DOES MONEY MATTER IN EDUCATION?

Third Edition

January 2025

Bruce D. Baker
David Knight

ALBERT SHANKER INSTITUTE
1998-2023

The Albert Shanker Institute, endowed by the American Federation of Teachers and named in honor of its late president, is a nonprofit, nonpartisan organization dedicated to three themes—children's education, unions as an advocate for quality, and both civic education and freedom of association in the public life of democracies. Its mission is to generate ideas, foster candid exchanges and promote constructive policy proposals related to these issues.

The institute commissions original analyses, organizes seminars, sponsors publications and subsidizes selected projects. Its independent board of directors is composed of educators, business representatives, labor leaders, academics and public policy analysts. This document does not necessarily represent the views of the members of its board of directors.

DOES MONEY MATTER IN EDUCATION?

Third Edition

January 2025

Bruce D. Baker
David Knight

ALBERT SHANKER INSTITUTE
1998-2023

ABOUT THE AUTHORS

Bruce D. Baker is Professor and chair of the Department of Teaching and Learning at the University of Miami.

David Knight is Associate Professor of Education Finance and Policy at the University of Washington.

ACKNOWLEDGMENTS

The authors wish to acknowledge the assistance of Jordan Beal, Matthew Di Carlo, Sean Lishansky, Mary Cathryn Ricker, and Vicki Thomas.

Report design by Megan Gilmore

COPYRIGHT AND PERMISSIONS

TABLE OF CONTENTS

EXECUTIVE SUMMARY

In this report, we provide a comprehensive review of the research on the effect of K-12 school funding on student outcomes. In other words, does money matter in education?

This is the third edition of this review, with the first two editions having been published in 2012 and 2016. When those previous reports were released, the nation's schools were still in the extended wake of the 2007-09 recession. School districts in virtually all states had been hammered by cuts, with the damage being particularly severe in higher-poverty districts and those serving larger shares of Black and Hispanic students. The impact of these cuts persists even today.

This erosion of investment in public schooling was, to be sure, a result of a catastrophic recession and the collapse of the housing market that accompanied it, but the draconian cuts were also justified in part by common arguments that more money wouldn't improve schools and student outcomes. Indeed, some went as far as to argue that the cuts would be beneficial, as they would force districts to be more efficient and achieve more with less. As we showed in our first two reports, such arguments were, at best, baseless claims contradicted by the empirical evidence at the time.

Today, a full eight years since the second edition of this report, the state of the "does money matter?" debate has improved in some respects, but not in others. On the positive side, a consistent flow of recent analyses, using better data and more sophisticated methods, has confirmed and elaborated on decades of prior research on the importance of adequate and equitable funding in K-12 schools. To whatever extent the idea that "money doesn't matter" was ever credible, it is no longer.

On the other hand, this emerging consensus that money does, in fact, matter is not yet reflected in many—perhaps most—states' K-12 school finance systems and policymaking. There is also persistent confusion on many of the critical issues underlying the general "money matters" conclusion. Such confusion is understandable. The research literature on the impact of school spending, both before and after the publication of our last report, is large and complex. It includes studies of whether additional K-12 spending improves outcomes (and whether less spending hurts outcomes), but it also includes dozens of analyses of how this impact varies between locations and student subgroups, as well as studies of the impact (and cost effectiveness) of individual policies on which education dollars are or might be spent.

In this report, we provide a fair survey of this school finance research landscape, one that we hope will inform and improve debates and policy. The body of this report offers a great deal of nuanced discussion of studies that may serve in this capacity, but our primary conclusions are summarized below.

MONEY MATTERS, WHETHER IT'S GOING UP OR DOWN.

The overwhelming bulk of studies we review show that infusions of additional money into schools lead to improved student academic achievement and outcomes later in life, while a handful of studies also validate that funding cuts, resulting from major events like the 2007-09 recession, lead to a decline in student outcomes.

MONEY MATTERS, WHETHER THAT MONEY IS DRIVEN INTO ANNUAL OPERATING EXPENDITURES OR CAPITAL INVESTMENTS.

The largest share of annual operating spending in public schooling goes toward (a) the competitiveness of teacher and other school staff wages; and (b) the quantities of school staff that can be hired. In other words, it goes to paying teachers more and/or hiring more teachers. Both matter, and a high-quality public schooling system requires a "both/and approach," rather than an "either/or approach." Competitive wages are needed to maintain or improve the quality of the teacher workforce, as

such quality matters for student outcomes. Reduced class sizes and staffing ratios (including tutoring) also lead to better student outcomes in the short or long term. On the capital investment side, spending on school facilities also improves student outcomes, both directly (e.g., providing healthy and safe spaces for student learning) and indirectly (e.g., supporting teacher recruitment and retention by offering high-quality, productive workspaces). For instance, improvements to heating, ventilation, and air conditioning systems offer a relatively large return on student achievement outcomes. Generally, investments in capital have a four- to six-year lag between the commitment of new funding and measurable positive effects on students.

MONEY MATTERS MORE–AND HAS A MORE PROFOUND IMPACT–FOR CHILDREN EXPERIENCING POVERTY AND IN SCHOOL DISTRICTS AND COMMUNITIES IN WHICH STATES HAVE HISTORICALLY UNDERINVESTED.

Several studies discussed herein validate that spending more on schools and communities that have previously been deprived of resources yields greater returns on investment than spending where prior investment has been high and student need relatively lower; the difference in return on investment may be as high as 20-fold. These findings validate the importance of promoting funding progressiveness in state school finance systems, with the goal of equal educational opportunity for all.

MONEY MATTERS, REGARDLESS OF HOW CHANGES IN FUNDING COME ABOUT.

Whereas school finance legislation and litigation receive the most attention, the reality is that changes in the amount and distribution of school dollars can occur due to a variety of reasons, including:

- legislatively initiated school finance reforms;
- legislative school finance reforms in response to judicial pressure (e.g., litigation);
- large-scale economic changes (global/national recessions);
- localized economic changes (changes to taxable property wealth); and
- democratic processes (bond elections, local spending referenda, and budget votes).

Multiple causal studies discussed herein validate that, whatever the cause of substantive changes in school funding, those substantive changes matter. They influence student outcomes. Many multistate studies broadly characterize school finance reforms, often as a collection of judicial pressures and legislative responses, on balance finding that those reforms lead to positive outcomes for children. Others focus on economic fluctuations, finding that, when economic shifts lead to changes in school funding, those changes also matter for student outcomes; increases help and cuts hurt. Still others focus on local referenda leading to investment in capital infrastructure, or specific features of school funding formulas that drive additional funding to individual school districts or protect them from losses; once again, these changes affect outcomes. Regardless of the cause, the research shows that increased funding improves student outcomes and that decreased funding harms student outcomes.

In short, the evidence reviewed in this report overwhelmingly suggests that additional investment improves student outcomes, particularly for underserved students, whereas funding reductions harm those outcomes. Although the school finance literature has generally supported this conclusion for decades, an ongoing flow of studies over the past 10 or so years offers particularly compelling proof of the consistency and magnitude of the impact. This growing body of evidence has helped to foster an emerging consensus among education researchers, advocates, and other stakeholders as to the importance of adequate and equitable K-12 funding. There are, to be sure, still important outstanding questions about issues such as how much funding is enough and the most cost-effective ways to spend additional dollars. We address at least some of these questions in this report. But the "money doesn't matter" argument has largely faded from the landscape.

We acknowledge the basic reality that school funding is and always will be a highly political arena. Even the highest-quality empirical evidence must contend with practical and political constraints, particularly when the conclusions call for additional spending. That said, we hope that the review of the evidence presented in this report will serve to inform school funding debates and policymaking going forward.

INTRODUCTION

A publicly funded, open education system is a cornerstone of a democratic society. To provide equal access to high-quality instruction, school systems require basic resources to support personnel and physical infrastructure. Yet there exists no federally guaranteed fundamental right to an education in the United States. The responsibility for providing equal access to high-quality instruction falls largely to individual states.

The U.S. system of financing schools leaves much of the responsibility to states and local school districts, which, on average, collectively pay for about 90 percent of K-12 funding, with the federal government accounting for the other 10 percent. Each state's constitution speaks to either the obligations of legislatures to provide public schooling or the rights of children to have access to sufficiently funded public schools.[1] These constitutional rights have provided an avenue through which education advocates can seek additional funding: the courts. Judicial deliberations over those rights often boil down to what makes for an educated citizenry and what constitutes an adequate or aspirational outcome for each child.[2] The central underlying question in such deliberations is often whether money matters in providing children equal opportunity to achieve those outcomes. This question is routinely debated by local boards of education, state legislatures, and to a lesser extent the U.S. Congress, but also in the context of state courts.

These venues are not optimal for sorting through the complex body of empirical evidence on the impact of school funding adequacy and equity. In addition, school funding—the process for generating tax revenues and allocating those funds to schools—is inherently political, and to some extent it always will be. That said, in this report, we lay out this evidence in a manner that we hope will be accessible to advocates, journalists, and the taxpaying public.

1.0 FRAMING THE DEBATE

Before delving more deeply into the empirical literature, we must first lay some groundwork for demystifying the education dollar and where it goes. We start with an overview of the basic mechanisms by which money matters in "brick-and-mortar" schooling contexts. Schooling remains a human resource intensive industry, requiring sufficient quantities of adequately qualified adults to get the job done and adequate spaces in which to do so. There is no great mystery as to where the money goes in public schooling. As of yet, there have been no breakthrough innovations to replace the human resource intensive model of schooling with something that achieves the same or better outcomes at substantially lower spending.

We follow up with a detailed breakdown of where the education dollar goes. Finally, we discuss the forces that lead to changes in the amounts or distribution of dollars to schools, which are sometimes due to intentional choices and the preferences of policymakers, sometimes due to judicial pressure, and other times influenced by economic conditions such as recessions and periods of growth. When the flow of money changes, we can best observe whether and how those changes mattered.

1.1 THE MECHANISMS BY WHICH MONEY MATTERS ARE RELATIVELY STRAIGHTFORWARD

As shown in Figure 1, there are two main "paths" that education dollars may travel. The first (top of the figure, indicated in green) includes those dollars spent on the annual operations of schools. The bulk of this spending, which constitutes about 85 percent of total spending, goes to the salaries and wages of all school personnel, from teachers and administrators to office, custodial, food service, and transportation staff. But it also includes the dollars spent on keeping the lights on, making classrooms warm or cool enough to foster student learning, fueling up the buses to get students to and from school buildings, and buying materials, supplies, equipment, insurance, and more (Materials, Supplies, and Equipment (and other), or "MS&E & Other" in Figure 1).

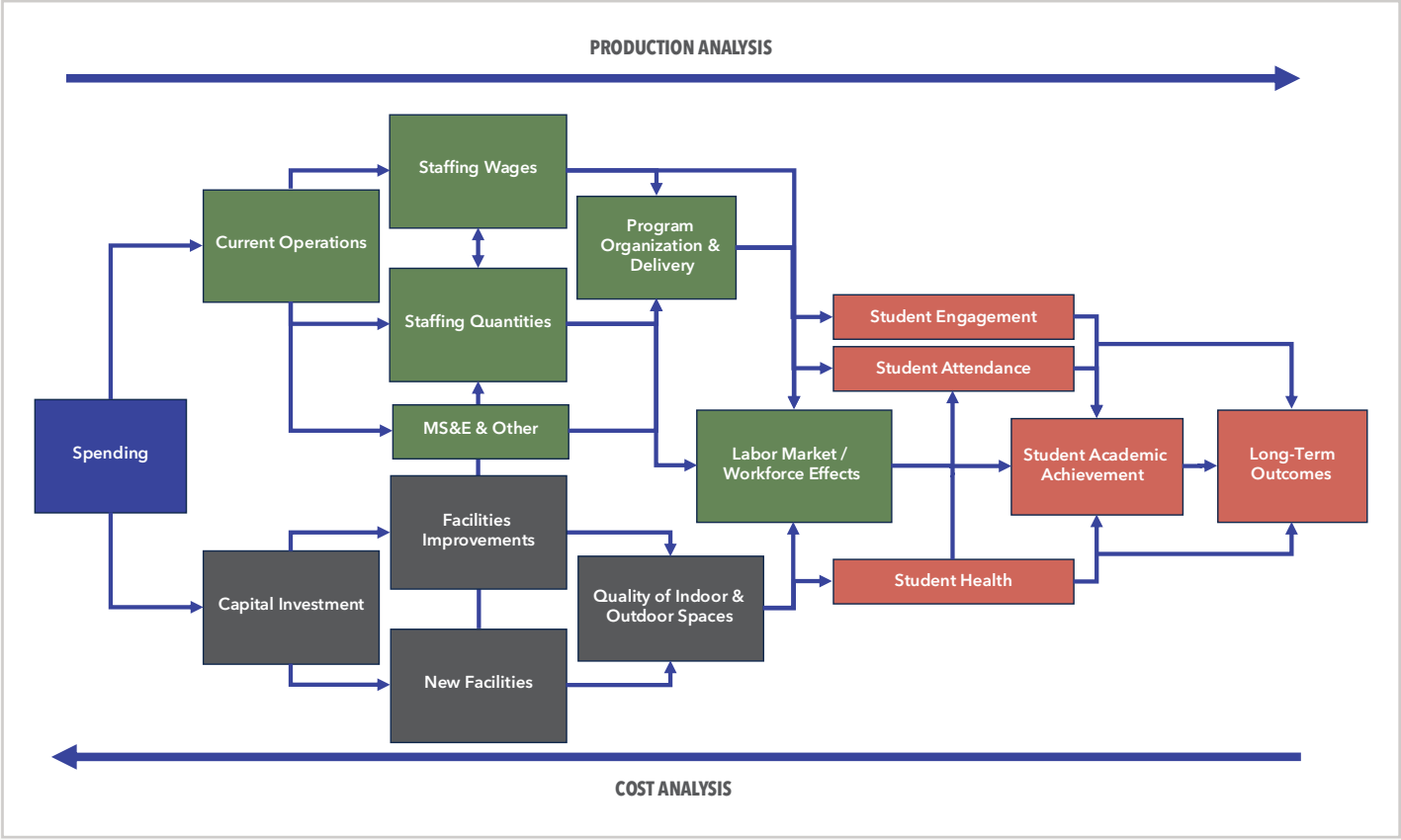
The bottom path (Capital Investment, indicated in black) involves spending on capital infrastructure for schooling—creating adequate, safe, and healthy spaces for student learning. This includes construction of new school facilities as well as improvements to existing

facilities. As we discuss in later sections of this report, these two paths—current operations and capital investment—have interactive effects on school quality. On the one hand, well-lit, safe, temperature-regulated spaces can enable student health and learning. Yet these conditions also represent good working conditions for employees, and may, for example, help in the recruitment and retention of qualified teachers. They create the spaces in which smaller classes can be provided (one cannot adopt a policy of class-size reduction without available classrooms). Investment in upgraded insulation and heating, cooling, and mechanical systems can reduce the amount of money spent to operate those systems during the year. Spending on the bottom path can free up money for the top path. As we unpack the research literature on how investments influence student outcomes, it is also important to understand the different time frames between investment of the school dollar and the effects of such spending, and how this presents challenges for researchers trying to study the impact of K-12 funding. The mechanisms by which K-12 funding is spent are generally straightforward, but studying the impact of those investments is often complicated. Schooling is a long-term process with cumulative effects. Such effects can be measured in the short term (e.g., testing outcomes after reforms are implemented) or over a longer period of time (e.g., college attendance, earnings upon entering the workforce).

If we think about a specific investment of school dollars as a treatment of sorts, the impact of that treatment will vary by the duration of exposure. If, for instance, we're looking at investments through the top path—say, hiring additional teachers to reduce class sizes—the students in lower grades will receive more of this treatment (i.e., additional years) than will their peers in higher grades, presumably generating greater improvement in short- and long-term outcomes.

FIGURE 1

How School Funding Matters



Furthermore, treatments vary in how quickly they can be expected to improve outcomes. All else being equal, class-size reduction or intensive tutoring will likely have a more immediate impact than, say, attempts to improve the quality of the teaching workforce by increasing salaries. The latter approach (salary raises) requires gradually replacing one generation of teachers with a higher-quality pool of entrants to the profession. Knowing that exposure happened, for whom, starting when, and for how long is important for evaluating the effects of any particular intervention.

Identifying the time frame of exposure to the benefits of investment in capital (the bottom path) can be still more complicated. On the surface, it seems quite simple: The treatment begins when children start attending the new or renovated facility and persists as long as they do. But the hard part from a researcher's perspective is identifying the average time lag between passing a new local levy to raise the revenue to spend on capital and the opening of the new building, or the lag between reported spending on construction and children's exposure to the benefits of that investment. As will be discussed later in this report, more recent studies of the effects of investment in capital have better accounted for and identified the average time frames from investment to exposure to outcomes.

1.2 UNDERSTANDING THE EDUCATION DOLLAR

In education research and policy, K-12 resources are typically expressed in terms of per-pupil spending. Researchers often examine how much a specific change in per-pupil spending affects student outcomes, or how much more per-pupil spending is needed to achieve a specific outcome goal.

Calculating per-pupil spending may sound easy enough, but, done correctly, it is often more complicated than it seems. Done incorrectly, it can misinform or lead to false conclusions. There are two components: (1) total school spending (the numerator); and (2) the number of pupils served by programs supported with that spending (the denominator). Available data aren't always sufficient for precisely or comprehensively matching the numerator and denominator, or for clearly aligning the treatment with the treated, including the timing issues discussed above. Attempting to precisely attach 100 percent of the dollar spent to the child served can itself be a fool's errand. Some noise and imprecision must be tolerated when evaluating complex publicly financed social systems. Such imprecision can be problematic, leading to bias when the imprecision affects some districts and children more than others.

For example, many public school districts provide services to their communities after school hours, including adult education after hours, leasing facilities to outside organizations for various activities or including students who attend virtual schools to participate in district athletic programs. These services are often important to the communities, but they do require spending money on something other than educating the children attending during school hours. If that spending remains in the numerator but is divided by school-year enrolled children in the denominator, spending per student is overstated. The same issue applies to spending on summer school programs, camps, or activities, which will be reported as an expenditure but not necessarily reflected in additional classroom time and students enrolled. This problem, moreover, arises even when these services are paid for with fees charged to participants. The money shows up as revenues and expenditures for and by the district, but does not serve school-day, school-year enrolled students. Similarly, if districts pay to send children to other schools or districts or to outside service providers, one must be able to precisely adjust both the numerator and the denominator to reflect this, or at least accept that such imprecision exists and have some understanding of how this might affect (bias) the measurement of per-pupil spending (treatment) across schools and districts and over time. This becomes even more complicated when districts pay for services, like special education or transportation, for children attending other, non-district schools (e.g., charter schools).[3]

Data on school district revenues, expenditures, and enrollments used in research typically come from one of two sources: (1) state data and financial reporting systems; or (2) the federal data collections of the U.S. Census Bureau and U.S. Department of Education (the Census Bureau's Annual Survey of School System Finances[4] and the National Center for Education Statistics' Common Core of Data[5]). Federal data sources on school finance are collected from states and harmonized to a reasonable degree, allowing for analyses of multiple states at the same

time. Despite the best efforts of federal agencies to ensure comparability, however, inconsistencies in revenue and spending data do exist from state to state. Individual state data systems, on average, do a better job tracking revenue and spending across districts within their borders in a manner that allows more precise matching of numerator and denominator (i.e., spending and number of pupils). But these state-collected data may not be suitable for analyses that include other states as well.

Either state-specific or federal data can be used to illustrate where the school dollar goes, and how much can reasonably be attributed to the current year and the potential impact on short-term outcomes of students. Using data from the Fiscal Survey of Local Governments, Public School Finance (F-33) survey:

- On average, about 85 percent of total spending goes toward current annual operations, but current annual operating spending still includes spending on non-instructional programs and activities (for others) and payments or transfers to other agencies.[6]
- About 5 to 6 percent of total spending is transferred to other institutions. In higher-poverty school districts, it's about 7.5 percent.[7]
- About 0.8 to 0.9 percent of total spending is on non-elementary and non-secondary programs (those that don't serve the current student population); it is slightly higher (roughly 1 percent) in high-poverty districts.[8]
- Capital spending constitutes about 10 to 12 percent of total spending in lower-poverty school districts but less than 8 percent in high-poverty districts.[9]

These figures help clarify a few key points. The middle two bullet points present the greatest likelihood of problems in determining who is served by these expenditures and whether they can or should be systematically removed from per-pupil spending calculations. We have also found that transfers to other institutions are also not consistently or thoroughly reported. These reporting problems are more problematic when they affect different schools differently—for example, overstating spending or spending changes in higher- versus lower--poverty school districts. See the appendix of this report for the complete breakdown.

