS from 1990 through 2015, and Column 8 includes CPS from 1990 through 2015 and reforms from 1990 through 2013 (N=86 events, in 28 states).
Outcomes: age in years.
Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A3: K-12 Staff: SFR and Age: robustness to triple difference

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	-0.060	-0.027	-0.038	-0.21*	-0.14**	-0.028	-0.14	-0.17
	(0.075)	(0.089)	(0.12)	(0.12)	(0.052)	(0.053)	(0.14)	(0.18)
K-12 Staff × SFR		-0.033	-0.022	0.15	0.078	-0.032	0.097	0.089
		(0.10)	(0.12)	(0.13)	(0.082)	(0.087)	(0.16)	(0.20)
Mean of D.V.	45	45	45	45	45	45	45	45
β/Mean	-0.001	-0.001	-0.000	0.003	0.002	-0.001	0.002	0.002
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Comparison Group		Gov't	State and Local	Non Profit	Profes. Service Industry	Educ. / Health Service	PSM Group A	PSM Group B
States	51	51	51	51	51	51	51	51
Respondents	564861	1478358	1227917	1113815	2509184	1988538	8893696	8893716
R^2	0.005	0.006	0.007	0.006	0.008	0.009	0.068	0.020

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month. Employment controls: employment status (present or absent).
Outcomes: age in years.
Regression: reform coefficient in Column 1 is estimated using Equation (1) and reform coefficients in Columns 2 through 8 are estimated using Equation (2). Comparison groups are other workers in: government, state and local government, non-profits, professional service industries, education and health service industries (a subset of professional service industries), and two groups of propensity-score matched employed workers. Group B is matched on gender, education, sector (class of worker), employment status, and earnings; and group A is additionally matched on age, usual hours worked, race, and hispanic. Both groups are re-weighted to match the distribution of propensity scores in the sample of K-12 industry workers, and propensity scores are controlled for in bins of 10 percentage points each. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A4: K-12 Staff: SFR and Weekly Earnings: robustness to samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	18.4***	18.3***	21.2***	17.0***	17.6***	23.0***	13.9***	4.98
	(3.84)	(3.78)	(5.36)	(3.52)	(3.50)	(5.61)	(4.16)	(3.57)
Mean of D.V.	909	909	910	869	898	905	889	889
β/Mean	0.020	0.020	0.023	0.020	0.020	0.025	0.016	0.006
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample Restriction		Trim Weights p5-p95	Frame Aug 05 Apr 14	Include Ever Employed	Include Enrolled Resp.	Drop Jun-Aug	Data 1990-2015	SFR 1990-2013
States	51	51	51	51	51	51	51	51
Respondents	141983	141983	102489	146291	145182	108965	273263	273263
R^2	0.033	0.033	0.035	0.139	0.031	0.034	0.035	0.035

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS ORG (months 4 and 8) from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month. Employment controls: employment status (present or absent). Samples: Column 2 winsorizes the CPS weights at the 5th and 95th percentiles of their distribution, Column 3 restricts the sample to the longest available CPS frame (PSUs) from August 2005 through April 2014, Column 4 adds workers ever employed, but never enrolled, Column 5 adds workers who reported being enrolled in school or university, Column 6 drops responses in the summer vacation months, Column 7 includes CPS from 1990 through 2015, and Column 8 includes CPS from 1990 through 2015 and reforms from 1990 through 2013 (N=86 events, in 28 states).
Outcomes: divided by CPI-U (base: school fiscal year July 2014-June 2015) to represent real values; recoded to zero if not employed as a wage or salary worker or otherwise not in earnings universe. 8.9 percent of values are missing in 1993. Weekly earnings before deductions are as reported in the CPS ORG: higher of direct report, or weekly hours multiplied by hourly wage rate.
Regression: reform coefficients are estimated using equation (1). Preferred specification in Column 1. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A5: K-12 Staff: SFR and Weekly Earnings: robustness to triple difference