These federal data are used in many of the multistate studies of the effects on student outcomes of increasing school spending. Money that is not spent in current-year teacher wages and benefits and instructional materials and supplies serves an important public purpose, but in a framework of linking the school dollar to student outcomes, that spending may be measured as "inefficient"—i.e., it may seem like it is not improving student outcomes. The reality, in contrast, is that these expenditures go to services (treatments) that may in fact be quite necessary or at least desired by the community, but they do not lead to changes in students' measured outcomes.

This imprecision in measurement of the treatment can create some fuzziness in the spending-to-outcomes relationship and introduce some bias across settings and over time where imprecision varies across settings and over time. These effects might reveal themselves as differences in the effects of an additional school dollar across settings or over time, and those differences may be more substantial across states where data linking students and spending are less precise than within states, because of systematic differences in how states report some spending and transfers to the federal data collection.[10]

1.3 FACTORS THAT SHIFT THE DISTRIBUTION OF DOLLARS TO SCHOOLS

One can evaluate whether money matters in two different dimensions: (1) across institutions (schools or districts) or states that spend different amounts and have different outcomes; and/or (2) over time within institutions or states when spending levels change, upward or downward. The latter is preferable from both an analytical standpoint and the standpoint of informing policy. Legislatures and courts are interested in whether increasing spending is likely to increase students' outcomes and provide more equal opportunity to achieve the desired outcome goals. This requires that researchers have good data tracking changes in spending and outcomes and other conditions that may affect those outcomes over a period of time during which real, measurable changes in spending have occurred, and which is long enough for effects on student outcomes to show up. Many of the annual collections of school district-level data in digital formats were initiated by states in the 1990s, with federal collections becoming more complete during that same time period. As

a result, only a few decades of complete annual, easily accessible fiscal data exist, and somewhat less for measured student outcomes.

Substantive changes in school funding, and changes that affect students in some settings more than others, occur for a variety of reasons. Below is a short list of particularly common and/or important factors:

- Contextual Shifts & Drift
  - Economic conditions (up or down) and economic capacity of states and local communities.[11]
  - Political ideological conditions and demographic shifts.[12]
- Policy Choices & Changes
  - Legislative reforms of school funding formulas and specific features of those formulas.[13]
    - Legislative (and executive) choices to fund those formulas (or not to do so).
  - Litigation challenging constitutional deprivation and judicial rulings (pressuring legislative reforms).[14]

As we'll discuss, many scholars focus specifically on how state high court rulings (HCRs) and/or school finance reforms (SFRs) affect school funding and student outcomes. But it is equally important to understand how the broader economy affects state and local capacity to fund public goods and services. Researchers like to study "shocks"—instant, large changes that occur at an identifiable point in time. The so-called Great Recession of 2007-09, which continues to affect school budgets even today, was a particularly severe shock, and it provided an important opportunity to evaluate (1) the extent to which school spending was affected by the recession;[15] (2) the differential extent of recessionary cuts across states;[16] (3) the differential extent of recessionary cuts across districts by types of students served;[17] and (4) the effects of those cuts on students' outcomes.[18]

Most changes due to economic conditions, in contrast, follow more subtle ebbs and flows. Similarly, across several election cycles, the political ideology of state legislatures, courts, or Congress can shift, leading to long-run changes in education spending.[19]

As such, much of the research focuses on changes in spending that result either from major overhauls or from specific features of state school funding systems that determine the funding available to schools. Researchers prefer opportunities to evaluate precise events and policies that affect school funding at specific moments in time. Not all such events are as precise as we'd like to imagine.[20]

Among the less precise or time-specific events are the aforementioned broad categories of HCRs and SFRs. Regarding the former (HCRs), state supreme court rulings on the constitutionality of existing state funding formula statutes would seem to be precisely timed and discrete events, but they are not. Such rulings typically occur as parts of sequences of separate rulings over time, and any one ruling may address only some specific claims (e.g., separating equity and adequacy concerns) or address only questions pertaining to specific provisions of states' statutes. Some HCRs overturn formula features intended to improve equity but are still counted as a "win" by plaintiffs to overturn existing policy. HCRs may also involve periods of continued court oversight across multiple years. HCRs are not simply court orders to increase funding to adequate levels or provide more funding to specific districts. The rulings don't always lead to school finance reforms (SFRs), or to real changes to school funding formulas (legislation) resulting in increases to funding.[21] Whether HCRs lead to more equitably targeted funding depends on political, economic, and demographic contexts as well.[22]

On average, however, high court rulings that lead to school finance reforms do lead to increased spending on schools, at least for some children.[23] Further, some school finance reforms occur as legislative initiative, without judicial pressure, or even with less formal or direct judicial pressure.[24]

It can be easier to identify and characterize precisely specific rulings, funding formula reforms, policy changes, and other events within a single state, rather than coming up with a common classification scheme that fits across all states. Research within states over time, or on select states where similar data and policies exist, tends to focus on what we might call State Formula Discontinuities or Local Referenda Discontinuities. State school funding formulas may be changed or tweaked in very specific ways that affect some districts differently than others, even when those districts are otherwise quite similar. For example, creating or changing criteria (e.g.,

minimum enrollment) that determine whether small district supplements are granted can exclude some districts from the supplements while including other districts of very similar size.

These features of state school finance systems allow researchers to study the effects of funding and funding changes on either side of these rather arbitrary lines of distinction—almost like randomization of treatments and controls. Similarly, states create rules governing how and whether local communities can, by referendum, adopt local tax increases (or raise taxes above a certain cap) to support capital investment or increase annual operating spending (Local Referenda Discontinuities). Some communities may narrowly pass such referenda, while others may not, creating opportunities to study the influence of these changes on students' outcomes.

2.0 HOW CAN WE DETERMINE WHETHER/HOW MONEY MATTERS?

Methods and data for evaluating whether and how school funding affects students' outcomes have evolved over time, but the highly contentious political and judicial contexts in which these issues are debated have not. Insofar as these contextual factors bear heavily on the role of empirical evidence in school finance policy, it is worth briefly discussing over a half century's debate on the "does money matter?" question, starting with the influential "Coleman Report" produced by sociologist James Coleman and his colleagues in the 1960s. Next, we address the subsequent analyses, re-analyses, and meta-analyses that continued to fuel debates through the mid-1990s, and how lawsuits in state courts over school funding equity and adequacy escalated from the 1970s to 1990s.

Finally, in this section, we address how rapidly growing national and state data sources available in electronic formats, coupled with increased computing power and new statistical methods for *causal modeling,* created opportunities for researchers to produce a flurry of recent, rigorous empirical analyses of whether, how, in what ways, and to what extent changes to school funding affect school quality and student outcomes.

2.1 THE COLEMAN LEGACY

The saga over whether money matters in American public education can be traced back to the broader question of whether *schools* matter. That is, whether schools and school quality have any influence on student achievement, educational attainment, and future earnings. The first national, large-scale quantitative analysis to explore this question was sociologist James Coleman's widely cited "Equality of Educational Opportunity" report, authorized as part of the Civil Rights Act of 1964.[25] The oft-cited conclusion of that report is:

> *"It is known that socioeconomic factors bear a strong relation to academic achievement. When these factors are statistically controlled, however, it appears that differences between schools account for only a small fraction of differences in pupil achievement" (pp. 21-22).*[26]

While the conclusion infers a relatively small role for schools in influencing student achievement, the policy implications of this finding can be interpreted in two vastly different ways:

- Interpretation 1: Family backgrounds matter a lot for student outcomes, therefore schools don't matter and further investment in schools is a waste of resources.
- Interpretation 2: Family backgrounds matter a lot for children's outcomes, therefore we must figure out how to leverage and target resources for schooling and other social services to mitigate disparities in outcomes.

The second of these interpretations is more closely aligned with the spirit of the legislation that authorized the report—i.e., that there are strong social, racial, and economic differences in children's outcomes that require a significant, sustained policy response to mitigate racial and economic inequality. Some scholars viewed the Coleman findings in this light and set out to estimate empirical models to determine the funding that might be needed to help mitigate disparities created by differences in

children's family backgrounds. Garms and Smith offered the following policy objective in 1970:

> *"Equality of educational opportunity exists when the average achievement of groups of students is roughly equal. This definition recognizes a duty of the public schools, as servants of society, to attempt to overcome environmental deficiencies that are not the fault of the individual students." (p. 305).*[27]

Garms and Smith asserted that the Coleman Report revealed the need to leverage school resources to provide for more equal educational opportunity (measured by equality of outcomes)—i.e., that family backgrounds and access to schooling resources were conflated, that schools matter, and that money (shifting the terrain to target funding according to needs, rather than inversely with respect to needs) could make the difference. Unfortunately, this work and argument was well ahead of its time, with such definitions and policy objectives not taking hold until decades later (with minimal recognition for Garms and Smith).[28]

By contrast, the first policy interpretation—that schools don't matter much, and investing more resources into them is folly—became the dominant interpretation of Coleman's report. In 1986, 20 years after the report, economist Eric Hanushek published an article that would become one of the most widely cited, yet now widely refuted, sources for the claim that money simply doesn't matter when it comes to improving school quality and student outcomes.[29] In that article (p. 1150), Hanushek notes the statistical shortcomings of the Coleman analyses, suggesting that the major contribution of the report was to set the stage for more rigorous analyses to follow, including offering the *production function* framework for evaluating the effectiveness of schooling resources. That is, estimating a statistical model to evaluate how schooling inputs (funding, teachers, class sizes, etc.), controlling for non-schooling inputs (student characteristics, schooling context), influence student outcomes. The approach was typically done at the time using data on students, school districts, states, or countries, with one or a few years of data, and regression modeling of the cross-sectional variation in spending, schooling resources, and outcomes, given student characteristics.

Hanushek provided a summary of a collection of post--Coleman studies, using data from a variety of contexts, small and large, in the United States and elsewhere. Hanushek's summary was not really a meta-analysis by modern standards, but rather a tally, or simple "vote count" of the findings of those studies, without filtering or screening of the studies based on data or methodological quality standards, and without any attempt to statistically equate findings across studies. Some of the studies found a positive relationship between spending and student outcomes, while others found no relationship or a negative one. Hanushek came to the following conclusion, which was italicized for emphasis in the original publication:

> *"There appears to be no strong or systematic relationship between school expenditures and student performance" (p. 1162).*[30]

The most direct rebuttal to this characterization of the findings of existing research (at the time) came in a series of re-analyses by University of Chicago scholars Rob Greenwald, Larry Hedges, and Richard Laine, who gathered the studies originally cited by Hanushek in 1986 and conducted meta-analyses of those that met certain quality parameters. They included studies that (a) had appeared in a refereed journal or book; (b) used U.S. data; (c) had outcome measures that were some form of academic achievement; (d) used data at the district or less aggregate level; (e) employed a model that controlled for socioeconomic characteristics, fit with longitudinal data; and (f) included data that were independent of other data included in the universe of studies considered by Hanushek. Notably, these "quality control measures" pruned a significant share of studies[31] used by Hanushek.

Specifically pertaining to aggregate per-pupil spending measures, Greenwald, Hedges, and Laine (1996) found that, among statistically significant findings, the vast majority of study findings were positive (11:1), and that most of the analyses that did not find a statistically discernible relationship between spending and outcomes still found a positive association (p. 368). They concluded:

> *"Global resource variables such as PPE [per-pupil expenditures] show strong and consistent relations with achievement. In addition, resource variables that*

> *attempt to describe the quality of teachers (teacher ability, teacher education, and teacher experience) show verystrong relations with student achievement" (p. 384).*

Digging deeper and exploring the relationship between a variety of resource and student outcome measures, Greenwald, Hedges, and Laine also came to the conclusion that "a broad range of resources were positively related to student outcomes, with 'effect sizes' large enough to suggest that moderate increases in spending may be associated with significant increases in achievement" (p. 361).[32] This finding stands in sharp contrast to Hanushek's statement of uncertainty.

Other researchers, including Wenglinsky (1996), went on to explore with greater precision the measures of financial inputs to schooling that are most strongly associated with variations in student outcomes. Largely confirming the meta-analyses of Greenwald, Hedges, and Laine, Wenglinsky's analysis found that "per-pupil expenditures for instruction and the administration of school districts are associated with achievement because both result in reduced class size, which raises achievement" (p. 221).[33]

Additionally, scholars have come to new conclusions using the original Coleman data and more up-to-date statistical techniques, finding that even Coleman's data indicate that schooling quality has significant effects on student outcomes. In one recent example, Konstantopoulos and Borman (2011) conclude:

> *"Our results also indicated that schools play meaningful roles in distributing equality or inequality of educational outcomes to females, minorities, and the disadvantaged."*[34]

In a related analysis, Borman and Dowling (2010) report:

> *"Even after statistically taking into account students' family background, a large proportion of the variation among true school means is related to differences explained by school characteristics."*[35]

In short, while family background certainly matters most, schools matter as well. Furthermore, there exist substantive differences in school quality that explain a substantial portion of the variation in student outcomes.

Studies from the late 1990s evaluated the relationship between financial resources and student outcomes, making incremental improvements to production function analyses by (a) adjusting the value of the education dollar for regional cost variation;[36] (b) testing alternative "functional forms" of the relationship between financial inputs and student outcomes; and (c) applying other statistical corrections for the measurement of inputs.[37] These studies have invariably found a positive, statistically significant (though at times small) relationship between student achievement gains and financial inputs.

These studies also, however, raised new, important issues about the complexities of attempting to identify a direct link between money and student outcomes. These difficulties include equating the value of the dollar across widely varied geographic and economic contexts, as well as accurately separating the role of expenditures from that of students' family backgrounds, which also play some role in determining local funding (e.g., via property taxation). Most of the studies included in Hanushek's review suffered from serious data and methodological limitations, which have since been addressed in more recent work (discussed below).[38]

Interest in direct dollar-to-outcomes analysis also stalled due to the imprecision of data on the financial resources available to school sites and students. Most existing financial data continue to be reported at the school district level, but resources may vary widely across schools within these districts. As a result, questions about whether money matters are often restricted to linking district-level funding with student-level outcomes, which ignores the manner in which district funds are distributed among schools. School-site spending data are increasingly available but have not generally been the subject of new production function studies. That is, few studies have as yet evaluated the relationship between school-level spending and student-level outcomes. Instead, researchers have increasingly focused on "within school" factors that are thought to influence student outcomes, including schooling resources such as class sizes and teacher characteristics, that are often more easily linked in datasets to schools and classrooms.[39]

To summarize this discussion on the evolution of the debate about whether resources matter, it is important to recognize that Hanushek's original conclusion from 1986

was merely a statement of "uncertainty" about whether a *consistent* relationship exists between spending and student outcomes—one that is big enough to be important. His conclusion was *not* that such a relationship does not exist. Nor was it a statement that schools with fewer resources are better, or that reducing funding can be an effective way to improve schools.

By the early 2000s, the cloud of uncertainty conjured by Hanushek in 1986 had largely lifted in the aftermath of the various, more rigorous studies that followed, with finance scholars using detailed datasets to examine more finely grained relationships between money and student outcomes. As we'll discuss below, the uncertainty has been replaced with an empirically grounded confidence that funding does matter.

2.2 RESEARCH AS EVIDENCE IN POLITICAL AND LEGAL CONTEXT

Note, again, that Hanushek's main conclusion from his 1986 article was simply that there existed no systematic relationship between school spending and student outcomes. That is, some studies were positive, others negative, and others inconclusive. The point was merely to cast doubt on whether new or additional investment in public schooling by state legislatures would have any positive effect, supporting legislative or judicial inaction. The 2012 book *Merchants of Doubt,* by Naomi Oreskes and Erik Conway, explains how this is a particularly effective strategy in the context of legislative or judicial debate.[40] John Perry (2010) in a review of *Merchants of Doubt* noted:

> *"Doubt is effective because belief in a hazard demands action to avert it. Action brings costs and usually loss of present benefits in exchange for speculative, uncertain and distant future gains, while inaction maintains present comforts perhaps at the expense of speculative, uncertain, and distant future hazards. A few grains of doubt easily tip the balance in favor of inaction" (p. 1541).*[41]

Debating empirical evidence from environmental, medical, or social science in political and judicial context:

- gives the misguided perception that evidence is balanced on both sides, or balanced by the political balance of state legislatures (in an almost anything goes environment); and
- leads to a body of alternative facts and misrepresentations often given equal or greater credibility than more rigorous analyses.

Legislative and judicial deliberations necessarily cater equally to both sides in any debate wherein the final conclusions are drawn by lay people who often hold political interests. They are necessary governing processes for which there are no obvious replacements, but not the most useful for evaluating the balance of rigorous empirical evidence.[42]

A related problem with seeking empirical truth in political or judicial context is that those contexts permit the submission of evidence that doesn't meet any particular standards of rigor to be explained by presumptive experts and interpreted by lay people. In judicial contexts, expert witnesses are subjected to relatively low standards for vetting expertise,[43] but once accepted as an expert, (almost) anything goes. In legislative testimony, no qualifications are applied.

As the earlier version of this report explains, a series of common strategies lacking in both empirical and intellectual rigor has been used to sow doubt over the effectiveness of schooling resources for improving student outcomes, including:

a) scatterplots of school spending (on the horizontal axis) and student test scores (on the vertical axis), without consideration of intervening factors (clouds of doubt);

b) anecdotes about states or cities that purportedly dramatically increased spending but did not see returns in student outcomes;[44] [45] [46] [47]

c) the assertion that the United States spends more than any other nation on public schooling but has much worse outcomes on international assessments;[48] and, most notably,

d) graphs showing large increases in spending on schools in the United States over time, alongside flat test scores over the same period, distorting axes to show large increases in the former, and none in the latter.

We dissect these methods in an earlier version of this report, and in a 2018 book,[49] and feel no need to revisit them in detail here. Our interest here is in moving forward to summarize the current state of more rigorous empirical research.

But it remains important to remember how context continues to shape this debate. Even the most cynical read of regression-based, cross-sectional production function studies in the 1970s and 1980s was that they found no systematic relationship between spending and outcomes (even though, again, the better studies actually did find such a relationship). No study has ever found that spending more harms children, or specifically that spending more on schools as a function of judicial orders harms children. Yet, the title of a 2006 book—*Courting Failure: How School Finance Lawsuits Exploit Judges' Good Intentions and Harm Our Children*[50]—edited by Eric Hanushek asserts as much, without ever backing up that assertion with empirical evidence.

More recently, in the face of mounting, irrefutable evidence of the positive effects of school funding increases on student outcomes, Hanushek tripled down and returned to the well of uncertainty, asserting that while many recent studies do find positive effects of increased spending on student outcomes, those affects vary widely and some are very small. As a result, he contends, we should primarily concern ourselves more with how money is spent rather than how much.[51] Others have piled on with similar arguments to be addressed in greater detail as we explore the recent empirical evidence in Section 3.0.[52]

2.3 EVOLUTION OF DATA AND RESEARCH METHODS

Over the past three decades in particular, we have experienced rapid growth in computing power and digital information. While broadly true across all aspects of life, this is particularly true of data on school spending, student outcomes, and economic conditions, with finer granularity and more consistent collection and access over time. Figure 2 illustrates major events and time periods in the financing of public schooling on the top portion of the timeline and research on whether and how money matters for student outcomes on the bottom portion.

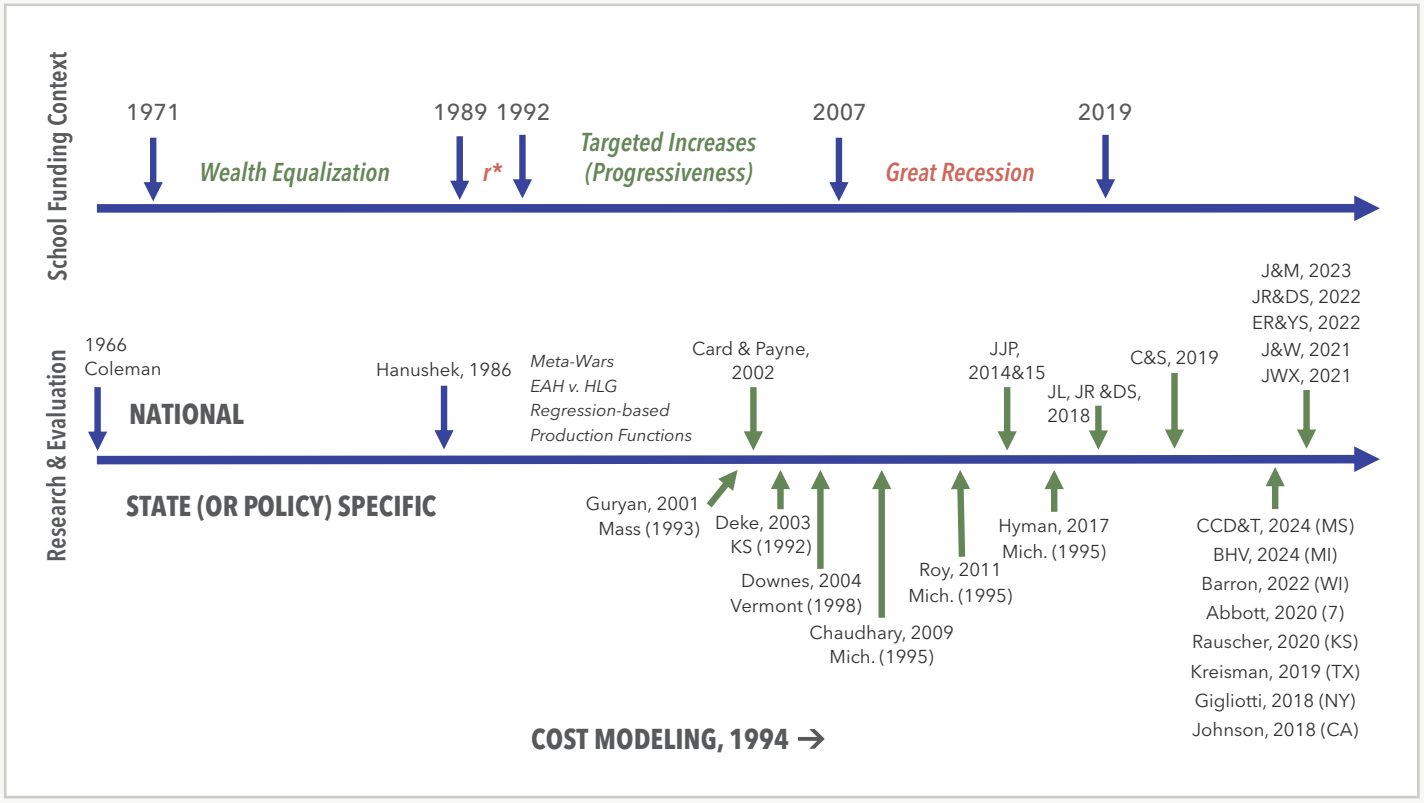
To simplify these time periods, in the years following Coleman, the 1970s and 1980s in particular, legal challenges to state school finance systems, and reforms that followed, focused primarily on reducing the connection between local taxable property wealth and school spending.[53] That is, they focused on using state aid to equalize for differences between districts in their taxable wealth and, thus, their ability to raise local revenue. In the late 1980s, interest—and legal strategy—shifted focus toward establishing an "adequate" floor of funding,[54] but there was also increased acknowledgement and acceptance that costs may vary from one child to the next (or one context to the next).[55] Pressure for reforms during the late 1980s to early 1990s also came about as a brief recession (small "r"*), coupled with a decade of property tax revolts, led to a decline in school spending and re-emerging inequalities in many states. School finance reforms and litigation from the 1990s through about 2008 led to the most significant increases in the progressiveness (funding targeted to higher poverty settings) of school funding and overall levels of funding. Then the Great Recession hit, leading to a reversal of those trends, one from which many school districts have yet to recover.

Meanwhile, state annual data collections of school finance, demographic, student achievement, and graduation rates were increasingly digitized and made either publicly downloadable or available by request. Throughout this same period, the federal government, through the U.S. Census Bureau in collaboration with the National Center for Education Statistics of the U.S. Department of Education, stepped up its annual collections of digitized school district-level fiscal data, harmonized to the extent possible to a common chart of accounts. The common charting is important, for example, in determining how pension and retirement funds are allocated to instructional staff, or for deciphering which funds are dedicated to capital investment versus annual operations, as well as which funds are used to support students in non-district schools. As noted previously, there remain irregularities across states, but having these data, however imperfect, is an important advancement. Other U.S. Census data collections, including annual collections of more dense samples of the American Community Survey, became more widely available from the 2000s forward.[56]

In terms of empirical research production, in the period from Hanushek's 1986 meta-analysis through the 1990s, most new studies continued to apply correlational and regression-based production function modeling to a limited supply of available, largely cross-sectional data. Meanwhile, the battle of meta-analyses and vote tallies of those studies continued.

FIGURE 2

Developments in Data, Modeling, and Money Matters Research



Concurrent with the data explosion from the 1990s forward, the early 2000s was a period of rapid advancement in statistical methods referred to as "causal modeling."[57] That is, statistical methods that provide opportunity to estimate a *causal* effect of a specific treatment or event on a measured outcome, compared against some counterfactual, or an otherwise similar condition under which that event or treatment did not occur. Such methods are particularly useful for studying the effects of state school finance reforms or specific policies on student outcomes across settings within states and across states over time.

While these studies begin appearing in the empirical literature in economics in the early 2000s, significant escalation occurred after 2015, as indicated in our timeline above.[58] Hence the need for this update of our prior report, which was produced in 2016, just as this new literature began to emerge. In the next section, we dive into the details of that new literature, which includes both multistate and single-state studies, applying a variety of causal modeling approaches to various data sources, and at different system levels.

While multistate or "national" analyses often garner the most attention, state-specific studies provide some considerable advantages, most notably in relation to the consistency and comparability of data and measures across settings within state boundaries and over time. State-specific studies may also more precisely identify treatments, events, and funding formula features that affect spending.[59]

3.0 DO INVESTMENTS IN SCHOOLING AFFECT CHILDREN'S OUTCOMES

In this section, we synthesize the research linking educational spending on children's outcomes. To aid the reader in making sense of the literature, we organize our synthesis according to three types of research designs. The first set of studies are meta-analyses that estimate the impacts of educational spending based on evidence drawn from a set of high-quality existing studies. Meta-analysis, done properly, represents a powerful research tool for gauging the level of consensus in the field on a particular research topic. The second section describes studies designed as multistate or national analyses of the impact of funding on outcomes. These empirical studies examine educational investments across many states, often over an extended time period, providing a wider population of causal inference. However, state-specific idiosyncrasies with both state aid programs and the finance variables in state datasets present challenges to the internal validity of these studies. Finally, in a third subsection, we focus on state-specific studies that leverage discontinuities in state funding formulas such as those described earlier. We argue these third type of studies—state-specific studies leveraging state finance datasets—have the strongest internal validity because they address many of the limitations in national cross-state studies.

3.1 META-ANALYSES

Meta-analysis is a technique in which the results of multiple studies are analyzed in order to synthesize the evidence on a particular intervention or phenomenon. The most comprehensive peer-reviewed synthesis of evidence linking increased school spending to improved student outcomes is Jackson and Mackevicius (2024), a meta-analysis published in *American Economic Journal: Applied Economics*.[60] The authors use rigorous meta-analytic methods, estimating weighted-mean effect estimates based on the precision of estimated effects from each individual study. To determine which studies to include, the authors use strict criteria. The sample of studies is limited to those that adopt a quasi-experimental method that leverages an exogenous shock in spending, and that clearly demonstrate meaningful policy-induced variation in school spending. Many other studies, in contrast, are correlational and/or examine a policy change but do not confirm the policy actually led to significant differences in school spending (with such differences necessary to confirm or deny their impact). From the hundreds of potential studies, the authors identify 31 that meet these stringent criteria. While the inclusion criteria do not include a time frame, most studies were published after 2015, after the so called "credibility revolution" that initially emerged in the early 2000s.

By combining a large number of effect estimates across many different studies, Jackson and Mackevicius (2024) can estimate more than just the overall impact of additional spending. Their meta-analysis facilitates estimation of (a) the average impact of both capital and operational spending; (b) the amount of variation or heterogeneity in effect estimates; and (c) the extent to which effects differ for different student subgroups (e.g., students from low--income households). They can also consider, from their set of studies, whether there are diminishing returns to educational investments and whether findings are systematically related to the study design or scope. Following contemporary meta-analysis techniques, overall average effect estimates are weighted by the precision of each individual estimate. The authors find that, on average, a policy increasing spending by $1,000 per-pupil for four years improves test scores by 0.032 standard deviations, which, for perspective, roughly translates to an increase in lifetime earnings of $4,550, according to estimates from Dewey et al. (2024).[61] The authors find the same funding increase of $1,000 per student over four years would also increase college-going by 2.8 percentage points.

Another key contribution of this analysis is that it shows how often these effects are likely to occur and how much the estimates vary across contexts. Jackson and Mackevicius estimate that the same investment will produce positive, statistically significant impacts on test scores over 90 percent of the time. The authors show that after harmonizing estimates across studies—setting to a common standard of the effects of an additional $1,000 per--pupil increase—estimated effects vary by 0.021 standard deviations, which is far less, by about one-fifth, than the variation in impacts from raw estimates. The authors also find that effects of capital spending are roughly similar to

those of operational spending when amortized over the life of a typical capital project. In other words, the effects of money vary across contexts and how it's spent, but not by as much as previously thought. Finally, the authors show that estimated effects are consistently larger for students who are low-income versus those who are not.

Some of the details of this meta-analysis are further explored in a separate paper by Jackson and Persico (2023).[62] They explain the benefit of harmonizing effect estimates across studies to determine the true variability in effect estimates across studies. In an article response, McGee (2023) argues that policymakers making investments in high-poverty areas are not able to do so with complete certainty of a positive impact—that there is a range of likely outcomes that may come from investing an additional $1,000 per pupil, some of which are quite small.[63] The countervailing point—that perfect certainty is not a precursor to addressing poverty—is offered in Jackson and Persico (2023b).[64] In short, the Jackson and Mackevicius (2024) meta-analysis provides clear and sufficient evidence that educational investments improve student outcomes. That is, policymakers can be relatively confident that additional investment in schooling will yield positive outcomes of important magnitude.

In a different meta-analysis by Handel and Hanushek (2023), presented in a recent working paper, the authors (including the same Hanushek discussed above) argue that estimates of the effect of additional spending range in magnitude by so much that drawing strong generalizations is difficult.[65] In conducting their meta-analysis, the authors limit the sample of studies to those with the strongest, causal research designs, using a specific set of search criteria over a specified time. Like Jackson and Mackevicius (2024), the authors weigh the estimates of each study by the level of precision using random effects estimators, recognizing that outlier estimates are often estimated with less precision and should therefore be given less weight in the calculation of both overall average impact estimates and the variation in impact estimates across studies. The authors also harmonize dollar amounts across studies, adjusting for both inflation and relative dollar amounts, converting all effect estimates to that of a 10 percent increase in funding.

With these adjustments, Handel and Hanushek (2023) conclude that money clearly matters for both student achievement and attainment: 14 of the 16 studies linking spending to test scores identify positive effects and all 18 of the included studies examining educational attainment show positive impacts. But the authors find that the magnitude of these effects varies widely across studies. They estimate that 78 percent of the variability in results across studies reflects true differences in effect estimates, suggesting the effect of K-12 investment varies widely across context. If the true impacts of educational investments vary this substantially, then policymakers are afforded little certainty about their expected impacts and may be more hesitant to invest taxpayer money into K-12 education. Jackson and Mackevicius (2024) also estimate the degree of effect heterogeneity across studies but reach different conclusions. The authors find that funding increases will produce modest impacts over 90 percent of the time, implying not only that "money matters" but that K-12 school spending represents a reliable investment of public funds. That is, that policymakers can be relatively confident that additional investment in schooling will yield positive outcomes.

3.2 EMPIRICAL STUDIES OF MULTIPLE STATES

The meta-analyses described above synthesize high-quality empirical studies that use data on educational investment and student outcomes. To further understand this body of research, including the recent advancements in the field pertaining to data and causal research design, we next describe some of the empirical studies upon which the meta-analyses discussed earlier are based. We start with national studies or studies of multiple states.

Building on earlier work by Card and Payne (2002), Jackson et al. (2016) was the first of a recent wave of national studies estimating the impact on student outcomes from educational investments made through court-mandated school finance reforms.[66] Researchers consider these school finance reforms to be "exogenous policy shock" because they stem from state supreme court rulings that require the state legislatures to pass laws to increase K-12 funding. The method is powerful because much of the variation in school resources that U.S. students experience is across states, not within, so leveraging statewide differences in resource levels over time is logical. The authors of these studies track the timing of individual state supreme court decisions and the state school finance reforms that

those court decisions mandated. When the new funding comes in, arguably from an "exogenous increase," researchers can study changes in student outcomes.

Jackson and colleagues show that school districts in states that implemented court-mandated school finance reforms have experienced significant increases in school funding compared to districts in states that did not have their systems ruled unconstitutional by courts. A key feature of their research design is that it exploits the randomness of the court decision and school finance reforms and the extended period in which children are "exposed to treatment" (i.e., attending better resourced schools). Many of the earlier studies were correlational, comparing outcomes in high- versus low-spending districts each year, or examining whether a spending increase in a given year might lead to an increase in test scores that same year. This approach, which in recent years has given way to more rigorous methods using better data and computing power, is subject to bias. For example, states may target more funding to lower-achieving districts, making the effect of higher funding look negative. Conversely, in other states, wealthier districts spend more and have higher test scores, making the effect of additional funding look positive. Researchers are rarely able to randomly assign large amounts of additional funding to schools and instead look for large increases that happen because of external, semi-random policy shocks (rather than decisions within the system).

Jackson et al. (2016) use the timing of state court decisions, and show that such decisions lead to meaningful differences in spending across states. These "exogenous" increases in spending more closely resemble the randomness of an experiment compared to the older regression studies. Moreover, the study examines students for a full 12 years of public education and also looks at outcomes into adulthood. This extended period acknowledges that the full impact of educational investments is cumulative and may take time to emerge. The authors find that students exposed to a 10 percent increase in per-pupil spending for their full time in K-12 public schools complete 0.31 more years of education, earn 7 percent higher wages, and experience a 3.2 percentage point reduction in the incidence of adult poverty. The authors also find that the effects are especially pronounced for children from low-income families.

The Jackson et al. study's findings were a major breakthrough, but several other studies have reached similar conclusions using similar research designs with different datasets and time periods (e.g., Abbott et al., 2020;[67] Miller, 2018;[68] Rothstein and Schanzenbach, 2022[69]). Candelaria and Shores (2019), for instance, examine court-mandated school finance reforms that took place from 1989 to 2010.[70] They find that seven years after a reform, districts serving higher percentages of low-income students experienced approximately a 12 percent increase in spending, and over the same period, those districts experienced a roughly 7-12 percentage point increase in graduation rates. Lafortune et al. (2018) find similarly positive impacts, focusing on test scores from the National Assessment of Educational Progress from 1990 to 2011.[71] Finally, a follow-up study by Rothstein and Schanzenbach (2022) draws on data from the U.S. Census American Community Survey to show that court-mandated reforms increase high school completion and college-going, and that effects are particularly pronounced among Black students and women.[72]

Several national studies further elaborate on factors driving these results. Rauscher and Shen (2022) show that investments targeted to more impoverished areas have larger impacts.[73] The authors aggregate the U.S. Census/NCES F-33 finance survey to the county level and use data (also county level) from the Stanford Education Data Archive (SEDA) and the National Vital Statistics System, the latter of which includes information about birthweight and child poverty. The authors find that educational investments have the largest impacts on test scores in counties with high rates of low-birthweight newborns and counties with lower prior investment. They note "the achievement return in low-birth weight counties is over 20 times larger than that in high-birth weight counties." Other studies that shed light on factors driving positive impacts of state school finance reforms examine how funds were spent and the political context. Bruner et al. (2020) find school finance reforms passed in states with stronger teacher unions tended to lead to larger increases in spending, with a greater share of funds devoted to teacher compensation levels, rather than to hiring new teachers or property tax relief.[74] Increased spending on teacher salaries led to improved student test score achievement, and positive impacts were driven by states with stronger teacher unions.

The majority of multistate studies, including all of those discussed above, examine funding increases that stem from a court-mandated legislative reform. Because the reforms are court-mandated, researchers consider the funding increase to be "exogenous" or unrelated to other factors that could influence student outcomes, such as property wealth, income, or education level. Some multistate studies have used other policy changes or formula discontinuities to assess the causal impact of spending increases on student outcomes. Two recent studies use this approach to estimate the effects of the federal pandemic stimulus funding provided through the Elementary and Secondary School Emergency Relief Fund.[75] The studies reached similar conclusions, with one finding that each $1,000 of federal stimulus funds a district received per student increased achievement by about 0.03 grade equivalents, with larger and more consistent effects for math and in high-poverty communities. Researchers estimated that the stimulus funds helped support academic recovery, but that an additional $7,000 to $13,000 per student of further investment would be needed to restore academic learning trajectories back to pre-pandemic levels.

In other cases, multistate studies have leveraged exogenous changes in property values, tax elections, or economic conditions to help estimate causal effects of school spending. Miller (2018) collects property value and tax data from over 7,000 districts in 24 states.[76] He observes that state-level changes in property values generate different spending increases for school districts based on differences in state finance formulas. Using graduation and test score data from the Common Core of Data and SEDA, respectively, he finds that "a 10 percent increase in spending improves graduation rates by 2.1 to 4.4 percentage points and student test scores by 0.05 to 0.09 standard deviations" (p. 1). Abott et al. use tax election results from multiple states, using the same graduation and test score data as other studies.[77] They find that a $1,000 per-pupil increase in spending caused an increase in test scores of 0.15 standard deviations and graduation rates by 9 percentage points. Like Rauscher and Shen (2022), effects are driven by districts below the median in spending per pupil.[78]

Finally, Jackson et al. (2021) examine school budget cuts related to the Great Recession—i.e., the impact of spending declines rather than increases.[79] The authors compare cohorts of students exposed to varying degrees of budget cuts, depending on the timing of their school entrance and the state in which they attend school. The exogenous shock to school spending allows the authors to estimate the causal effects of budget cuts. They find that a $1,000 reduction in per-pupil spending reduces test scores by 0.039 standard deviations and college-going rates by 1.24 percentage points (about 2.5 percent).

To summarize, whether funding increases are induced by judicial pressure, legislative initiative, legislative response to judicial pressure, large-scale economic changes (global/national recessions), localized economic changes (changes to taxable property wealth), or democratic process (bond elections), increased investment in operational spending or capital investment matters. While the manner in which the funds are spent will influence the magnitude of effects (that is, better uses of funds will, of course, yield better results), the literature shows that, on average, increases in educational spending improve outcomes for students.

A key benefit of the national/multistate studies discussed in this section is their broad generalizability. The population of causal inference for any studies represents a wider population with similar characteristics as that of the study's sample. Thus, by estimating impacts across many states and over a substantial time period, the results more likely apply to a larger number of districts and are therefore more generalizable.

3.3 STATE-SPECIFIC STUDIES

We next synthesize studies that examine reforms within a specific state. As discussed earlier, multistate studies are limited by their need to identify funding increases driven by "exogenous policy shocks" that take place in different states at different times. As noted above, multistate studies are beneficial because much of the variation in school resources that U.S. students experience is across states, not within. However, this method typically requires use of a national school finance dataset with variables that must be harmonized across contexts in ways that do not always align with state policies. Moreover, high-court rulings and legislative responses to those rulings may have nuances that differ across states, making it difficult to compare the same per-pupil dollar increases in different states.