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SFR	18.4***	7.23	13.1***	4.49	4.79	5.35	5.37	2.74
	(3.84)	(5.29)	(4.19)	(8.09)	(3.60)	(3.67)	(5.39)	(6.21)
K-12 Staff × SFR		11.2**	5.27	13.9	13.6***	13.0**	17.6**	18.8**
		(4.76)	(4.32)	(9.72)	(4.89)	(5.54)	(6.90)	(7.52)
Mean of D.V.	909	909	909	909	909	909	909	909
β/Mean	0.020	0.012	0.006	0.015	0.015	0.014	0.019	0.021
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Comparison Group		Gov't	State and Local	Non Profit	Profes. Service Industry	Educ. / Health Service	PSM Group A	PSM Group B
States	51	51	51	51	51	51	51	51
Respondents	141983	372261	309115	280068	631927	500585	2238473	2238477
R^2	0.033	0.044	0.034	0.027	0.016	0.018	0.066	0.072

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). CPS ORG (months 4 and 8) from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month. Employment controls: employment status (present or absent).

Outcomes: divided by CPI-U (base: school fiscal year July 2014-June 2015) to represent real values; recoded to zero if not employed as a wage or salary worker or otherwise not in earnings universe. 8.9 percent of values are missing in 1993. Weekly earnings before deductions are as reported in the CPS ORG: higher of direct report, or weekly hours multiplied by hourly wage rate.

Regression: reform coefficient in Column 1 is estimated using Equation (1) and reform coefficients in Columns 2 through 8 are estimated using Equation (2). Comparison groups are other workers in: government, state and local government, non-profits, professional service industries, education and health service industries (a subset of professional service industries), and two groups of propensity-score matched employed workers. Group B is matched on gender, education, sector (class of worker), employment status, and earnings; and group A is additionally matched on age, usual hours worked, race, and hispanic. Both groups are re-weighted to match the distribution of propensity scores in the sample of K-12 industry workers, and propensity scores are controlled for in bins of 10 percentage points each. Standard errors in parentheses, clustered by state.

* $p < .10$, ** $p < .05$, *** $p < .01$

Table A6: K-12 Staff: SFR and Diary Hours of Work: robustness to samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Diary Hours of Work								
SFR	115.6*	105.8*	67.8	134.5**	84.8*	112.4*	163.1**	117.4*
	(63.8)	(57.1)	(71.4)	(57.5)	(49.4)	(56.2)	(63.1)	(62.9)
Mean of D.V.	1688	1657	1669	1712	1675	1823	2203	1690
β/Mean	0.068	0.064	0.041	0.079	0.051	0.062	0.074	0.069
Panel B: Usual Hours of Work, for Comparison								
SFR	36.6**	30.6	76.0***	28.1	34.7**	61.2***	32.8	38.2**
	(17.9)	(19.6)	(27.9)	(16.9)	(16.6)	(22.2)	(24.1)	(17.3)
Mean of D.V.	2107	2108	2108	2148	2086	2099	2110	2107
β/Mean	0.017	0.015	0.036	0.013	0.017	0.029	0.016	0.018
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample Restriction		Trim Weights p5-p95	Frame Aug 05 Apr 14	Eligible in CPS ORG	Include Enrolled Resp.	Drop Jun-Aug	Drop W'end	Drop Low-Q.
States	51	51	51	51	51	51	51	51
Respondents (Diary)	6891	6891	4480	6258	7621	5353	3418	6802
Respondents (Usual)	6604	6604	4277	6046	7303	5126	3275	6521

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month, day, and holiday. Employment controls: employment status (present or absent) and full or part time status. Samples: Column 2 winsorizes the ATUS weights at the 5th and 95th percentiles of their distribution, Column 3 restricts the sample to the longest available CPS frame (PSUs) from August 2005 through April 2014, Column 4 drops respondents who would have not been in my ATUS sample based on their responses in their CPS month 8 interview, Column 5 adds K-12 workers who reported being enrolled in school or university, Column 6 drops responses in the summer vacation months, Column 7 drops responses on weekends, and Column 8 drops responses flagged as low-quality or incomplete by ATUS interviewers.
Outcomes: multiplied by 365.25 to represent annual hours. Diary hours of work: activity code 0501. Usual hours of work: as reported in ATUS interview, divided by 5; variable hours recoded to missing.
Regression: reform coefficients are estimated using Equation (1). Preferred specification in Column 1. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A7: K-12 Staff: SFR and Diary Hours of Work: robustness to triple difference