State-specific studies provide somewhat more precision in isolating the effects of specific policy interventions and the types of state policy interventions that lead to sub-

stantive changes in funding, affecting student outcomes. State-specific studies also ensure more consistency of measurement of both schooling inputs and outcomes across schools and districts and over time. Over two decades ago, Figlio (2004) argued that the influence of state school finance reforms on student outcomes is perhaps better measured *within* states over time, explaining that national studies confront problems that include the enormous diversity in the nature of state aid reform plans and the paucity of national student performance data.[80] Accordingly, more recent peer-reviewed studies of state school finance reforms have applied longitudinal analyses within specific states. And several such studies provide compelling evidence of the potential positive effects of school finance reforms.

In the subsections below, we group studies according to the type of policy that led to the funding or spending increase, specifically (a) local elections or hold harmless policies; (b) small district and block grants; and (c) significant state finance reforms. Throughout these three subsections, we highlight the value of state-specific studies for modeling the policy change and measuring student outcomes. In a final subsection, we address the question of diminishing returns, highlighting multiple studies that identify consistent and positive linear effects of funding on student outcomes as levels of per-pupil investment increase.

3.3.1 | SPECIFIC POLICIES: LOCAL ELECTIONS AND ENROLLMENT HOLD HARMLESS

Local property tax levy elections exist in part to provide voters with some autonomy over spending levels in their local schools. While this feature of U.S. education policy provides some measure of local control, a foundational principle of the U.S. education system, it can create challenges for school districts. Whether a levy election fails or passes can have significant consequences for resource levels in schools and for educational opportunity for students and families.

A recent study published in *American Economic Journal: Economic Policy* compared districts in Wisconsin that narrowly passed their local levy election, to districts in which such elections narrowly failed. By comparing districts in which levies narrowly passed and failed, the authors can compare districts with similar characteristics but very different resource levels. The analysis shows that increased operational spending (passing levies) had substantial positive impacts on student test scores, high school dropout rates, and postsecondary enrollment. The analysis is unique from multistate studies because it takes into account several state-specific nuances in policy and data. For example, the author uses state agency data to examine in detail how election winners allocated new funds, providing insights into the mechanisms through which additional school spending affects student outcomes. Districts that invested more in salary increases showed greater gains relative to those hiring additional staff members. Identifying those mechanisms is difficult with multistate studies that draw on national datasets.

Along similar lines, Gigliotti and Sorensen (2022) examine New York's enrollment "hold harmless" provisions, which provide additional state aid for districts experiencing declining enrollment. That study finds test score effects similar to multistate studies discussed above: 0.047 SD in math and 0.042 in English for each $1,000 in additional per-pupil spending. As the authors explain, their use of state-specific test score data better establishes the links between educational investments, state-specific learning standards, and learning assessments that measure progress toward those standards.

3.3.2 | SMALL DISTRICTS AND BLOCK GRANTS

Many states have funding policies for smaller districts, recognizing the challenges of operating within diseconomies of scale, where districts have few students over which to spread the fixed costs such as central office administration and some capital investments. In Texas, districts with fewer than 1,600 students receive an extra per-student amount that prorates upward for smaller districts. Kreisman and Steinberg (2019) compare districts directly above and below the enrollment threshold to receive additional funds, and they find that each $1,000 per pupil in foundation funding increased math test scores by 0.08 SD and reading scores by 0.10 SD. Additional funds also led to increased graduation rates and college enrollment, while reducing dropout rates.

In Kansas, the state transitioned to a block grant that funded districts in 2015-16 and 2016-17 based on 2014-15 enrollment levels. By freezing revenue amounts regardless of enrollment changes, the policy provided additional funds to districts with declining enrollment, but fewer per-pupil funds to growing districts. Drawing on Kansas Department of Education data, Rauscher (2020)

found that math and reading test scores increased among districts benefiting from additional funds. She compares these results to an across-state analysis based on test score data from the Stanford Education Data Archive. The across-state analysis showed that the overall funding reduction in Kansas from 2008 to 2014 significantly harmed districts in Kansas in terms of math and reaching achievement. While the across-state estimates are less precise, they further substantiate the important role resources played during this period of fiscal uncertainty in Kansas school systems.[81]

3.3.3 | STATE SCHOOL FINANCE REFORM STUDIES

Another set of state-specific studies examine the impacts of "whole system" (overhauling the state general aid formula) finance reforms. For example, Michigan's Proposal A, passed in 1994, substantially altered the distribution of funding, narrowing gaps between especially high-spending wealthier school districts and lower-spending districts. Several studies examine the effects of these funding changes. Roy (2011) found significant positive effects on student performance among the lowest-spending districts, consistent with several prior studies.[82] Most recently, Hyman (2017) found that students exposed to a $1,000 increase in per-pupil spending (representing a 10 percent increase) were 3.0 percentage points more likely to enroll in college and 2.3 points more likely to earn a postsecondary degree, a 7 percent and 11 percent increase, respectively.[83] Use of student-level panel administrative data allowed the author to track individual students across districts, better measuring exposure to additional resources, especially for students who move between districts. The study is also unique in its linking of individual data to National Student Clearinghouse postsecondary enrollment and completion data, allowing for the assessment of long-term outcomes.[84] Most recently, Barron, Hyman, and Vasquez (2024) revealed the broader, longer-term social benefits of Proposal A, including crime reduction.[85]

Around the same time Michigan passed Proposal A, Massachusetts enacted its Education Reform Act of 1993. Three studies of this reform find similar results.[86] The first, a non-peer-reviewed report by Downes, Zabel, and Ansel (2009) explored the influence on student outcomes of accountability reforms and changes to school spending. They found that the 1993 reform act increased funding in higher-poverty school districts and that statewide income-based achievement gaps narrowed following passage of the law. The second study, an NBER working paper by Guryan (2001), focused specifically on the redistribution of spending resulting from changes to the state school finance formula. Guryan found that "increases in per-pupil spending led to significant increases in math, reading, science, and social studies test scores for 4th- and 8th-grade students," estimating that a "$1,000 increase in per-pupil spending leads to about a third to a half of a standard-deviation increase in average test scores" (p. 1). The most recent of the three, published in 2014 in the *Journal of Education Finance,* found that "changes in the state education aid following the education reform resulted in significantly higher student performance" (p. 297).

Researchers have also identified positive impacts resulting from major state finance reforms in Vermont and Kansas, around the same time period (the early 1990s).[87] Downes (2004) studied reforms in Vermont associated with Act 60, finding that the law "dramatically reduced dispersion in education spending and has done this by weakening the link between spending and property wealth" and that "student performance has become more equal in the post-Act 60 period" (p. 312).[88] Deke (2003) evaluated "leveling up" of funding for very low-spending districts in Kansas, following a 1992 lower court threat to overturn the funding formula (without formal ruling to that effect). The study found that a 20 percent increase in spending was associated with a 5 percent increase in the likelihood of students going on to postsecondary education (p. 275).[89]

California implemented the Local Control Funding Formula (LCFF) beginning in 2013-14, a major system overhaul that increased overall state funding but also removed many of the state's categorical funding programs and replaced them with a weighted student funding model. Under this approach, districts receive a base per-student allotment, and funding is then adjusted based on student weights that correspond to additional cost needs. Students who are low-income, multilanguage learners, or classified as foster youth generate an additional 20 percent in funding, and concentration grants provide additional funding for districts with especially high percentages of these students. Analyses of the impact of LCFF are considerably important for the field because the design of the policy closely aligns with what research suggests are best practice approaches for effective state school finance policies, including a progressive allocation of funds across

school districts based on student need, local autonomy to make spending decisions, and an accompanying accountability plan that has built in local flexibility.

Several studies estimate the causal impact of LCFF on student outcomes, finding that the increased investment both increased test scores and narrowed economic and racial differences in achievement by reducing funding disparities across school districts.[90] By leveraging state agency data, the studies also show LCFF reduced the probability of grade repetition, increased high school graduation rates and rates of college readiness, and decreased suspensions and expulsions. Use of detailed state administrative data also allows for more closer analysis of how funds are used. Johnson and Tanner (2018), for example, found that "LCFF-induced increases in district revenue led to a significant reduction in the average school-level student-to-teacher ratio and led to significant increases in average teacher salaries and instructional expenditures."

Some studies of LCFF implementation highlight challenges, including the fact that within-district spending was not always as progressive with respect to student need as intended by the law, and that new funding sometimes required new hiring of inexperienced teachers.[91] Overall, the studies of LCFF provide valuable evidence of the power of educational investments paired with local flexibility and accountability. Importantly, state-specific analyses have provided insights into the mechanisms driving positive results.

Lastly, a recent set of studies examines funding reforms in Washington state that followed a 2012 state Supreme Court decision. The funding increases led to teacher salary increases, which in turn improved retention. However, salary increases disproportionately benefited more veteran teachers and those working in lower-poverty school settings, and one study found only modest impacts of these spending increases on student outcomes.[92] Causal studies that do not find significant impacts of funding on resources are often those focused on policies that mandate funds be used for specific purposes, such as computers,[93] or those in which the time horizon is potentially too short to see positive impacts, especially in the case of capital investments.

3.4 INTERPRETING HETEROGENEITY AND DIMINISHING RETURNS?

Given the resounding finding that educational investments positively impact students across contexts and policy designs, policymakers naturally wonder whether returns diminish beyond some level of spending. Researchers have taken up this question using a variety of analytic techniques. A few studies identify larger effects for districts with lower initial spending, perhaps suggesting that returns to investments diminish at higher spending levels.[94] However, Jackson and Mackevicius (2024), in the meta-analysis discussed above, examine this issue statistically, synthesizing impact estimates from a large number of rigorous studies, including contexts where funding increases were implemented on top of already high spending levels. The authors find "little statistical evidence that school spending exhibits diminishing returns" (p. 3).[95] Moreover, while funding progressiveness increased up until the Great Recession, high-poverty districts suffered larger cuts to funding after the recession and those cuts led to lower student outcomes. There is little or no evidence that states have "maxed out" spending for student outcomes, especially for children in high-poverty communities. Rather, since the 2007-09 recession, state legislatures retreated and saw the damages of that retreat for their public school students.[96]

But every state, even the most robustly funded, has less well-funded districts that also tend to be those serving higher-need student populations. In fact, our own work shows that the most robustly funded states overall often have the greatest disparities in equal opportunity.[97]

As noted above, state-specific studies are useful for exploring how additional dollars are spent. Studies of California's LCFF, for example, show that increased instructional spending was associated with improved outcomes. While increased spending on instruction, instead of student support and administration, is often seen as a positive, studies show non-instructional resources also provide important benefits for schools.[98]

One of the most consistent findings across multistate studies and state-specific studies is that effects are larger for students classified as lower income or who have previously struggled academically. Jackson and Mackevicius (2024) highlight this finding in their meta-analysis, showing that the probability of producing a positive impact

through educational spending is greater when new funds are targeted to higher-poverty settings. The finding has led to the conclusion that more progressive or equitable policy reforms are also more efficient, since targeting funding to high-poverty settings produces greater benefits to society.

4.0 WHAT SPECIFIC INVESTMENTS MATTER?

As we explained in Figure 1 of this report, how the education dollar is spent is reasonably well understood and tracked. Figure 1 laid out two general paths—which tend to be discrete (operational spending and capital spending are largely kept separate) by nature of public finance (tax) policy and government accounting: (1) funds for current fiscal year operations; and (2) funds raised to pay down long--term debt on major capital projects. As we explained earlier, funds for current operations tend to be spent primarily on school staff, including teachers, administrators, and other school staff (librarians, counselors, clerical, custodial, etc.). The primary budgetary tradeoff to be made within staffing planning is that between the wages or salaries that can be paid to school staff and the quantities that can be hired. That is, put simply, the tradeoff is between having smaller class sizes or higher pay for teachers—where higher pay should lead to a higher-quality teacher workforce. But this decision may be constrained by available classroom spaces (the capital side). We do, however, know from some of the research discussed in the previous section that funding increases that were shown to improve student outcomes went primarily to increased wages and/or smaller class sizes as well as more instructional time.[99] That is, money matters because it pays for educational programs and resources that matter, and vice versa—educational programs and resources that matter cost money.

In this section, we review the empirical literature on whether and to what extent:

a) teacher compensation matters for improving school quality and student outcomes;

b) smaller classes or increased ratios of adults to children (instructors, tutors, and other supports) matter for improving school quality and student outcomes; and

c) investment in new or improved physical spaces for learning matters for improving school quality and student outcomes.

We also discuss the interplay between these investments.

4.1 DOES "HOW YOU SPEND MONEY" MATTER MORE THAN HOW MUCH YOU SPEND?

While the assertion that "how money is spent is important" is certainly valid, we cannot reasonably make the leap to assert that how money is spent is necessarily more important than how much money is available. How money is spent matters, but if you don't have it, you can't spend it. Opponents of increasing school funding have argued repeatedly, including in the context of state courts, that school districts have more than enough money to do what needs to be done—to provide constitutionally adequate education—if they just used their money more efficiently.

Some have even gone as far as arguing that cutting funding would not compromise districts' ability to serve their students and might actually force them to become more efficient, or that school districts serving low-income and minority children should simply spend smarter than neighboring, more affluent districts that have substantially more resources.

These kinds of arguments are, as we've shown, strongly contradicted by the evidence. They are also dangerous. As a panel of judges in Kansas noted so eloquently: "Simply, school opportunities do not repeat themselves and when the opportunity for a formal education passes, then for most, it is most likely gone. We all know that the struggle for an income very often—too often—overcomes the time needed to prepare intellectually for a better one. If the position advanced here is the State's full position, it is experimenting with our children which have no recourse from a failure of the experiment."[100]

There are legitimate questions regarding how K-12 dollars should be spent, but debates on these questions are sometimes oversimplified and/or misleading. For instance, there is a long-running debate about whether money is better spent on higher wages for teachers to recruit better teachers or on class-size reduction (teacher quality vs. teacher quantity). A common argument on the

teacher quality side is that a good teacher in front of 100 kids is better than a bad teacher in front of 20 kids. One problem, right off the bat, is that teachers don't come in simple categories of good and bad, and the teacher who is effective in front of 20 kids might not be so good if given 100. Furthermore, while competitive wages help in recruiting higher-quality entrants to the teaching profession, wages interact with working conditions to influence both recruitment and retention. Teachers, like all workers, evaluate their options—both where they teach and whether they want to be or remain teachers—based on an array of preferences for compensation, location, and working conditions. Those who are more qualified and more mobile have more choices. Smaller classes and fewer cumulative students across all sections or courses taught (total student load) are preferred by teachers, all else equal. So too are high-quality facilities and learning spaces. To achieve any particular set of educational goals and student outcomes requires both reasonable class sizes and competitive pay. It's both/and, not either/or. If class sizes are too large for instruction to be effective, and pay only reasonably competitive, there's little room to trade one for the other. All of these decisions occur in a context of competing school districts in geographic proximity, with varied class sizes and wages. School districts don't make these decisions in a vacuum.[101]

As we discuss later in this section, even when flexibility is available for new, different, innovative educational models, as in charter and private schooling, most schools still adhere to a similar human resource intensive model. For example, well-funded, prominent charter school operators with a successful track record are typically those that provide extended days and school years, and compensate teachers accordingly, while elite, high-tuition private schools market their small classes, advanced and elective courses, and personal attention, including college counseling. Those that have attempted technological substitution, most notably online schooling, have produced especially poor outcomes, as have bargain-basement private schools participating in statewide voucher programs.

While there may be some important variations in the resource allocation choices that public school districts, charter schools, or private schools make, in the end, it comes down to having enough money to make good choices—to pay for the things needed to run school and provide the programs and services that meet students' needs toward achieving the desired outcome goals.

4.2 INVESTMENTS WITH A TRACK RECORD

In this section, we focus on literature addressing (a) teacher compensation, teacher quality, and student outcomes; and (b) instructor-to-student ratios (or teacher quantity) and student outcomes. Both of these resource measures have financial implications: as noted, personnel compensation represents over 80 percent of a typical school budget, and teachers are the largest contingency. Thus, it is natural, when exploring whether money matters, to explore whether *things that cost money* matter.[102] In our previous edition, we focused on "class size" in particular, which is one important part of the "teacher quantity" conversation. But it's really more broadly about the trained adult/instructor/tutor time with and attention to students. In this edition, we expand our scope to include some discussion of other human resource intensive approaches, such as high-dosage tutoring. These are all staffing quantity measures that require hiring additional professional-level instructional certificated staff members for a given enrollment level, which requires additional spending. We also address literature that attempts to unpack the tradeoffs between spending on increased teacher quantity versus increased teacher wages.

Additionally, since our last edition, the literature linking spending on school facilities and student outcomes has expanded, and we address the pathways by which investment in schooling capital matters for students and influences teacher retention.

4.2.1 | TEACHER QUALITY AND WAGES

One of the earliest comprehensive studies linking indicators of teacher quality to educational outcomes was, again, the 1966 Coleman Report. The report looked at a variety of specific schooling resource measures, finding a positive relationship between school district average teacher experience and education levels and the overall outcomes of their students. A multitude of studies on the relationship between teacher characteristics and student outcomes have followed, producing mixed messages as to which characteristics matter most and by how much.[103] In this section, we synthesize research linking teacher

quality and wages to student outcomes. We first explain how and why teacher compensation matters for improving student outcomes, drawing on evaluations of across-the-board pay increases, retention bonuses, merit pay, and the role of alternative non-teacher wages. We then synthesize research examining the link between teacher credentials, especially teacher experience, education level, and certification—the main drivers of salary schedules—and student outcomes. We close by reiterating why comprehensive salary investment, beyond simple redesigning of teacher pay to include more incentives, is needed to sustain a high-quality teacher workforce.

4.2.1.a | How and Why Teacher Compensation Matters

Adequate teacher compensation is essential for ensuring a stable and high-quality educator workforce. Teacher salaries affect educational outcomes in a variety of ways. Numerous studies examine whether more competitive salaries (or salary incentives) improve (a) district *hiring* and the number of candidates entering the field of teaching; (b) job *retention* and workforce stability; and (c) *effort*, including increased effort in classroom instruction, grading, and lesson planning. The research shows money plays a critical role for improving hiring and retention, but has less impact on teacher effort. Importantly, as we discuss below, there is little evidence that school districts can improve hiring, retention, and effort simply by redesigning their approach to compensation—for example, by basing teacher pay on performance and evaluation—without incurring additional costs.

Teacher hiring. A substantial body of literature validates the conclusion that higher overall salaries attract a greater number and/or more qualified set of teacher candidates into preparation programs, into school district applicant pools, and into new teaching positions.[104] For example, Murnane and Olson (1989) found that salaries affect the decision to enter teaching and the duration of the teaching career,[105] while Figlio (1997, 2002) and Ferguson (1991) concluded that higher salaries are associated with more qualified teachers.[106] In addition, studies have demonstrated the important role of non-teacher salaries in shaping the teacher labor market (wages of other workers at similar age and education level in the same labor market).[107] Loeb and Page (2000) showed that:

> *"Once we adjust for labor market factors, we estimate that raising teacher wages by 10 percent reduces high school dropout rates by 3 percent to 4 percent. Our findings suggest that previous studies have failed to produce robust estimates because they lack adequate controls for non-wage aspects of teaching and market differences in alternative occupational opportunities."*[108]

Hough and Loeb (2013) similarly found evidence that higher salaries at the school district level increased the size of that district's teacher applicant pools and increased the quality of new hires, with the latter based on estimates of newly hired teachers' impact on test scores during their first years.