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Diary Hours of Work								
SFR	115.6*	-46.4	-73.4*	-38.5	25.6	19.1	48.5	63.8
	(63.8)	(29.0)	(38.4)	(47.9)	(19.1)	(30.6)	(38.1)	(39.9)
K-12 Staff × SFR		162.0**	189.0**	154.1**	90.0	96.5	60.9	48.4
		(65.4)	(76.0)	(59.7)	(70.8)	(82.7)	(70.8)	(80.8)
Mean of D.V.	1688	1688	1688	1688	1688	1688	1688	1688
β/Mean	0.068	0.130	0.151	0.123	0.072	0.077	0.049	0.039
Panel B: Usual Hours of Work, for Comparison								
SFR	36.6**	-18.0	-14.2	3.26	9.97	-2.19	-9.63	6.55
	(17.9)	(17.9)	(23.8)	(15.7)	(7.92)	(10.9)	(15.8)	(16.1)
K-12 Staff × SFR		54.6*	50.8	33.4	26.6	38.8*	45.6*	32.2
		(28.0)	(34.7)	(23.2)	(16.6)	(20.7)	(23.4)	(21.6)
Mean of D.V.	2107	2107	2107	2107	2107	2107	2107	2107
β/Mean	0.017	0.026	0.024	0.016	0.013	0.018	0.022	0.015
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y	Y	Y	Y
Comparison Group		Gov't	State and Local	Non Profit	Profes. Service Industry	Educ. / Health Service	PSM Group A	PSM Group B
States	51	51	51	51	51	51	51	51
Respondents (Diary)	6891	17752	14948	13412	45758	23496	92787	97824
Respondents (Usual)	6604	17126	14403	12845	43331	22501	92787	92787

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents reporting employment in K-12 industry, not enrolled. Date controls: month, day, and holiday. Employment controls: employment status (present or absent) and full or part time status.
Outcomes: multiplied by 365.25 to represent annual hours. Diary hours of work: activity code 0501. Usual hours of work: as reported in ATUS interview, divided by 5; variable hours recoded to missing.
Regression: reform coefficient in Column 1 is estimated using Equation (1) and reform coefficients in Columns 2 through 8 are estimated using Equation (2). Comparison groups are other workers in: government, state and local government, non-profits, professional service industries, education and health service industries (a subset of professional service industries), and two groups of propensity-score matched employed workers. Group B is matched on gender, education, sector (class of worker), employment status, and earnings; and group A is additionally matched on decadal age group, full or part time status, usual hours worked, black, and hispanic. Both groups are re-weighted to match the distribution of propensity scores in the sample of K-12 industry workers, and propensity scores are controlled for in bins of 10 percentage points each. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A8: Children: SFR and Diary Hours: robustness to samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Education								
SFR	68.0**	46.1	182.8***	65.9**	70.7*	63.9	92.5*	67.6**
	(33.4)	(33.9)	(25.4)	(31.2)	(37.9)	(44.5)	(53.5)	(32.6)
Mean of D.V.	1231	1220	1263	1267	1594	1529	1619	1234
β/Mean	0.055	0.038	0.145	0.052	0.044	0.042	0.057	0.055
Panel B: Class								
SFR	44.3	30.1	174.3***	42.8	40.9	42.4	56.0	41.7
	(34.1)	(32.0)	(27.4)	(33.8)	(36.7)	(44.2)	(52.9)	(32.6)
Mean of D.V.	956	945	981	985	1259	1199	1327	959
β/Mean	0.046	0.032	0.178	0.043	0.032	0.035	0.042	0.043
Panel C: Homework								
SFR	13.8*	9.59	2.08	12.8	16.8*	12.1	26.0*	16.3*
	(8.18)	(7.97)	(15.1)	(8.30)	(9.47)	(10.2)	(13.8)	(8.31)
Mean of D.V.	240	242	249	246	296	293	254	241
β/Mean	0.057	0.040	0.008	0.052	0.057	0.041	0.102	0.068
Panel D: Travel Related to Education								
SFR	10.9**	9.00**	38.4***	12.4***	12.0**	12.9**	16.1*	8.87*
	(5.01)	(4.00)	(7.69)	(4.25)	(5.73)	(5.98)	(8.78)	(4.76)
Mean of D.V.	74	72	73	75	92	90	99	73
β/Mean	0.148	0.125	0.523	0.166	0.131	0.144	0.162	0.121
Panel E: Market and Household Production								
SFR	-70.5*	-61.8	-192.2***	-81.6*	-80.6*	-83.6***	-73.8	-69.1
	(40.9)	(39.4)	(70.4)	(43.2)	(40.7)	(30.9)	(47.9)	(42.8)
Mean of D.V.	748	752	748	722	568	637	687	749
β/Mean	-0.094	-0.082	-0.257	-0.113	-0.142	-0.131	-0.107	-0.092
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Grade Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample Restriction		Trim Weights p5-p95	Frame Aug 05 Apr 14	Age 18 and Under	Report Enrolled in HS	Drop Jun-Aug	Drop W'end	Drop Low-Q.
States	51	51	51	51	51	51	51	51
Respondents	9150	9150	5926	8821	6790	6854	4440	9001