In short, more adequate teacher salary levels influence the quality of the teaching workforce, which in turn affects student outcomes. Teacher salaries also send a signal to potential teacher candidates earlier in the preparation pipeline, such that any significant reductions in teacher compensation, such as those due to a tax levy limit, can alter the behavior of would-be teacher applicants. For example, Figlio and Rueben (2001) note: "Tax limits systematically reduce the average quality of education majors, as well as new public school teachers in states that have passed these limits."[109]

Even at a national scale, higher salaries appear to expand the quality and quantity of new applicants into the field. Kraft and Lyon (2024), for example, identify a nationwide increase in real teacher wages during the 1990s and 2000s, one which corresponded with an increase in the perceived prestige of the teaching profession and an increase in the number of college students preparing to be teachers.[110]

Recent analyses also identify money leveraged as sign-on bonus incentives as effective mechanisms for improving hiring practices. For example, Hawai'i's statewide district implemented a $10,000 bonus program for special education teachers following significant shortages in that area. Researchers found the salary increase led to a 32 percent reduction in the proportion of vacant special education positions and reduced the proportion of unlicensed teachers among incumbents.[111] Most new applicants to special education positions came from general education teachers already working in the system, suggesting that during periods of acute labor shortages, districts can leverage financial incentives to address staffing challenges.

Retention. A substantial body of work links teacher salary increases to improved retention, which in turn improves teacher-student relationships, learning environments, and student outcomes.[112] Studies either examine how across-the-board salary increases improve retention rates for different teachers or examine the influence of targeted retention incentive policies on retention.

One recent study by Sun et al. (2024) examined large, across-the-board salary increases in Washington state that were funded through a major court-mandated state finance reform.[113] The authors found those districts receiving the largest increases in state aid, and providing the largest salary increases, had higher teacher retention in the first year compared to their counterparts in districts receiving smaller salary increases. These effects are actually likely to be underestimated because all districts in the state enacted substantial salary increases, and the authors causal research design requires examining only variation in salaries *across* districts. Several other studies explore the influence of across-the-board salary increases. Hendricks (2014, 2015) in particular has used differences in the structuring of traditional salary schedules in Texas to demonstrate how different design approaches can improve retention for specific groups of teachers. His work shows, for example, that, on average, a 1 percent increase in salary is associated with a 1.4 percent decrease in turnover, with larger effects occurring among early career teachers, and null effects among teachers with more than 19 years of experience.[114]

Other salary policies provide bonuses for teachers with specific qualifications or performance ratings, or to work in specific schools, such as those serving higher-poverty student populations or those facing staffing challenges. Clotfelter et al. (2005, 2008a, 2008b) found that, despite some initial implementation challenges, a retention bonus of $1,800 in North Carolina led to a decline in turnover by 4.2 percentage points.[115] Cowan and Goldhaber (2018) found a program in Washington state providing bonuses for National Board Certified teachers to work in high--poverty schools reduced teacher turnover rates by about 4 percentage points for eligible schools. Several other studies identify positive effects of monetary retention bonuses, including programs that distribute financial incentives through teacher loan forgiveness.[116]

Effort. Contemporary efforts to tie teacher pay directly to productivity, including performance bonuses based on student test results, have generally failed to produce concrete results in the United States.[117] The primary purpose of these performance-based incentive programs, which received a great deal of attention and advocacy during the late 2000s and early 2010s, is to boost teacher effort and, therefore, to improve their students' measured performance.

However, recently published studies of individual and group financial incentives continue to find mixed to null effects,[118] although alternative compensation models in some settings have yielded positive results. Dee and Wyckoff (2015) study the DC Impact program and find some evidence that a comprehensive strategy combining frequent teacher evaluations and financial incentives can yield marginal improvements to the average rate of student achievement growth among retained teachers.[119] Similarly, in a study of an Austin, Texas, pay-for-performance (P4P) program, Balch and Springer (2015) found that the school district's REACH program resulted in positive gains in both math and reading in the first year of implementation.[120] An evaluation of Minnesota's Q Comp program reached similar results raising achievement by 0.03 standard deviations.[121] Fryer and colleagues (2012) study the effect of providing teachers' bonuses in advance and taking the money back if students do not improve sufficiently, and find even larger impacts on test scores.[122] A few more recent analyses have identified small positive effects from teacher performance pay policies.[123]

Still missing in this literature on teacher "pay structure" reforms, such as merit pay, are cost-effectiveness comparisons of the alternatives.[124] That is, if we take the same total payroll dollars and allocate those dollars traditionally across teachers with incremental differences in salaries by experience and credentials held, as opposed to implementing those salaries and bonuses by the above alternatives (along with paying for the associated costs of the evaluation metrics used for allocating salaries), do we see differences in the production of student outcomes? Can comparable or better outcomes be achieved where the summed costs of alternative pay structures are equal to or less than current costs?

Rothstein (2012) critiques the presumption that tying teacher pay directly to measures of performance outcomes would necessarily improve the efficiency of money allocated to compensation. He explains:

> *"Simulations indicate that labor market interactions are important to the evaluation of alternative teacher contracts. Typical bonus policies have very small effects on selection. Firing policies can have larger effects, if accompanied by substantial salary increases. However, misalignment between productivity and measured performance nearly eliminates the benefits while preserving most of the costs."*[125]

Those observations align with more recent empirical evidence from the DC Impact program discussed earlier, where teachers react to the threat of job loss, but monetary incentives tied to student performance often do not provide sufficient motivation for teachers to change instructional behaviors.[126]

Yuan et al. (2013) identify similar issues from three pay-for-performance programs in three different school districts, all of which were evaluated through randomized controlled trials, with limited impacts on student outcomes. The programs included the Project on Incentives in Teaching in Tennessee (Springer et al., 2010), the Pilot Project on Team Incentives in Texas (Springer et al., 2012), and the School-Wide Performance Bonus Program in New York City Public Schools (Marsh et al., 2011). In addition to estimating the test-based effects of these incentive programs, the authors also surveyed teachers involved and found that teachers felt that changes in their personal efforts alone were unlikely to lead to significant changes in student achievement due to concerns about the influence of family environment on student achievement. Other teachers expressed concerns about a lack of fairness in the program, and that the monetary incentives were too low. In sum, teachers felt the monetary bonuses that were ultimately distributed were not actually "linked to their performance in the classroom" and few reported changes in instructional practices, number of hours worked, or collegiality. In one exception, teachers in the Tennessee program reported greater emphasis on test preparation as well as collaboration among colleagues, but, again, the program did not significantly alter student outcomes.

Several other studies find small positive effects associated with teacher merit pay; however, a consistent finding from these studies is the difficulty in implementing the program with fidelity, as well as the difficulty in gaining teacher buy-in.

In short, while there exists some new evidence that alternative compensation methods and evaluation metrics may yield some positive results, the evidence is mixed at best, and researchers do not as of yet have a deeper understanding of the relative cost-effectiveness of alternatives. What scholars have established in each case is that the overall level of teacher compensation continues to matter for recruitment and retention of talent into the teaching profession, relative to other labor market opportunities, which directly improves student outcomes through a higher-quality and more stable instructional workforce.

4.2.1.b | TEACHER CHARACTERISTICS THAT MATTER

An extensive literature explores teacher characteristics associated with improved student outcomes; studies link positive effects with teacher experience, education, and certification. As with other areas in school finance and economics of education research, many of the older studies are correlational, examining statistical relationships between, for example, teacher experience and student outcomes without sufficient statistical controls to account for nonrandom assignment of students to teachers. If, for instance, students facing more significant impediments to learning tend to be assigned to more experienced teachers within a grade level, the correlation between teacher experience and student outcomes at that school may be downwardly biased. Estimates could just as easily be upwardly biased if early career teachers are assigned a greater share of struggling students, but both examples highlight the difficulty in determining the influence of teacher experience on student outcomes.

Nevertheless, inconsistent findings on the relationship between teacher "effectiveness" and how teachers get paid—by experience and education—added fuel to the "money doesn't matter" fire during the first decade and a half of the 21st century.[127] Since a large proportion of school spending necessarily goes to teacher compensation, and (according to this argument) since school districts do not pay teachers in a manner that reflects or incentivizes their productivity, then spending more money won't help.[128] In other words, the assertion is that money spent on the

current system doesn't matter because teacher salary is distributed based on characteristics not directly tied to job performance.

This argument of course misses the important point about the general role of experience and education in determining teachers' salaries, and what that means for student outcomes. While teacher salary schedules may determine pay differentials across teachers *within* districts, the simple fact is that *where* one teaches (i.e., in which district) is also very important in determining how much they make.[129] Arguing over attributes that drive the raises in salary schedules also ignores the bigger question of whether paying teachers more in general might improve the quality of the workforce and, ultimately, student outcomes. Teacher pay is increasingly uncompetitive with salaries offered by other professions, and the "penalty" teachers pay increases the longer they stay on the job.[130]

Moreover, the tenuous argument that experience and education-based salary schedules are necessarily inefficient conflicts with more recent evidence about the effects of teacher experience. In the wake of growing literature and policy rhetoric asserting the inefficiency of paying teachers according to experience and credentials, a handful of new studies have surfaced revealing that the gains in student outcomes resulting from increased teacher experience may extend well beyond the first few years of experience.[131] Thus, it would not be entirely inefficient for salaries to continue scaling upward with increased experience, especially given the additional costs of implementing alternative measures on which to base salaries. Wiswall (2013) finds that for mathematics achievement "there are high returns to later career teaching experience, about twice as much dispersion in initial teacher quality as previously estimated, and a pattern of negative selection where high quality teachers are more likely to exit."[132] Papay and Kraft (2015) find that while teachers experience rapid productivity improvement in their early careers, teachers continue to build human capital beyond these first years.[133] And Ladd and Sorensen (2014) similarly find additional returns to experience in terms of higher test scores and improvements to student behavior beyond the first few years of teaching.[134] Perhaps most importantly, the overall efficiency and effectiveness of teacher compensation does not depend exclusively on the extent to which each dollar allocated to any and every teacher's salary can be associated precisely with a measurable, marginal gain to the test scores of that teacher's students. First, the benefits of schooling extend beyond short-term achievement gains. Second, teacher compensation exists, and exerts whatever influence it may have, within a complex social and economic system.

Finally, we return to the argument that districts can redesign teacher compensation to improve hiring, retention, and effort—e.g., through merit pay—without increasing overall teacher salary costs. Assertions that performance-based pay is necessarily more cost-effective than traditional salary structures falsely assume traditional step-and-lane salary schedules to be monolithic. In practice, salary differentials associated with experience and credentials vary widely. Some are compressed from top to bottom, while others are not, and they may favor experience over credentials or vice versa. Hendricks (2015a, 2015b) explores these issues:

> *"Increasing salaries for teachers with 3 or more years of experience differentially retains high-ability teachers, while higher salaries for teachers with 0-2 years of experience differentially retain low-ability teachers. This likely occurs because higher early-career salaries disrupt a positive sorting process that exists among novice teachers."*[135]

That is, one might restructure traditional salary schedules to improve teacher hiring and retention rather than trying to change the salary schedule to align with difficult-to-measure outcome-based incentives. For example, Hendricks also finds that changing salary structures may alter recruitment potential and the recruiting pool:

> *"Overall, a 1% increase in base salary for teachers of a particular experience level increases the proportion of the targeted teachers hired by 0.04-0.08 percentage points. Pay increases have the largest effect on hire rates among teachers with 2-3 years of experience and the effect diminishes with experience. I show that higher teacher salaries provide a dual benefit of retaining and attracting a more effective distribution of teachers. Districts may also improve student achievement growth at no cost by reshaping their salary schedules so that they are increasing and concave in teacher experience."*[136]

To summarize, despite all the uproar about paying teachers based on experience and education, and its misinterpretations in the context of the "does money matter?" debate, the evidence suggests that the simplest forms of the merit pay approach too often miss the point. To whatever degree teacher pay matters in attracting high-quality educators into the profession and retaining them, it's less about *how* they are paid than *how much*. Furthermore, the average salaries of the teaching profession, with respect to other labor market opportunities, can substantively affect the quantity and quality of applicants to preparation programs, entrants to the teaching profession, retention within the field, and, ultimately, student outcomes.[137] Diminishing resources for schools can constrain salaries and reduce the quality of the labor supply, while salary differentials between schools and districts might help to recruit or retain teachers in high-need settings. In other words, when it comes to improving teacher quality, there are no cheap solutions. Resources matter.

4.2.2 | CLASS SIZE AND TEACHER QUANTITY

Class-size reduction is often characterized as a prohibitively expensive use of additional school dollars.[138] To be sure, reducing class sizes costs money, since you have to hire additional teachers, but the question of whether it's *too* expensive must rely on detailed comparisons of alternative uses of the same dollars, or the effects on student outcomes of those alternative uses.

Instead, most arguments against class-size reduction proceed by noting that there are significant costs to adding more teachers and classrooms (which, again, is an unsurprising revelation),[139] followed by a (often vague) statement as to the differences between the most and least "effective" teachers (as measured by their effects on test scores). In other words, so this argument goes, we should improve teacher quality rather than quantity. The problem here—in addition to the implication that this is an either/or proposition—is that one cannot compare the cost-effectiveness of class-size reduction with "improving teacher quality," which is an outcome, not a concrete policy with measurable costs and benefits. As discussed above, policies designed to improve teacher quality have a very mixed track record of success, though there is some evidence that frequent high-quality evaluations and feedback can improve performance of existing teachers.

What we do know, in contrast, is that ample research indicates that children in smaller classes achieve better outcomes, both academic and otherwise, and that class-size reduction can be an effective strategy for closing racial and socioeconomic achievement gaps.[140] For example, Krueger (1999), in a re-analysis of data from the large-scale randomized Tennessee Project STAR class-size reduction study, concluded:

> *"The main conclusions are 1) on average, performance on standardized tests increases by four percentile points the first year students attend small classes; 2) the test score advantage of students in small classes expands by about one percentile point per year in subsequent years; 3) teacher aides and measured teacher characteristics have little effect; 4) class size has a larger effect for minority students and those on free lunch."*[141]

In a more recent analysis, also re-evaluating the Tennessee STAR data, Konstantopoulos and Chun (2009) wrote:

> *"We used data from Project STAR and the Lasting Benefits Study to examine the long-term effects of small classes on the achievement gap in mathematics, reading, and science scores (Stanford Achievement Test). The results consistently indicated that all types of students benefit more in later grades from being in small classes in early grades. These positive effects are significant through grade 8. Longer periods in small classes produced higher increases in achievement in later grades for all types of students. For certain grades, in reading and science, low achievers seem to benefit more from being in small classes for longer periods. It appears that the lasting benefits of the cumulative effects of small classes may reduce the achievement gap in reading and science in some of the later grades."*[142]

Researchers continue to revisit data from the Tennessee STAR study, which in more recent years has permitted researchers to explore long-term outcomes of those students randomly assigned to smaller class sizes versus their peers who were not. For example, Dynarski, Hyman, and Schanzenbach (2013) find:

> *"Assignment to a small class increases the probability of attending college by 2.7 percentage points, with effects more than twice as large among blacks. Among students*

> *enrolled in the poorest third of schools, the effect is 7.3 percentage points. Smaller classes increase the likelihood of earning a college degree by 1.6 percentage points and shift students towards high-earning fields such as STEM (science, technology, engineering and mathematics), business and economics. We find that test score effects at the time of the experiment are an excellent predictor of long-term improvements in postsecondary outcomes."*[143]

Admittedly, there remain some naysayers on whether class-size reduction yields *cost-effective* benefits in terms of student outcomes. But the findings upon which these counterarguments are based often lack the weight of large-scale randomized studies, such as Tennessee's Project STAR, relying instead on natural variations in class sizes across schools.[144]

In addition, assertions regarding the excessive cost and inefficiency of class-size reduction often lack rigorous cost-effectiveness analysis. In a 2011 brief for the Center for American Progress, for example, author Matthew Chingos asserted that class-size reduction is the "most expensive school reform."[145] But that same report provided no direct cost or cost-effectiveness comparisons between class-size reduction and other alternatives. A later version of the review by Chingos (2013) published as a policy paper in the *Journal of Policy Analysis and Management* criticized class-size reduction as broad state-imposed policy, revisiting the costs and potential downsides of statewide class-size reduction policies implemented in California and Florida.[146] Chingos suggests that estimates of long-term earnings of students subjected to class-size reduction do not justify the cost,[147] but he also acknowledges that direct comparisons between spending on class-size reduction and other alternatives are few and far between.

Dynarski, Hyman, and Schanzenbach (2013) provide the most direct cost-effectiveness comparison of class-size reduction policies with other options for which sufficient data on costs and outcome benefits were available:

> *"A fair conclusion from this analysis is that the effects we find in this paper of class size on college enrollment alone are not particularly large given the costs of the program. If focused on students in the poorest third of schools, then the cost-effectiveness of class size reduction is within the range of other interventions. There is no systematic evidence that early interventions pay off more than later ones when the outcome is limited to increased college attendance."*[148]

It's true that a large body of the literature on the effectiveness of class-size reduction relies on data from a relatively small group of sources, most notably, the Tennessee STAR experiment.[149] Further, many class-size reduction studies finding substantial benefits have focused on class-size reduction in early grades (K-3), and most of these programs are pilots implemented on a relatively small scale. (A comprehensive review of the literature on class-size reduction is beyond the scope of this brief, but see endnote 66 for additional resources.)[150]

One recent study takes a somewhat different approach, studying class-size changes in their dynamic social context, along with other reform policies and strategies, more like the work of Gilpin and Kaganovich (2011) and Rothstein (2012) discussed in the previous section on teacher compensation. Specifically, Gilraine and colleagues (2020) revisit data on California's statewide class-size reduction program of the late 1990s. The program was often criticized as being very costly and leading to a shifting of higher-quality teachers to more desirable schools and districts, leaving low-income and minority children with smaller classes taught by less qualified teachers. That is, this particular case is frequently cited for why class-size reduction is both costly and ineffective as broad-based policy. But, retrospectively, Gilraine and colleagues find quite the opposite, and identify additional causal mechanisms. First, the authors find a large, positive direct effect of smaller classes on test scores. Second, they conclude that providing smaller classes made schools more appealing to families with children attending private school, who then switched to the public schools. Gilraine and colleagues find a positive indirect effect of these demographic changes that was even larger than the direct effect, increasing the program's benefit-cost ratio significantly.[151]

In a separate recent study, using data from New York City from 2009 to 2013, Gilraine (2020) evaluated the effects of newly introduced class-size caps, finding that the class-size reductions resulted in increased student achievement.[152] Over the short time period studied, Gilraine did find some loss to those gains as a function of newly hired

teachers. This study illustrates how the causal methods used to evaluate policy change effects in school finance reform studies above can also be used for evaluating changes to class-size policies, expanding and updating the literature on class-size effects beyond STAR.

It's true that reducing class sizes costs more than not reducing class sizes. But class-size reductions, implemented effectively, have positive effects. As such, one can reasonably infer that using increased resources to reduce class sizes would have positive effects, or that resources matter.

While it's certainly plausible that other uses of the same money might be equally or even more effective, there is little evidence to support this. For example, while we are quite confident that higher teacher salaries lead to increases in the quality of applicants to the teaching profession and increases in student outcomes, we do not know whether the same money spent toward salary increases would achieve better or worse outcomes if it were spent toward class-size reduction. Indeed, some have raised concerns that large-scale class-size reductions can lead to unintended labor market consequences that offset some of the gains attributable to class-size reduction (such as the inability to recruit enough fully qualified teachers).[153] There is clearly a need for a more precise cost-benefit analysis.[154] Still, the preponderance of existing evidence suggests that the additional resources expended on class-size reductions do result in positive effects.

Finally, in the aftermath of the COVID-19 pandemic and declining student test scores that resulted from the disruption of student lives and the delivery of online instruction, researchers have sought to identify interventions that might be most effective at recovering that pandemic-induced "learning loss." Much of the focus has been on the potential for high-dosage tutoring as a supplement to, but not a replacement for, traditional public schooling. Many studies on the effects of tutoring predate the pandemic and have been summarized in a handful of recent articles. Kraft and Falken (2021), in a blueprint for scaling up tutoring and mentoring, provide a synthesis of the tutoring and mentoring literature, noting that "tutoring can be effective for students across the full distribution of achievement, with mixed findings about which students benefit most," and that "students might experience the broadest benefits when elements of school-based mentoring programs are integrated within tutoring programs that allow for sustained relationships and focus on clear academic and social/emotional development goals with frequent feedback" (p. 3).[155]

Nickow, Oreopolis, and Quan (2020) conducted a meta-analysis of nearly 100 studies on the effectiveness of tutoring programs, finding that:

> *"Tutoring programs yield consistent and substantial positive impacts on learning outcomes, with an overall pooled effect size estimate of 0.37 SD. Effects are stronger, on average, for teacher and paraprofessional tutoring programs than for nonprofessional and parent tutoring. Effects also tend to be strongest among the earlier grades. While overall effects for reading and math interventions are similar, reading tutoring tends to yield higher effect sizes in earlier grades, while math tutoring tends to yield higher effect sizes in later grades. Tutoring programs conducted during school tend to have larger impacts than those conducted after school" (p. 1).*[156]

The authors also unpack the comparative costs and benefits of tutors having different levels of professional training. While they find that classroom teachers "often make effective tutors, it is far from clear that the effectiveness differentials between trained teachers and paraprofessionals outweigh the costs" (p. 53). They find paraprofessionals to be more effective than non-professional or volunteer tutors and argue that "paraprofessional school staff members and recent graduates in professional fellowship programs represent promising bodies of potential tutors" (p. 54).

In a very recent study, Guryan and colleagues also find large positive effects of tutoring in a randomized controlled trial involving 9th- and 10th-grade students in Chicago Public Schools. The authors conclude that the cost-benefit ratio for the specific intervention is comparable to many successful early childhood programs.[157]

4.2.3 | CAPITAL INVESTMENT

Facilities and infrastructure investments have both direct and indirect influence on student outcomes. In the meta-analysis discussed previously by Jackson and Mackevicius (2023), nine studies looked specifically at capital construction projects.[158] According to the results of the meta-analysis, the "Marginal effects of capital spending are similar to non-capital, and effects are similar across baseline spending levels and geography."[159] Furthermore,

the authors show that four to six years after the capital investment, effects are consistently positive, even at the low end of estimates. Among the studies included in the meta-analysis, one 2019 evaluation found that "results consistently suggest that passing a bond measure increases achievement among low- but not high-SES students. However, these benefits for low-SES students are delayed and emerge 6 years after an election."[160] That is, it takes time for bond measures to result in capital expenses resulting in fully equipped facilities. And extended exposure to those new facilities yields significant, long-term positive effects.

In another recent study, Biasi, Lafortune, and Schonholzer (2024) evaluate the effects of different types of capital investments on student achievement outcomes and capitalization in housing values. The study included data from 29 states and yielded a number of novel and more nuanced findings than previous studies linking only the front-end investment, or choice to invest, with the back-end student outcomes. The authors do generally find that investment in schooling capital yields both improved student test scores and housing prices. But the authors also find the greatest benefits from, in descending order, investments in (a) heating, ventilation, and air conditioning (HVAC) systems; (b) science facilities; (c) health and safety measures; (d) plumbing, roofing, and furnaces; and (e) classroom improvements. Investments in athletic facilities and classrooms yielded the greatest returns in terms of housing values. The authors also find that "socio-economically disadvantaged districts benefit more from capital outlays, even conditioning on project type and the existing capital stock," and that "closing the spending gap between high- and low-SES districts and targeting spending towards high-impact projects may close as much as 25% of the observed achievement gap between these districts" (p. 1).[161] The effects of investment in HVAC systems on student outcomes coincides with a handful of studies over time validating the importance of well-regulated temperature for student learning and specifically student test taking.[162] Investments in textbooks also matter.[163]

Several studies also point to the indirect effects of high-quality facilities in promoting teacher recruitment, retention, and overall effectiveness.[164] Ladd (2011), in a study of North Carolina public schools, addresses the indirect effects of school facilities through their influence on teacher perceptions of their working environments, noting that "For reading, teachers' perceptions of facilities are also predictive of positive school effects." [165] Maxwell (2016) finds that "academic achievement is linked to building condition mediated by the social climate and student attendance."[166]

4.3 ARE ALTERNATIVE DELIVERY SYSTEMS MORE EFFICIENT?

In this final section, we address whether two common types of interventions—charter school expansion and private schools (including vouchers)—show promise for creating better student outcomes at lower cost. That is, do charter schools and vouchers offer the same or better student outcomes at lower cost?

Prior to addressing this question, it's important to understand that public expenditure is not necessarily a reflection of full cost. A state might choose, for example, to provide families with $8,000 per child to spend toward private school tuition. The tuition at a typical private school might be double that amount and even full tuition may not cover that school's operating costs of educating that student. There exist no comprehensive data sources on private school expenditures or outcomes, complicating any effort to evaluate whether private schools in states subsidizing attendance or not do more with less or otherwise.

Charter schooling raises similar concerns in that state data systems lack precision in establishing comparability between spending in charter schools and public district-operated schools.[167] In many cases, regular public districts spend their own resources to provide services to children in charter schools, or for managing student choice systems and school assignment. That spending shows up as a district, not charter, expenditure, thus overstating district spending per pupil while at the same time understating charter school spending. Other inconsistencies exist in the reporting of private contributions to charter schools or the organizations that manage them. Little data exist on the spending of the management organizations overseeing charter networks, which are roughly analogous to district central offices but often not subject to the same requirements for public financial disclosure. That said, we begin with the literature on charter schooling, as there is at least somewhat better reporting and a body of literature comparing their relative efficiency to district schools.

4.3.1 | DO CHARTER SCHOOLS PROVE MONEY DOESN'T MATTER?

Some argue that charter schools generally achieve more for less or the same funding than traditional public schools serving similar student populations, thus providing support for the idea that money doesn't matter for improving school quality.[168] The core assumption is that charter schooling improves efficiency because the flexibility afforded through chartering permits charter schools to engage in more creative teacher compensation strategies and technological substitution, such as trading small class sizes for efficient use of technology through blended and online learning. Further, efficiency improvement yielded by charter innovations creates competitive pressure on traditional public schools to improve.[169]

Regarding productivity improvements from technological substitution, a recent review of charter school literature by Epple, Romano, and Zimmer (2015) characterized online and cyber charter schools in particular as a "failed innovation, delivering markedly poorer achievement outcomes than traditional public schools" (p. 55).[170] That said, we do not know if these markedly poor achievement outcomes were achieved with markedly fewer resources, and thus, a break-even on efficiency.

The assertion of large efficiency gains through chartering is often built on poor and/or misestimation of the resources received and used by charter schools. Specifically, advocates assert that charter schools generally receive less funding than do traditional public schools and achieve the same or better outcomes, thus making them more efficient.[171] The first assertion, that charter schools receive less funding and spend less, is certainly not uniformly true.[172] Baker, Libby, and Wiley (2015) explain that charter school spending varies substantially by context and by operator within context.[173] Knight and Toenjes (2020), in a study of Texas charter schools, found "after accounting for differences in accounting structures and cost factors, charter schools receive significantly more state and local funding compared to traditional public schools with similar structural characteristics and student demographics."[174] In a study completed on behalf of the Maryland Department of Education, authors from the American Institutes for Research (AIR) found: "In all districts except Frederick, the predicted expense is less than the actual charter expense, indicating that average spending would be less for these charter schools if they followed the spending patterns of traditional schools in their district."[175] That is, when modeled by regression analysis, given a variety of student and school characteristics, charter schools were spending more than expected (meaning, more than otherwise similar traditional public schools). Authors from AIR arrived at similar findings using similar methods in a study completed as part of the Getting Down to Facts project in California.[176] Some charter operators, in some contexts, spend substantially more than both other charter schools and traditional public schools in the same area, while others spend much less. The second assertion, that charters systematically outperform traditional district schools, is also unsupported by the evidence, [177] and the specific assertion that those that do spend much less perform similar to or better than traditional district schools stands largely untested.

A handful of studies identify significant positive achievement effects of schools from the Knowledge Is Power Program (KIPP) network, but this same research provides only weak, imprecise measures of the resources available in these schools.[178] Baker, Libby, and Wiley (2012, 2015) indicate that KIPP schools in New York and Texas tend to spend substantially more than traditional district schools serving similar populations.[179] Dobbie and Fryer (2011) declare that high standards and "no excuses" strategies of select charter school operators are more important than spending differences in producing improved student outcomes.[180] But spending measures in the study are poorly documented and incomplete. Baker, Libby, and Wiley's (2012, 2015) review of financial documents and public data, applying model-based comparisons of school-site expenditures to schools serving similar student populations, reveals that many of the school operators involved in Dobbie and Fryer's study spent far more than similar district schools.[181] Baker, Libby and Wiley (2012) also explain that much of the additional spending among high-spending charter operators is allocated to maintaining smaller class sizes, providing longer school days and years, and paying more to teachers, holding experience and education levels constant, for working those additional hours. That is, the investments by charter operators follow traditional wisdom and are not especially innovative. [182]

Perhaps the strongest empirical evidence of charter school efficiency advantages comes from the work of Gronberg, Taylor, and Jansen (2012) on Texas charter schools.[183]

The authors find that, generally, Texas "charter schools are able to produce educational outcomes at lower cost than traditional public schools—probably because they face fewer regulations—but are not systematically more efficient relative to their frontier than are traditional public schools" (p. 302).[184] In other words, while the overall cost of charter schools is lower for comparable output, the variations in relative efficiency among Texas charter schools are substantial. Efficiency is neither uniformly nor consistently achieved.

Related work by these authors reveals that the lower overall expenses among charters are largely a function of lower salaries and inexperienced staff (Taylor et al., 2011).[185] That is, the difference in total staffing cost, and resulting difference in total instructional expense per pupil, was largely due to the reduced experience levels of teachers, resulting in part from the fact that many of the schools existed for fewer than 10 years (many fewer than five), in addition to high turnover among teachers in their first few years. That is, compensation was lower not because of creative technological substitution or alternative compensation, but because of relative inexperience and high turnover among educators. Epple, Romano, and Zimmer (2015) suggest that these patterns are similar across studies of charter school teachers.[186] Thus, estimated efficiency gains, where they do exist, may rely on maintenance of high turnover and relatively inexperienced staff, a questionably scalable and sustainable option.

Put simply, research on the charter school sector in the aggregate tells us little about whether and to what extent money matters, or if money can be made to matter more or less than it currently does leveraged through traditional investments in public schooling. Some charter schools spend much more than both other charter schools and traditional public schools and appear to yield benefits to students from that spending. Others spend less and do poorly, and still others spend less but do less poorly than expected (and are thus more efficient). Still, the variations in the charter school sector, and the variations across traditional public schools, may provide insights down the line in how to more effectively and efficiently leverage resources. By and large, charter schools that spend more appear to do so by providing competitive compensation for their teachers, offering longer school days and years, and maintaining smaller classes, while those that spend less do so by maintaining inexperienced staff and high turnover.

4.3.2 | WHAT DO WE KNOW ABOUT PRIVATE SCHOOL SPENDING AND OUTCOMES?

Far less evidence exists on private schools and their relative efficiency compared to public school systems. This, again, is mainly because there exists no comprehensive source of data on private school spending, demographics, and other structural features that would be necessary for making such comparisons. The National Center for Education Statistics does biennially collect data on private schools, their affiliations, and their enrollments and locations. However, despite their efforts, these data are inconsistent and incomplete (and do not include finance measures).[187] Private schools registered as nonprofits and not governed or operated by religious organizations are compelled to file IRS tax forms (990), but these forms lack precise information on operational and capital spending.

Lubienski and Lubienski (2013) explain, in a book-length volume, that private school student outcomes are largely a function of the students those schools select, retain, and serve.[188] While early studies of smaller-scale voucher programs in cities like Milwaukee, Washington, D.C., and New York City showed gains in graduation rates and educational attainment, more recent studies of larger-scale statewide voucher programs have found negative effects on student outcomes, especially in math.[189] But this finding neither supports nor refutes whether private schooling, publicly subsidized or not, can do more with less. It may simply be the case that the voucher payments provided only allow for access to lower-spending, lower-quality schools—that is, private schools that get less for less.

What we do know about private school expenditures and how those expenditures translate to the delivery of schooling resources and programs comes largely from a 2009 report that gathered the tax filings (IRS form 990) from approximately 1,500 private schools across the country,[190] merged the revenue and spending data with data on enrollments from the NCES Private School Universe Survey, and compared spending by school type and spending level with reported tuition rates, pupil-to-teacher ratios, teacher salaries, and teacher characteristics from the NCES Schools and Staffing Survey.[191] That study found that private independent day schools (a) spend nearly double the per-pupil spending of surrounding public districts; (b) operate for fewer days per year (about 165 compared to 180 to 185); (c) pay their teachers similarly on an annual basis; and (d) have pupil-to-teacher ratios

significantly lower than public districts. Schools belonging to Christian school associations (a) spent about 70 percent of the surrounding public district level; (b) had pupil-to-teacher ratios somewhat lower than surrounding districts but higher than private independent schools; but (c) had much lower teacher salaries than surrounding public districts or private independent schools. Catholic schools reporting spent similarly to district schools and had similar staffing ratios, but lower salaries. In short, private schools with more resources seem to leverage those resources to pay better salaries than other private schools, but even more so, to provide smaller classes. However, we have no way of knowing how or whether these choices translate to differences in student outcomes, just that private schools with more resources tend to put a larger share into smaller classes than into wages.

5.0 SORTING OUT COST/QUALITY RELATIONSHIPS

One additional body of literature addresses a related question: Does it cost more to achieve higher outcomes, or a broader set of outcomes? That question is essentially the inverse of our first question—would spending more improve outcomes?—and depends on the existence of a relationship between spending and schooling quality, as measured by student outcomes. The method of analysis is the *education cost function*. The education cost function is the conceptual flip side of the education production function. Like production function research, cost function research seeks to identify the link between spending variation and outcome variation, cross-sectionally and longitudinally. Notably, as we discussed earlier, this type of statistical modeling does not establish the causal link between spending and outcomes, but it does allow for inferences to be made regarding the amount of spending needed, on average under certain conditions, to achieve certain outcome goals. The goal of the education cost function is to discern the levels of spending associated with efficiently producing specific outcome levels (the "cost" per se) across varied geographic contexts and schools serving varied student populations. Most published studies applying cost function methodology use multiple years of district-level data, within a specific state context, and focus on the relationship between cross-district (over time) and sometimes cross-school variations in spending and outcome levels, considering student characteristics, contextual characteristics such as economies of scale, and labor cost variation. Some cost function studies evaluate whether varied expenditures are associated with varied levels of outcomes, all else being equal, while other cost function studies evaluate whether varied expenditures are associated with varied growth in outcomes.

The existing body of cost function research has produced the following (in some cases obvious) findings:

1. The per-pupil costs of achieving higher-outcome goals tend to be higher, across the board, than the costs of achieving lower-outcome goals, all else being equal.[192]
2. The per-pupil costs of achieving any given level of outcomes are particularly sensitive to student population characteristics. In particular, as concentrated poverty increases, the costs of achieving any given level of outcomes increase significantly.[193]
3. The per-pupil costs of achieving any given level of outcomes are sensitive to district structural characteristics, most notably, economies of scale.[194]

Researchers have found cost functions of particular value for evaluating the different costs of achieving specific outcome goals across settings and children. In a review of cost analysis methods in education, Downes (2004) explains: "Given the econometric advances of the last decade, the cost-function approach is the most likely to give accurate estimates of the within-state variation in the spending needed to attain the state's chosen standard, if the data are available and of a high quality" (p. 9).[195]

A flurry of peer-reviewed studies and rigorous state-specific analyses have been published in recent years, using increasingly consistent methodologies and yielding increasingly consistent findings.[196] For example, in a study of New Hampshire school costs, researchers from the American Institutes for Research found that cost models fit to state-sourced, school-level data and then rolled up to district-level predictions were correlated at .77 with cost estimates from Baker and Weber's national cost model estimated using federal-sourced, district-level measures.[197] Similarly, researchers from AIR, in a more recent study

of costs in Delaware schools, again comparing estimates derived from state-sourced, school-level data rolled up to district-level costs, found that their cost models were correlated with estimates from the National Education Cost Model at .89.[198] Importantly, what all of these estimates show is that it costs more to achieve more, and that it costs more in some settings than others to achieve a similar set and level of outcomes. That is, money matters.

Notably, the dollar-to-outcomes relationship from these cost modeling studies differs from the estimated return to investment of new dollars identified by Jackson and colleagues. But it does so in logical ways and for logical reasons. Often, the additional cost to raise district achievement to a desired level is less than the achievement gain estimated from causal studies for that same amount of spending increase. These differences are sometimes used by critics to assert that, therefore, cost function estimates are not valid representations of the spending-to-outcomes relationship. But it stands to reason that the two would produce different results in this particular direction and that this is not a valid falsification test for the cost function.[199] While the cost function is, to an extent, the flip side of the production function, one critically important difference is that when spending is the dependent variable, it becomes possible to control for those variations in spending that are not related to differences in outcomes. That is, to control for and remove from the spending-to--outcomes relationship those inefficiencies.[200] As such, one would expect that the estimated "cost" to achieve a specific improvement to outcomes may be less than the expected increase in outcomes from a given expenditure increase, if some of that increase was spent less efficiently (or not immediately realized in student outcomes).

Typically, efficiency is considered relative to the average actual behavior of local public school districts, which is another point of contention for critics of cost modeling as a method for parsing spending-to-outcome relationships.[201] The argument goes that one cannot possibly identify the efficient spending level associated with achieving any desired outcome level by evaluating the spending behavior of existing schools and districts, whose spending is largely inefficient (because, as discussed above, district expenditures are largely tied up in labor agreements that, according to these authors, are in no way linked to the production of student outcomes). If all schools and districts suffer such inefficiencies, then one cannot possibly discern underlying minimum costs by studying those institutions. However, this argument rests on the assumption that desired outcomes could be achieved while spending substantially less and entirely differently than any existing school or district spends, all else being equal. Evidence to this effect is sparse to nonexistent.

Authors of cost function research explain that the goal of cost modeling is more modest than exact predictions of minimum cost, and that much can be learned by better understanding the distribution of spending and outcomes across existing schools and districts, and the varied efficiency with which existing schools and districts achieve current outcomes.[202] That is, the goal of the cost model is to identify, among existing "outcome producing units" (districts or schools), the more (and less) efficient spending levels associated with given outcomes, where those more efficient spending levels associated with any given outcome provide a real-world approximation, approaching the minimum costs of achieving those outcomes.

In summary, while education cost function research is not designed to test specifically whether and to what extent money matters, the sizeable body of cost function literature does suggest that achieving higher educational outcomes, all else being equal, costs more than achieving lower educational outcomes. Further, achieving common educational outcome goals in settings with concentrated child poverty, children for whom English is a second language, and children with disabilities costs more than achieving those same outcome goals with less needy student populations. Cost models provide some insights into how much more money is required in different settings and with different children to achieve measured outcome goals. Such estimates are of particular interest in this period of time when more and more states are migrating toward common standards frameworks and common assessments but are still providing their schools and districts with vastly different resources. Cost modeling may provide insights into just how much more funding may be required for all children to have equal opportunity to achieve these common outcome goals.

6.0 SUMMING UP THE EVIDENCE

In this third edition of our "Money Matters" report, we have laid out a comprehensive review of the research on the effect of K-12 school funding on student outcomes. This review shows that a consistent flow of analyses since the mid-2010s, using better data and more sophisticated methods, have confirmed and elaborated on decades of prior research on the importance of adequate and equitable funding in K-12 schools. In summarizing this large body of evidence, we reach the following main conclusions:

Money matters, whether it's going up or down. The overwhelming bulk of studies we review show that infusions of additional money into schools lead to improved student academic achievement and outcomes later in life, while a handful of studies also validate that funding cuts, resulting from major events like the 2007-09 recession, lead to a decline in student outcomes.

Money matters, whether that money is driven into annual operating expenditures or capital investments. The largest share of annual operating spending in public schooling goes toward (a) the competitiveness of teacher and other school staff wages; and (b) the quantities of school staff that can be hired. In other words, it goes to paying teachers more and/or hiring more teachers. Both matter, and a high-quality public schooling system requires a "both/and approach," rather than an "either/or approach." Competitive wages are needed to maintain or improve the quality of the teacher workforce, as such quality matters for student outcomes. Reduced class sizes and staffing ratios (including tutoring) also lead to better student outcomes in the short or long term. On the capital investment side, spending on school facilities also improves student outcomes, both directly (e.g., providing healthy and safe spaces for student learning) and indirectly (e.g., supporting teacher recruitment and retention by offering high-quality, productive workspaces).