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents eligible for K-12 schooling: aged 15 through 19, and either reporting not completed 12th grade or enrolled in high school. Date controls: month, day, and holiday. Grade controls: education, age. Samples: Column 2 winsorizes the ATUS weights at the 5th and 95th percentiles of their distribution, Column 3 restricts the sample to the longest available CPS frame (PSUs) from August 2005 through April 2014, Column 4 drops respondents aged 19, Column 5 drops respondents who do not report being enrolled in high school either in error or because of holidays, Column 6 drops responses in the summer vacation months, Column 7 drops responses on weekends, and Column 8 drops responses flagged as low-quality or incomplete by ATUS interviewers.
Outcomes: multiplied by 365.25 to represent annual hours. Education: activity code 06. Class: activity code 0601. Homework: activity code 0603. Travel related to education: related travel code 1806. Market and household production: activity codes 02 (household activities), 05 (work and work-related activities), 07 (consumer purchases), and related travel activity codes 1802, 1805, and 1807.
Regression: reform coefficients are estimated using Equation (1). Preferred specification in Column 1. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

Table A9: Parents: SFR and Diary Hours: robustness to samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Activities with Own Household Children								
SFR	8.49	8.37	-25.7	10.3	6.56	-9.42	-12.7	8.79
	(21.1)	(20.1)	(43.7)	(23.8)	(19.9)	(16.0)	(30.1)	(19.5)
Mean of D.V.	1684	1697	1696	1684	1684	1640	1393	1685
β/Mean	0.005	0.005	-0.015	0.006	0.004	-0.006	-0.009	0.005
Panel B: Childcare with Own Household Children								
SFR	5.89	6.62	4.71	5.68	5.41	3.85	2.82	7.80
	(8.78)	(8.13)	(15.6)	(9.03)	(9.08)	(11.8)	(10.1)	(9.31)
Mean of D.V.	319	322	321	318	319	336	330	320
β/Mean	0.018	0.021	0.015	0.018	0.017	0.011	0.009	0.024
Panel C: Childcare Related to Education with Own Household Children								
SFR	-4.17	-4.30	-3.83	-4.33	-3.50	-5.69	-8.08	-0.67
	(2.93)	(2.94)	(2.81)	(3.19)	(3.37)	(4.64)	(5.91)	(9.02)
Mean of D.V.	50	50	50	50	50	50	50	50
β/Mean	-0.084	-0.087	-0.077	-0.087	-0.070	-0.114	-0.162	-0.013
Panel D: Travel Related to Childcare Related to Education with Own Household Children								
SFR	0.89	0.88	-2.09	0.95	1.16	1.05	0.81	1.20
	(1.07)	(1.06)	(2.13)	(1.04)	(1.07)	(1.48)	(1.42)	(1.09)
Mean of D.V.	35	35	41	34	35	37	39	35
β/Mean	0.026	0.025	-0.051	0.027	0.034	0.028	0.021	0.035
Panel E: Market and Household Production								
SFR	9.69	8.96	10.0	15.2	9.71	13.3	10.8	11.7
	(14.4)	(14.7)	(32.8)	(14.7)	(15.2)	(19.1)	(19.6)	(15.1)
Mean of D.V.	2693	2687	2677	2715	2690	2694	3025	2696
β/Mean	0.004	0.003	0.004	0.006	0.004	0.005	0.004	0.004
SFR Definition	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.	Cumul.
State FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Date Controls	Y	Y	Y	Y	Y	Y	Y	Y
Grade Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample Restriction		Trim Weights p5-p95	Frame Aug 05 Apr 14	Drop Enrolled in Coll.	Drop Educ. Industry	Drop Jun-Aug	Drop W'end	Drop Low-Q.
States	51	51	51	51	51	51	51	51
Respondents	47170	47170	30477	44801	46025	35414	23224	46477