Money matters more—and has a more profound impact—for children experiencing poverty and in school districts and communities in which states have historically underinvested. Several studies discussed herein validate that spending more on schools and communities that have previously been deprived of resources yields greater returns on investment than spending where prior investment has been high and student need relatively lower. These findings validate the importance of promoting funding progressiveness in state school finance systems, with the goal of equal educational opportunity for all.

Money matters, regardless of how changes in funding come about. Changes in the amount and distribution of school dollars can occur due to a variety of reasons, including legislation initiated by state lawmakers, legislation in response to judicial pressure, local and national/global economic changes (e.g., recessions), and democratic processes (e.g., bond elections). Our review indicates that, whatever the cause of substantive changes in school funding, those substantive changes matter. They influence student outcomes; increased funding improves student outcomes and decreased funding harms student outcomes.

This growing body of evidence has helped to foster an emerging consensus among education researchers, advocates, and other stakeholders as to the importance of adequate and equitable K-12 funding. Yet this consensus is not reflected in many—perhaps most—states' school finance policymaking. We hope that the review of the evidence presented in this report will serve to inform school funding debates and policymaking going forward.

APPENDIX: DECOMPOSING THE EDUCATION DOLLAR

APPENDIX TABLE A

PERCENT OF K-12 CURRENT SPENDING BY CATEGORY AND DISTRICT POVERTY QUINTILE, 2016-2021

Poverty Group	Current Elem/Sec (% of Total)	Salaries & Benefits (% of Total)	Instruction (% of Current)	Salaries & Benefits for Instruction (% of Current)
1-Lowest	83.64%	69.57%	61.88%	54.31%
2-Low	85.32%	69.93%	61.01%	52.44%
3-Middle	86.79%	70.31%	60.57%	51.50%
4-High	86.89%	70.35%	59.63%	50.85%
5-Highest	87.51%	67.55%	61.95%	49.86%

APPENDIX TABLE B

PERCENT OF NON-ELEMENTARY/SECONDARY SPENDING, CAPITAL OUTLAY SPENDING, AND INTERGOVERNMENTAL TRANSFERS BY DISTRICT POVERTY QUINTILE, 2016-2021

Poverty Group	Non-Elem/Secondary (% of Total)	Capital Outlay (% of Total)	Transfers (% of Total)
1-Lowest	0.79%	11.02%	4.98%
2-Low	0.82%	9.66%	4.67%
3-Middle	0.76%	8.67%	4.31%
4-High	0.85%	8.17%	4.80%
5-Highest	0.93%	7.53%	7.14%

The total current expenditures for public elementary and secondary education are those that are associated with the day-to-day operations of the school district. These include expenditures for charter schools, if they exist in the district. They exclude long-term expenditures (like capital outlays), debt service, and expenditures beyond the scope of public elementary and secondary education. These data are taken from the CCD LEA Finance (F-33) survey.

NOTES

1 Baker, B., & Welner, K. (2011). School finance and courts: Does reform matter, and how can we tell?. Teachers College Record, 113(11), 2374-2414.

2 Black, D. W. (2018). The constitutional compromise to guarantee education. Stan. L. Rev., 70, 735.

Fujioka, K., & Knight, D. S. (forthcoming). McCleary at twelve: Examining policy designs following court-mandated school finance reform in Washington State. West Education Law Reporter.

3 Levin, J., Baker, B., Atchison, D., Brodziak, I., Boyle, A., Hall, A., & Becker, J. (2016). Study of funding provided to public schools and public charter schools in Maryland. American Institutes for Research.

4 https://www.census.gov/programs-surveys/school-finances/data/tables.html

5 https://nces.ed.gov/ccd/

6 pct_current= tcurelsc_f33full/totalexp_f33red

7 pct_transfers= (l12_f33full+ m12_f33full+ q11_f33full+ i86_f33full+v91_f33full+v92_f33full)/ totalexp_f33full

8 pct_nonelsec=nonelsec_f33full/totalexp_f33full

9 pct_capout= tcapout_f33full/totalexp_f33full

10 Because there are substantial differences in how states report certain data to the federal data collection, including whether they report all current year contributions to pension obligations, or whether school districts thoroughly report funding passed along to charter schools for charter school students (where charter schools are fiscally dependent on districts). Some states with particularly large independent charter school sectors (AZ) have simply given up on reporting their revenues and expenditures in the Census Fiscal Survey.

11 Knight, D. S. (2017). Are high-poverty school districts disproportionately impacted by state funding cuts? School finance equity following the Great Recession. Journal of Education Finance, 169-194.

Baker, B. D. (2014). Evaluating the recession's impact on state school finance systems. Education Policy Analysis Archives/ Archivos Analíticos de Políticas Educativas, 22, 1-30.

Baker, B. D., Farrie, D., & Sciarra, D. G. (2016). Mind the gap: 20 years of progress and retrenchment in school funding and achievement gaps. ETS Research Report Series, 2016(1), 1-37.

12 Oberfield, Z. W., & Baker, B. D. (2022). The politics of progressivity: Court-ordered reforms, racial difference, and school finance fairness. American Educational Research Journal, 59(6), 1229-1264.

13 Baker, B., & Duncombe, W. (2004). Balancing district needs and student needs: The role of economies of scale adjustments and pupil need weights in school finance formulas. Journal of Education Finance, 29(3), 195-221.

Baker, B. D., & Corcoran, S. P. (2012). The Stealth Inequities of School Funding: How State and Local School Finance Systems Perpetuate Inequitable Student Spending. Center for American Progress.

14 Baker, B., & Welner, K. (2011). School finance and courts: Does reform matter, and how can we tell?. Teachers College Record, 113(11), 2374-2414.

Candelaria, C. A., McNeill, S. M., & Shores, K. A. (2022). What Is a School Finance Reform? Uncovering the Ubiquity and Diversity of School Finance Reforms Using a Bayesian Changepoint Estimator. EdWorkingPaper No. 22-587. Annenberg Institute for School Reform at Brown University.

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

McNeill, Shelby M., and Christopher A. Candelaria. (2024). Paying for School Finance Reforms: How States Raise Revenues to Fund Increases in Elementary-Secondary Education Expenditures. (EdWorkingPaper: 24-892). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/vg4y-ga67

15 Baker, B. D., Farrie, D., & Sciarra, D. G. (2016). Mind the gap: 20 years of progress and retrenchment in school funding and achievement gaps. ETS Research Report Series, 2016(1), 1-37.

16 Baker, B. D. (2014). Evaluating the recession's impact on state school finance systems. Education Policy Analysis Archives/ Archivos Analíticos de Políticas Educativas, 22, 1-30.

17 Knight, D. S. (2017). Are high-poverty school districts disproportionately impacted by state funding cuts? School finance equity following the Great Recession. Journal of Education Finance, 169-194.

18 Jackson, C. K., Wigger, C., & Xiong, H. (2021). Do school spending cuts matter? Evidence from the Great Recession. American Economic Journal: Economic Policy, 13(2), 304-335.

19 Oberfield, Z. W., & Baker, B. D. (2022). The politics of progressivity: Court-ordered reforms, racial difference, and school finance fairness. American Educational Research Journal, 59(6), 1229-1264.

20 Candelaria, C. A., McNeill, S. M., & Shores, K. A. (2022). What Is a School Finance Reform? Uncovering the Ubiquity and Diversity of School Finance Reforms Using a Bayesian Changepoint Estimator. EdWorkingPaper No. 22-587. Annenberg Institute for School Reform at Brown University.

Baker, B., & Welner, K. (2011). School finance and courts: Does reform matter, and how can we tell?. Teachers College Record, 113(11), 2374-2414.

21 Baker, B., & Welner, K. (2011). School finance and courts: Does reform matter, and how can we tell?. Teachers College Record, 113(11), 2374-2414

22 Oberfield, Z. W., & Baker, B. D. (2022). The politics of progressivity: Court-ordered reforms, racial difference, and

school finance fairness. American Educational Research Journal, 59(6), 1229-1264.

23 Jackson, C. K., Johnson, R., & Persico, C. (2014). The effect of school finance reforms on the distribution of spending, academic achievement, and adult outcomes (No. w20118). National Bureau of Economic Research.

Candelaria, C. A., McNeill, S. M., & Shores, K. A. (2022). What Is a School Finance Reform? Uncovering the Ubiquity and Diversity of School Finance Reforms Using a Bayesian Changepoint Estimator. EdWorkingPaper No. 22-587. Annenberg Institute for School Reform at Brown University.

24 Baker, B. D. (2022). School finance and education equity: Lessons from Kansas. Harvard Education Press.

25 J. S. Coleman, E. Q. Campbell, C. F. Hobson, J. McPartland, A. M. Mood, et al., *Equality of Educational Opportunity* (Washington, DC: U.S. Office of Education, 1966).

26 J. S. Coleman, E. Q. Campbell, C. F. Hobson, J. McPartland, A. M. Mood, et al., *Equality of Educational Opportunity* (Washington, DC: U.S. Office of Education, 1966).

27 Garms, W. I., & Smith, M. C. (1970). Educational need and its application to state school finance. Journal of Human Resources, 304-317.

28 Baker, B. D., & Green, P. C. (2008). Conceptions of Equity and Adequacy in School Finance. Handbook of Research in Education Finance and Policy, 203.

29 To date, this piece has been cited over 5,500 times.

A later article by Hanushek, reiterating and updating his earlier findings, also shows up as widely cited in the Social Science Citation Index: E. A. Hanushek, "Assessing the Effects of School Resources on Student Performance: An Update," *Educational Evaluation and Policy Analysis* 19, no. 2 (1997): 141-164.

30 E. A. Hanushek, "Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24, no. 3 (1986): 1141-1177. A few years later, Hanushek paraphrased this conclusion in another widely cited article as "Variations in school expenditures are not systematically related to variations in student performance." E. A. Hanushek, "The Impact of Differential Expenditures on School Performance," *Educational Researcher* 18, no. 4 (1989): 45-62. Hanushek describes the collection of studies relating spending and outcomes as follows: "The studies are almost evenly divided between studies of individual student performance and aggregate performance in schools or districts. Ninety-six of the 147 studies measure output by score on some standardized test. Approximately 40 percent are based upon variations in performance within single districts while the remainder look across districts. Three-fifths look at secondary performance (grades 7-12) with the rest concentrating on elementary student performance" (fn #25).

31 Greenwald and colleagues explain: "Studies in the universe Hanushek (1989) constructed were assessed for quality. Of the 38 studies, 9 were discarded due to weaknesses identified in the decision rules for inclusion described below. While the remaining 29 studies were retained, many equations and coefficients failed to satisfy the decision rules we employed. Thus, while more than three quarters of the studies were retained, the number of coefficients from Hanushek's universe was reduced by two thirds" (p. 363). Greenwald and colleagues further explain that: "Hanushek's synthesis method, vote counting, consists of categorizing, by significance and direction, the relationships between school resource inputs and student outcomes (including but not limited to achievement). Unfortunately, vote-counting is known to be a rather insensitive procedure for summarizing results. It is now rarely used in areas of empirical research where sophisticated synthesis of research is expected" (p. 362). Hanushek (1997) provides his rebuttal to some of these arguments, and Hanushek returns to his "uncertainty" position: "The close to 400 studies of student achievement demonstrate that there is not a strong or consistent relationship between student performance and school resources, at least after variations in family inputs are taken into account" (p. 141). E. A. Hanushek, "Assessing the Effects of School Resources on Student Performance: An Update," *Educational Evaluation and Policy Analysis* 19, no. 2 (1997): 141-164. See also E. A. Hanushek, "Money Might Matter Somewhere: A Response to Hedges, Laine and Greenwald," *Educational Researcher* 23 (May 1994): 5-8.

32 R. Greenwald, L. Hedges and R. Laine, "The Effect of School Resources on Student Achievement," *Review of Educational Research* 66, no. 3 (1996): 361-396.

33 H. Wenglinsky, "How Money Matters: The Effect of School District Spending on Academic Achievement," *Sociology of Education* 70, no. 3 (1997): 221-237.

34 S. Konstantopoulos and G. Borman, "Family Background and School Effects on Student Achievement: A Multilevel Analysis of the Coleman Data," *Teachers College Record* 113, no. 1 (2011): 97-132.

35 G. D. Borman and M. Dowling, "Schools and Inequality: A Multilevel Analysis of Coleman's Equality of Educational Opportunity Data," *Teachers College Record* 112, no. 5 (2010): 1201-1246.

36 C. Taylor. "Does Money Matter? An Empirical Study Introducing Resource Costs and Student Needs into Educational Production Function Analysis," in *Developments in School Finance, 1997*, ed. W. J. Fowler Jr. (Washington, DC: U.S. Department of Education, National Center for Education Statistics, 1998), http://nces.ed.gov/pubs98/98212.pdf#page=83.

37 B. D. Baker, "Can Flexible Non-Linear Modeling Tell Us Anything New about Educational Productivity?," *Economics of Education Review* 20, no. 1 (2001): 81-92; D. N. Figlio, "Functional Form and the Estimated Effects of School Resources," *Economics of Education Review* 18, no. 2 (1999): 242-252; and J. Dewey, T. Husted and L. Kenny, "The Ineffectiveness of School Inputs: A Product of Misspecification," *Economics of Education Review* 19, no. 1 (2000): 27-45.

38 Specifically, Dewey and colleagues explain that many previous studies attempting to distill school resource effects on student outcomes concurrently correct for the economic background of

students. However, the economic background measures, such as family income, are also strong determinants of the demand for schooling resources. Thus, including the two simultaneously in regression models violates both conceptual appropriateness (resource levels are endogenous to family characteristics) and statistical properties associated with those conceptual problems (that the error term is correlated with the school input measures, requiring a different statistical approach). Dewey and colleagues review the previous studies summarized by Hanushek, identifying that several suffer from this problem and that those that do tend to understate the influence of resources. Then Dewey and colleagues estimate alternative production functions.

We conducted our own empirical analysis using the Project Talent student-level dataset from 1960 and pooled state data for 1987-1992. In regressions from both datasets that were not plagued by misspecification, there is evidence that each school input had an impact on achievement (p. 42). Figlio's study of alternative specifications of the "shape" of the relationship between money and outcomes raises similar issues about previous literature, including studies summarized by Hanushek, as does Corrine Taylor's analysis that applies adjustments for the costs of hiring teachers. Indeed, many of the same studies considered rigorous enough for inclusion in Greenwald and colleagues analyses also suffer from the problems addressed by Husted and Kenny, and by Taylor (geographic cost adjustment) and Figlio. But, note that in each case, Dewey and colleagues, Taylor and Figlio find that when applying functional form and labor cost corrections, they tend to find stronger effects of schooling resources—specifically money. So, one might then argue that Greenwald and colleagues' decisively positive findings are, in fact, understated. In conducting this review, we went back to a handful of the original studies summarized by Hanushek (1986) and listed in the sources note to Table 8 of that article. Several were not easily accessible, having been non-peer-reviewed reports and doctoral theses. But among those available, consistent with the findings of Husted and Kenny, none attempted to account for the endogeneity of expenditures, often either evaluating simple correlations between spending and outcome measures (thus suffering significant omitted variables bias) or including a spending measure alongside determinants of spending. Arguably, teacher characteristics, including teacher salaries, are also endogenous to local demand factors. Original Hanushek studies reviewed include:

A. Boardman, O. Davis and P. Sanday, "A Simultaneous Equations Model of the Educational Process," *Journal of Public Economics* 7, no. 1 (1977): 23-49. This study does not explore expenditures directly but does include measures of schooling facilities and teacher characteristics, but not salary. Thus, regional cost variation is less (or not) for the value of teacher salaries or education spending is less at issue. The authors of this study find that "many educational outputs jointly determine one another. Also, the results suggest that school and teacher variables have important effects on educational outcomes" (p. 23).

G. E. Johnson and F. P. Stafford, "Social Returns to Quantity and Quality of Schooling," *Journal of Human Resources* 8, no. 2 (1973): 139-155. In this study, the authors find "high but diminishing marginal returns to investment in expenditures per pupil per year" (p. 139). This is among the studies that arguably understates the sensitivity of expenditures to outcomes by inclusion of the spending measure (natural log of expenditures) in the model with determinants of expenditure (family socioeconomic status). In addition, the model uses a national sample but fails to control for regional variation in the value of expenditures.

C. R. Link and E. C. Ratledge, "Social Returns to Quantity and Quality of Education: A Further Statement," *Journal of Human Resources* 10, no. 1 (1975): 78-89. Link and Ratledge find: "Large but diminishing returns to incremental expenditures are observed" (p. 78). Link and Ratledge also use national survey data (National Longitudinal Study of the Labor Force). For the expenditure measure, like the above study, they use a measure of the 1968 district-level per-pupil expenditures (natural logarithm) and also do not correct for regional variation, though some of the urbanicity variables included may capture a portion of this variation (unintentionally). The endogeneity problems are less clear in this study, because in place of controlling for direct demand determinants (family income, education), the authors control for individual IQ. However, IQ is arguably simultaneously determined with education spending, both IQ and school spending being a function of parental economic status and education level. Sensitive to this point, the authors explore direct and indirect effects of IQ, years of education, and expenditures.

R. Raymond, "Determinants of the Quality of Primary and Secondary Public Education in West Virginia," *Journal of Human Resources* 3, no. 4 (1968): 450-470. Raymond studied 5,000 students in West Virginia. Raymond did not explore per-pupil expenditures but did explore several teacher salary measures, but does not correct for regional variation in the value of those salaries across West Virginia. Raymond finds salaries to be associated with output measures of quality.

T. I. Ribich and J. L. Murphy, "The Economic Returns to Increased Educational Spending," *Journal of Human Resources* 10, no. 1 (1975): 56-77. Ribich and Murphy used data from the national Project Talent survey. Ribich and Murphy found: "School expenditures are found to influence how many years of schooling an individual eventually receives, and the chief effect of spending differences on lifetime income is found to work through this school continuation link" (p. 56). Ribich and Murphy partly (though far from completely) correct for regional differences in the value of expenditures by including region variables. But, regression estimates likely suffer endogeneity addressed by Dewey, Husted, and Kenny (including both family socioeconomic measures and expenditures alongside one another). Interestingly, the authors instead attribute the insensitivity of their outcome measures to spending (when directly estimated including all regions) to regional differences, specifically racial differences within southern states.

F. Welch, "Measurement of the Quality of Schooling," *American Economic Review* 56 (1966): 379-

392. This study explored the return to elementary and secondary schooling of the male rural farm population in 1959, focusing on those who had not attended college in an effort to isolate differences in elementary and secondary schooling quality. This study is problematic on a number of levels when viewed in hindsight. First, the ultimate analysis of factors associated with the quality of schooling is aggregated to the state level (and noted by the author as a significant limitation). Second, expenditure measures are included in models with (a) potential determinants of expenditures (racial composition, labor composition, enrollment per secondary school); and (b) schooling resources dependent on expenditures (salaries, staff per 100 pupils) (see regression output in Table 4, p. 390). Further, expenditures are not adjusted for regional differences in value, nor are salaries.

39 In tangentially related work, Hanushek, Rivkin, and Taylor (1996) explore the influence of aggregation bias and omitted variables on estimates of the relationship between teacher characteristics and student outcomes, using data from the High School and Beyond survey. They find that at higher levels of aggregation, studies tend to overstate the strength of the relationship between resources and student outcomes, but raise the most significant concerns about studies using data aggregated to the state level with crude aggregate state level measures of student and population characteristics, far beyond the aggregation of most recent studies. E. A. Hanushek, S. Rivkin and L. L. Taylor, "Aggregation Bias and the Estimated Effects of School Resources," *Review of Economics and Statistics* 78, no. 4 (1996): 611-627. Along these lines, there does exist a separate body of literature that endeavors to prove that education spending is not associated with student outcomes by making national aggregate comparisons of spending and outcomes. That is, by showing that on average, countries that spend more per pupil don't perform better on international assessments. See, for example, H. J. Walberg, "Spending More While Learning Less," *Fordham Report* 2, no. 6 (1998). These studies suffer sufficiently from aggregation issues to be of little importance to the discussion herein. While aggregation might lead to overstating the money-outcome relationship in some studies, these studies also suffer from numerous substantial measurement problems regarding both input and outcome measures. For example, education spending data are simply not directly comparable across nations, partly because they include vastly different programs and services (athletics, arts, special education) as well as other specific expenses, such as health insurance costs for U.S. school employees, that may be covered via other government programs in other nations.