Data: school finance reforms from 2003 through 2013 (N=29 events, in 19 states). ATUS from 2003 through 2015, sample of respondents not working in or enrolled in K-12 schools, with own household children eligible for K-12 schooling: aged 5 through 19, reporting enrolled in grade school or not completed 12th grade. Date controls: month, day, and holiday. Samples: Column 2 winsorizes the ATUS weights at the 5th and 95th percentiles of their distribution, Column 3 restricts the sample to the longest available CPS frame (PSUs) from August 2005 through April 2014, Column 4 drops respondents enrolled in college, Column 5 drops respondents who work in the education industry, Column 6 drops responses in the summer vacation months, Column 7 drops responses on weekends, and Column 8 drops responses flagged as low-quality or incomplete by ATUS interviewers.
Outcomes: multiplied by 365.25 to represent annual hours. Childcare: activity codes 0301 (caring for and helping household children), 0302 (activities related to —'s education), and 0303 (activities related to —'s health). Childcare related to education: activity code 0302. Travel related to childcare related to education: related travel activity code 180302. Market and household production: activity codes 02 (household activities), 05 (work and work-related activities), 07 (consumer purchases), and related travel activity codes 1802, 1805, and 1807.
Regression: reform coefficients are estimated using Equation (1). Preferred specification in Column 1. Standard errors in parentheses, clustered by state.
* $p < .10$, ** $p < .05$, *** $p < .01$

A.3 Theoretical Model

I use a model to fix ideas about why reforms might not always raise achievement, and about the true cost of school finance reforms.

A.3.1 Human capital production and achievement

Schools spend funds E to produce a stock of human capital K. The flow of human capital is given by time derivative $\dot{K}$.

The human capital production function is $\dot{K} = F(H_S, H_P, H_C)$, where inputs are hours provided by school staff, H_S, parents, H_P, and children, H_C.[21]

Schools purchase school staff hours according to the labor supply function $H_S = f(E)$: marginal wages, $\frac{dE}{dH_S}$, may increase or decrease.

Parents and children supply inputs in response to the school's supply decision, summarized by functions $H_P = g(H_S)$ and $H_C = h(H_S)$; these functions depend on complementarity or substitutability between inputs.

Achievement A is a function of human capital, $A = a(K)$.

A.3.2 Simplifying assumptions

To simplify the following discussion of measures, assume that achievement is a linear function of human capital, $A = k + aK$.

To simplify the following discussion of elasticities, building intuition for what could make achievement effects small, assume that the production function is CES in adult and child inputs, and CES in school staff and parent inputs, with potentially diminishing returns: $F(H_S, H_P, H_C) = [b(cH_S^p + (1-c)H_P^p)^{\frac{r}{p}} + (1-b)H_C^r]^{\frac{m}{r}}$, where $p, r \leq 1$ and $m > 0$.

Further, assume $f(E) = E^f$, where f is the elasticity of school staff hours to expenditure; and $g(H_S) = H_S^g$ and $h(H_S) = H_S^h$, so that g and h are the elasticities of parent and child hours to school staff hours, and fg and fh are the elasticities to expenditure.

A.3.3 Measures

Schools want to measure the efficiency of marginal spending, given by the elasticity of output to expenditure $\epsilon_{\dot{K},E} = \frac{d\dot{K}}{dE}\frac{E}{\dot{K}}$.

Is there a way to measure efficiency? Since human capital is not observed, most studies estimate the total derivative of achievement with respect to public expenditure, $\frac{dA}{dE}$. This is not a measure of efficiency: it is in units of dollars, time, and the arbitrary scale of the test. We may be able to measure efficiency, if we can eliminate units.

To address time units, it suffices to divide by the duration of exposure, or take the time derivative, to get $\frac{d\dot{A}}{dE} = \frac{d}{dt}\frac{dA}{dE} = \frac{d}{dt}a\frac{d\dot{K}}{dE}\int dt = a\frac{d\dot{K}}{dE}$. The former method relates to long-run average effects, which are commonly estimated, while the latter relates to short-run window

[21] Compared to the canonical production function in Todd and Wolpin (2003), this function distinguishes measured achievement from real output, distinguishes expenditures and supply functions from real input, leaves out current human capital which is not measurable in my setting, and includes child inputs which are (an unusual feature) measurable in my setting.

effects, which are estimated for the first time in this paper. Both methods assume constant effects. In practice, the two methods are not equivalent because effects are not constant.

To address test units, it suffices to divide by the time derivative $\dot{A} = a\dot{K}$, which is not the same as $A = k + aK$ and is not known for all tests, but is estimated for the NAEP in Klopfer (2025). This results in $\frac{d\dot{A}}{dE}\frac{E}{\dot{A}} = \frac{d\dot{K}}{dE}\frac{E}{\dot{K}}$, the desired measure of efficiency.

A.3.4 Elasticities

I estimate that elasticity $\epsilon_{\dot{K},E} = \frac{d\dot{A}}{dE}\frac{E}{\dot{A}}$ is between zero and two-fifths for all NAEP estimates in the literature, and my own long-run and short-run estimates (see Table 3 and Sections 5 and 6). Relative to business as usual, this result suggests that marginal investments due to school finance reforms are very inefficient.

Why is the elasticity so low? The simplified formula provides intuition for what explanations can be ruled out, and how:

$$\epsilon_{\dot{K},E}\big|_{E=1} = mf[b(c + (1 + c)g) + (1 - b)h]$$

Notice that if $m = f = g = h = 1$, then $\epsilon_{\dot{K},E} = 1$: with constant returns to scale, if inputs grow like expenditure, then output should grow like expenditure, too.

Unusually low returns to scale, m, or staff hours elasticities, f, could explain low output elasticities. However, it is unlikely that added days of identical inputs would exhibit severely diminishing returns to scale, and staff hours elasticities are high.

Because output depends on both child and adult inputs, child hours elasticities, h, would have to be even lower, to explain low output elasticities.

Because output depends on both parent and staff hours, it is harder for parent hours elasticities, g, to explain low output elasticities; parent hours elasticities may matter when parent hours actively offset staff hours, $g < 0$, but parent hours do not adjust. Could strong complementarity between staff and parent hours explain low output elasticities? Notice that the formula is evaluated for $E = 1$, which (i) implicitly assumes staff, parent, and child time inputs are similar, and (ii) leaves the role of the CES parameters unexamined. Allowing for large expenditures, $E \to \infty$, if the elasticity of parent hours to expenditure is less than the elasticity of school staff hours to expenditure, $g < 1$, and if school staff and parents are perfect complements, $p \to -\infty$, then $\epsilon_{\dot{K},E} = bg + (1 - b)h$, and a low parent elasticity of hours to expenditure may explain low output elasticities. However, it is more common to assume that school staff and parent hours are substitutes, not strong complements.

Measurement is the most promising candidate explanation for low achievement elasticities, which are theoretically identical to human capital elasticities. The empirical evidence on hours elasticities predicts high human capital elasticities, and both theory and empirical evidence argue against strongly diminishing returns to scale.

A.4 Empirical Model

To understand how test timing and fadeout might affect achievement effects based on school finance reform exposures, consider the empirical achievement production function,[22]

$$A_{igst} = \alpha_0 x_{gst} + \alpha_1 x_{g-1st} + \ldots + \alpha_g x_{0st} + \mu_{gs} + \epsilon_{igst}$$

where A_{igst} is achievement for student i, in grade g, school s, and year t; α_τ gives the impact of input τ grades before grade g, and $0 \leq x_{gst} \leq 1$ reports the fraction of a full year's input, due to a reform, that is delivered before testing in grade g: the fraction must be zero or one for $g' < g$, but for $g' = g$ it can be any number in between. Then, taking means by years $h = \ldots, 0, 1, 2, 3, 4, 5, \ldots$ since an event at time t we have, for example in 4th grade:

$$\begin{aligned}
\bar{A}_{4st+5} &= \alpha_0 x + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \mu_{4s} + \bar{\epsilon}_{4st} \\
\bar{A}_{4st+4} &= \alpha_0 x + \alpha_1 + \alpha_2 + \alpha_3 + \mu_{4s} + \bar{\epsilon}_{4st} \\
\bar{A}_{4st+3} &= \alpha_0 x + \alpha_1 + \alpha_2 + \mu_{4s} + \bar{\epsilon}_{4st} \\
\bar{A}_{4st+2} &= \alpha_0 x + \alpha_1 + \mu_{4s} + \bar{\epsilon}_{4st} \\
\bar{A}_{4st+1} &= \alpha_0 x + \mu_{4s} + \bar{\epsilon}_{4st} \\
\bar{A}_{4st+0} &= \mu_{4s} + \bar{\epsilon}_{4st}
\end{aligned}$$

The regression of A on years of reform exposure $\max(0, \min(h, g))$[23] and school fixed effects yields a weighted sum of difference estimators. The difference estimators themselves yield weighted averages of the input coefficients, with the first input coefficient multiplied by the fraction of same-grade input delivered before the test.

For example, consider the following difference estimators:

$$\begin{aligned}
a^4_{5,0} =& (\bar{A}_{4st+5} - \bar{A}_{4st+0})/5 = (\alpha_0 x + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)/5 \\
a^4_{1,0} =& (\bar{A}_{4st+1} - \bar{A}_{4st+0})/1 = (\alpha_0 x)/1 \\
a^4_{2,1} =& (\bar{A}_{4st+2} - \bar{A}_{4st+1})/1 = (\alpha_1)/1
\end{aligned}$$

For any dataset with a significant pre- or post-period, where most observations have no exposure or full exposure, the most common difference estimator is the long-difference estimator, equal to $a^4_{5,0}$; in such a dataset, as T grows, the regression estimate will converge to $a^4_{5,0}$. This is the basis for my fadeout calculations.

Other estimators illustrate both some possible strategies to remove bias from test timing and fadeout, and the pitfalls of those strategies:

$a^4_{1,0}$ is an unbiased estimator of α_0 if $x = 1$, which could be checked directly. If x were large enough, estimates could be scaled by $1/x$. In my setting, I find that x is closer to zero. However, even if $x = 1$, the comparison would need to be made for a large enough sample: one reason many authors do not attempt to estimate $a_{1,0}$ is that their samples and research designs are underpowered to make that comparison.

[22] See Todd and Wolpin (2003), section 2.3.2., equation (8) for a more general specification.

[23] I assume here that the analyst would cap years of exposure at the grade of testing. This is not always the case in the literature.

$a^4_{2,1}$ is an unbiased estimator of α_1, which, for small x, may be larger than $a^4_{1,0}$, $a^4_{5,0}$, and other estimates that contain the term $\alpha_0 x$, or that contain the coefficients $\alpha_2, \alpha_3, \alpha_4 < \alpha_1$. However, the same power considerations apply.

Alternatively, if we knew both input timing x and fadeout $\gamma_\tau = \alpha_\tau / \alpha_0$, we could correct all difference estimates. Since that would require dividing difference estimates by γ_τ, fadeout estimates should be precise and distinct from zero: this condition is not met, empirically. Further, γ_τ should be relevant to the level of the exposures and outcomes: fadeout is typically estimated for classroom-level exposures and student-level outcomes.

Where these conditions cannot be met, it may be more productive to focus on model-based estimates of achievement, or to focus on economic outcomes.