40 Oreskes, N., & Conway, E. M. (2011). Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming. Bloomsbury Publishing USA.

41 John Perry (2010) Book Review: Merchants of Doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming by Naomi Oreskes and Erik Conway. Bulletin of the American Meteorological Society , Vol. 91, No. 11

42 Lewandowsky, S., Ecker, U. K., Cook, J., Van Der Linden, S., Roozenbeek, J., & Oreskes, N. (2023). Misinformation and the epistemic integrity of democracy. Current Opinion in Psychology, 101711.

43 Rosen, L. (2020). Expert Testimony in the Social Sciences: A Historical Overview of Contemporary Issues. Law and History Review, 38(1), 123-142.

44 As with other bodies of literature on the effectiveness of schooling resources, the research on state school finance reforms is a mixed bag in terms of analytic rigor. Secondhand references to dreadful failures following massive infusions of new funding can often be traced to methodologically inept, anecdotal tales of desegregation litigation in Kansas City, Mo., or to the court-ordered financing of urban districts in New Jersey.

Two reports from Cato are illustrative: P. Ciotti, *Money and School Performance: Lessons from the Kansas City Desegregations Experience*, Cato Policy Analysis 298 (1998); and D. Coate and J. VanderHoff, "Public School Spending and Student Achievement: The Case of New Jersey," *Cato Journal* 19, no. 1 (1999): 85-99. Edspresso, *New Jersey Learns Kansas City's Lessons the Hard Way* (2006), retrieved Oct. 23, 2009, from www.edspresso.com/index.php/2006/10/new-jersey-learns-kansas-citys-lessons-the-hard-way-2.

45 In 2009, Eric Hanushek and a consulting defense attorney for states facing school funding challenges, Alfred Lindseth of Sutherland Asbill & Brennan, produced a book in which one chapter is dedicated to trying to prove that court-ordered school funding reforms in New Jersey, Wyoming, Kentucky, and Massachusetts resulted in few or no measurable improvements. These conclusions, however, are based on little more than a series of graphs of student achievement on the National Assessment of Educational Progress in 1992 and 2007. The authors show little change in these states' scores and conclude that the reforms didn't work.

The authors assume that, during this period, each of the four states infused substantial additional funds into public education in response to judicial orders, and that these funds were targeted at low-income and minority students. They also necessarily assume that in all other states that serve as a comparison group, similar changes did not occur. Yet they validate neither assertion.

E. A. Hanushek and A. Lindseth, *Schoolhouses, Courthouses and Statehouses* (Princeton, NJ: Princeton University Press, 2009). See also http://edpro.stanford.edu/Hanushek/admin/pages/files/uploads/06_EduO_Hanushek_g.pdf.

46 In contrast, Welner and Baker review several studies applying more rigorous and appropriate methods for evaluating the influence of state school finance reforms. Among these analyses is one national study by Card and Payne (2002) that evaluates whether changes in spending inequality generally lead to changes in outcome inequality.

For additional discussion of the strengths and weaknesses of this particular study, see Baker and Welner (2011).

47 The authors measure both the extent and the timing of changes in each. These analyses, while imperfect, rise to a level far above those conducted by Hanushek and Lindseth. Card and Payne found "evidence that equalization of spending levels leads to a narrowing of test score outcomes across family background groups" (p. 49).

D. Card and A. A. Payne, "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores," *Journal of Public Economics* 83, no. 1 (2002): 49-82. See Baker and Welner (2011) for a more thorough discussion of the Card and Payne analysis and its strengths and weaknesses.

48 Baker, B. D., & Weber, M. (2016). Deconstructing the Myth of American Public Schooling Inefficiency. Albert Shanker Institute.

49 Baker, B. D. (2021). Educational inequality and school finance: Why money matters for America's students. Harvard Education Press.

50 Hanushek, E. A. (2006). Courting failure: How school finance lawsuits exploit judges' good intentions and harm our children (Vol. 551). Hoover Press.

51 Handel, D. V., & Hanushek, E. A. (2022). US School Finance: Resources and Outcomes (No. w30769). National Bureau of Economic Research.

52 Armor, D. J., & Ossian, A. M. (2024). 2 The State of School Resource Research. School Resources, the Achievement Gap, and the Law: Reconsidering School Finance, Policies, and Resources in US Education Policy.

53 Minorini, P. A., & Sugarman, S. D. (1999). School finance litigation in the name of educational equity: Its evolution, impact, and future. Equity and adequacy in education finance: Issues and perspectives, 34-71.

54 Clune, W. H. (1994). The shift from equity to adequacy in school finance. Educational Policy, 8(4), 376-394.

55 Berne, R., & Stiefel, L. (1999). Concepts of school finance equity: 1970 to the present. Equity and adequacy in education finance: Issues and perspectives, 7-33.

56 Steven Ruggles, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler. IPUMS USA: Version 14.0 [dataset]. Minneapolis, MN: IPUMS, 2023. https://doi.org/10.18128/D010.V14.0

57 Morgan, S. L., & Winship, C. (2015). Counterfactuals and causal inference. Cambridge University Press.

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Journal of economic perspectives, 24(2), 3-30.

58 Jackson, C. K., & Mackevicius, C. L. (2024). What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States. American Economic Journal: Applied Economics, 16(1), 412-446.

59 As state specific longitudinal studies of school finance reforms were beginning to appear in the early 2000s, David Figlio (2004) noted that the influence of state school finance reforms on student outcomes is perhaps better measured *within* states over time, explaining that national studies of the type attempted by Card and Payne (2002) confront problems that include (a) the enormous diversity in the nature of state aid reform plans; and (b) the paucity of national student performance data.

D. N. Figlio, "Funding and Accountability: Some Conceptual and Technical Issues in State Aid Reform," in *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. J. Yinger (MIT Press, 2004), 87-111.

This synopsis of Figlio's main points actually comes from an earlier chapter in the same volume, by volume editor John Yinger: J. Yinger, "State Aid and the Pursuit of Educational Equity: An Overview," in *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. J. Yinger (MIT Press, 2004), 39.

60 Jackson, C. K., & Mackevicius, C. L. (2023). What impacts can we expect from school spending policy? Evidence from evaluations in the US. *American Economic Journal: Applied Economics*.

61 Dewey et al. (2024) note a finding by Doty et al. (2022) who estimate the present value of lifetime earnings of $1.2 million for the average individual who completes at least high school; and a finding by Watts (2020) who finds that a one-standard deviation increase in test score achievement is associated with a 12 percent increase in lifetime earnings between the ages of 33 and 50.

Dewey, D., Fahle, E., Kane, T. J., Reardon, S. F., Staiger, D. O. (2024). Federal pandemic relieve and academic recovery. https://educationrecoveryscorecard.org/wp-content/uploads/2024/06/June2024ERS-Report.pdf

Doty, E., Kane, T. J., Patterson, T., & Staiger, D. O. (2022). What do changes in state test scores imply for later life outcomes? NBER Working Paper No. 30701, National Bureau of Economic Research, Cambridge, MA December.

Watts, T. (2020). Academic achievement and economic attainment: Reexamining associations between test scores and long-run earnings. AERA Open 6 (2):1-16. https://doi.org/10.1177/2332858420928985.

62 Jackson, C. K., & Persico, C. (2023). Point column on school spending: Money matters. *Journal of Policy Analysis and Management*, *42*(4), 1118-1124.

63 McGee, J. B. (2023). Yes, money matters, but the details can make all the difference. *Journal of Policy Analysis and Management*, *42*(4), 1125-1132.

64 Jackson, C. K., & Persico, C. (2023). Counterpoint–Don't let the perfect be the enemy of the good. *Journal of Policy Analysis and Management*, *42*(4), 1133-1135.

65 Handel, D. V., & Hanushek, E. A. (2023). *Contexts of Convenience: Generalizing from Published Evaluations of School Finance Policies* (No. w31653). National Bureau of Economic Research.

66 Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, *131*(1), 157-218.

D. Card and A. A. Payne, "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores," *Journal of Public Economics* 83, no. 1 (2002): 49-82. See Baker and Welner (2011) for a more thorough discussion of the Card and Payne analysis and its strengths and weaknesses.

67 Abott, C., Kogan, V., Lavertu, S., & Peskowitz, Z. (2020). School district operational spending and student outcomes: Evidence from tax elections in seven states. *Journal of Public Economics*, *183*, 104142.

68 Miller, C. L. (2018). The effect of education spending on student achievement: Evidence from property values and school finance rules. Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association, 111, 1–121. https://www.jstor.org/stable/26939356

69 Rothstein, J., & Schanzenbach, D. W. (2022). Does money still matter? Attainment and earnings effects of post-1990 school finance reforms. *Journal of Labor Economics*, *40*(S1), S141-S178.

70 Candelaria, C. A., & Shores, K. A. (2019). Court-ordered finance reforms in the adequacy era: Heterogeneous causal effects and sensitivity. Education Finance and Policy, 14 (1), 31–60.

71 Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, *10*(2), 1-26.

72 Rothstein, J., & Schanzenbach, D. W. (2022). Does money still matter? Attainment and earnings effects of post-1990 school finance reforms. *Journal of Labor Economics*, *40*(S1), S141-S178.

73 Rauscher, E., & Shen, Y. (2022). Variation in the relationship between school spending and achievement: Progressive spending is efficient. *American Journal of Sociology*, *128*(1), 189-223

74 Brunner, E., Hyman, J., & Ju, A. (2020). School finance reforms, teachers' unions, and the allocation of school resources. *Review of Economics and Statistics*, *102*(3), 473-489.

75 The two recent studies are:

Dewey, D., Fahle, E., Kane, T. J., Reardon, S. F., Staiger, D. O. (2024). Federal pandemic relieve and academic recovery. https://educationrecoveryscorecard.org/wp-content/uploads/2024/06/June2024ERS-Report.pdf

Goldhaber, D., & Falkan, G. (2024). ESSER and student achievement: Assessing the impacts of the largest one-time federal investment in K12 schools. CALDER Working Paper No. 301-0624. https://caldercenter.org/publications/esser-and-student-achievement-assessing-impacts-largest-one-time-federal-investment-k12

As described in Dewey et al. (2024), Goldhaber and Falkan (2024), and Gordon and Reber (2023), the federal Title I program is allocated according to a series of formulas, based primarily on the number and percent of funding-eligble children (students from lower-income households) that live within a school district's geographic area, based on annual U.S. Census estimates. However, the data are lagged such that 2017-18 values for funding-eligble children were used to determine Title I allocations for the 2020-21 school year, and the formula includes state minimums and adjustments based on state finance policies. As a result, per-pupil Title I allocations to school districts in a given year are not perfectly aligned with the actual student poverty level that year. This somewhat random misalignment allows researchers to estimate the causal impact of Title I dollars. Gordon (2002) uses a different approach, leveraging the adjustments to Title I allocations related to the release of decennial census data. For additional background on Title I formula allocations see:

Gordon, N., & Reber, S. (2023). *Title I of ESEA: How the formulas work* (Understanding and Improving Title I of ESEA). All4Ed. https://all4ed.org/publication/title-i-of-esea-how- the-formulas-work/

Gordon, N. (2002). Do federal grants boost school spending? Evidence from Title I. https://users.nber.org/~confer/2002/urcw02/gordon.pdf

76 Miller, C. L. (2018). The Effect of education spending on student achievement: Evidence from property values and school finance rules. Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association, 111, 1–121. https://www.jstor.org/stable/26939356

77 Abott, C., Kogan, V., Lavertu, S., & Peskowitz, Z. (2020). School district operational spending and student outcomes: Evidence from tax elections in seven states. Journal of Public Economics, 183, 104142.

78 Rauscher, E., & Shen, Y. (2022). Variation in the relationship between school spending and achievement: Progressive spending Is efficient. *American Journal of Sociology*, *128*(1), 189-223.

79 Jackson, C. K., Wigger, C., & Xiong, H. (2021). Do school spending cuts matter? Evidence from the Great Recession. *American Economic Journal: Economic Policy*, *13*(2), 304-335.

80 D. N. Figlio, "Funding and Accountability: Some Conceptual and Technical Issues in State Aid Reform," in *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. J. Yinger (MIT Press, 2004), 87-111. This synopsis of Figlio's main points actually comes from an earlier chapter in the same volume, by volume editor John Yinger: J. Yinger, "State Aid and the Pursuit of Educational Equity: An Overview," in *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. J. Yinger (MIT Press, 2004), 39.

81 Baker, B. D. (2022). School finance and education equity: Lessons from Kansas. Harvard Education Press.

Rauscher, E. (2020). Does money matter more in the country? Education funding reductions and achievement in Kansas, 2010–2018. AERA open, 6(4), 2332858420963685.

82 J. Roy, "Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence from Michigan," *Education Finance and Policy* 6, no. 2 (2011): 137-167.

L. Papke, "The Effects of Spending on Test Pass Rates: Evidence from Michigan," *Journal of Public Economics* 89, no. 5-6 (2005): 821-839;

83 Hyman, J. (2017). Does money matter in the long run? Effects of school spending on educational attainment. American Economic Journal: Economic Policy, 9(4), 256–280. https://doi.org/10.1257/pol.20150249

84 In contrast to my other research on the impact of school resources, Hyman (2017) finds additional funds had the largest impacts on postsecondary outcomes for students attending lower-poverty schools as well as suburban and urban schools. The results also showed that, on average, districts targeted a greater amount of funds to these schools.

85 Baron, E. J., Hyman, J., & Vasquez, B. (2024). Public school funding, school quality, and adult crime. *Review of Economics and Statistics*, 1-46.

86 T. A. Downes, J. Zabel and D. Ansel, *Incomplete Grade: Massachusetts Education Reform at 15* (Boston, MA: MassINC, 2009). https://massinc.org/wp-content/uploads/2009/05/incomplete_grade_executive-summary.ashx_.pdf

J. Guryan, "Does Money Matter? Estimates from Education Finance Reform in Massachusetts," Working Paper No. 8269 (Cambridge, MA: National Bureau of Economic Research, 2001). While this paper remains an unpublished working paper, the advantage of Guryan's analysis is that he models the expected changes in funding at the local level as a function of changes to the school finance formula itself, through what is called an instrumental variables or two-stage least squares approach. Then, Guryan evaluates the extent to which these policy-induced variations in local funding are associated with changes in student outcomes. Across several model specifications, Guryan finds increased outcomes for students at grade 4 but not at grade 8. A counter study by the Beacon Hill Institute suggests that reduced class size and/or increased instructional spending either has no effect on or actually worsens student outcomes. S. Jaggia and V. Vachharajani, *Money for Nothing: The Failures of Education Reform in Massachusetts* (2004), www.beaconhill.org/BHIStudies/EdStudy5_2004/BHIEdStudy52004.pdf.

P. Nguyen-Hoang and J. Yinger,"Education Finance Reform, Local Behavior, and Student Performance in Massachusetts," *Journal of Education Finance* 39, no. 4 (2014): 297-322.

87 T. A. Downes, "School Finance Reform and School Quality: Lessons from Vermont," in *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. J. Yinger (Cambridge, MA: MIT Press, 2004).

88 Two studies of school finance reforms in New Jersey also merit some attention, in part because they directly refute findings of Hanushek and Lindseth and of the earlier Cato study, and do so with more rigorous and detailed methods. The first, by Alex Resch of the University of Michigan (doctoral dissertation in economics), explored in detail the resource allocation changes during the scaling-up period of school finance reform in New Jersey. Resch found evidence suggesting that New Jersey Abbott districts "directed the added resources largely to instructional personnel" (p. 1), such as additional teachers and support staff. She also concluded that this increase in funding and spending improved the achievement of students in the affected school districts. Looking at the statewide 11th-grade assessment ("the only test that spans the policy change"), she found "that the policy improves test scores for minority students in the affected districts by one-fifth to one-quarter of a standard deviation" (p. 1). Goertz and Weiss (2009) also evaluated the effects of New Jersey school finance reforms but did not attempt a specific empirical test of the relationship between funding level and distributional changes and outcome changes. Thus, their findings are primarily descriptive. Goertz and Weiss explain that on state assessments, achievement gaps closed substantially between 1999 and 2007, the period over which Abbott funding was most significantly scaled up. A. M. Resch, *Three Essays on Resources in Education* (dissertation) (Ann Arbor: University of Michigan, Department of Economics, 2008), retrieved Oct. 28, 2009, from http://deepblue.lib.umich.edu/bitstream/2027.42/61592/1/aresch_1.pdf; and M. Goertz and M. Weiss, *Assessing Success in School Finance Litigation: The Case of New Jersey* (New York City: Campaign for Educational Equity, Teachers College, Columbia University, 2009). The authors explain: "State Assessments: In 1999 the gap between the Abbott districts and all other districts in the state was over 30 points. By 2007 the gap was down to 19 points, a reduction of 11 points or 0.39 standard deviation units. The gap between the Abbott districts and the high-wealth districts fell from 35 to 22 points. Meanwhile performance in the low-, middle-, and high-wealth districts essentially remained parallel during this eight-year period" (Figure 3, p. 23).

89 J. Deke, "A Study of the Impact of Public School Spending on Postsecondary Educational Attainment Using Statewide School District Refinancing in Kansas," *Economics of Education Review* 22, no. 3 (2003): 275-284.

90 Johnson, R. (2023). School funding effectiveness: Evidence from California's local control funding formula. Learning Policy Institute. https://doi.org/10.54300/529.194

Johnson, R. C., & Tanner, S. (2018) Money and freedom: The impact of California's school finance reform on academic achievement and the composition of district spending. Getting Down To Facts II. https://gettingdowntofacts.com/publications/money-and-freedom-impact-californias-school-finance-reform-academic-achievement-and

Lafortune, J., (2021). Targeted K–12 funding and student outcomes: Evaluating the Local Control Funding Formula. Policy Brief. Public Policy Institute of California. https://www.ppic.org/publication/targeted-k-12-funding-and-student-outcomes/

Lee and Fuller (2017) find "schools in districts receiving concentration grants during the initial two years of Local Control Funding did engage in organizational change that parallels gains in pupil achievement, compared with schools in almost identical districts not receiving concentration grants. These benefits were largely experienced by Latino students and not by other groups at significant levels" (p. 1).

Lee, J., & Fuller, B. (2017). How decentralizing school finance may narrow achievement gaps. Institute of Human Development Working Paper. https://gse.berkeley.edu/sites/default/files/gse-archive-10_0/CASchoolFinance_Lee_Fuller_June2017_workingpaper.pdf

Johnson and Tanner (2018) found that "a $1,000 increase in district per-pupil spending experienced in grades 10-12 leads to a 5.9 percentage-point increase in high school graduation rates on average among all children, with similar effects by race and poverty. On average among poor children, a $1,000 increase in district per-pupil spending experienced in 8th through 11th grades leads to a 0.19 standard deviation increase in math test scores, and a 0.08 standard-deviation increase in reading test scores in 11th grade."

Similarly, Fan and Liang (2020) find "positive effects of LCFF on student achievement measured by percentage of graduates meeting university of California/California State University (UC/CSU) entrance requirements, and the effects are causal and non-random. Results suggest a delayed effect that became significantly positive two years after LCFF implementation. The positive effects likely result from LCFF-induced additional funding, while the effects of non-funding-related aspects of the reform tend to be arbitrary in the short run. In addition, we find that the marginal effect on student achievement tends to be higher in high-poverty school districts, relative to low-poverty school districts. The empirical results of a significantly positive effect of LCFF, especially in high-poverty school districts, support reduction of educational inequity as one LCFF target."

Fan, Q., & Liang, J. (2020). The effects of California's school finance reform: Empirical evidence from the Local Control Funding Formula (LCFF). Journal of Education Finance, 46(2), 140-157.

91 Lee and Fuller (2022) find "Schools that enjoyed greater funding modestly reduced average class size and the count of teaching periods assigned to staff in five subsequent years. Yet, high-poverty schools receiving higher budget augmentations more often assigned novice teachers to English learners (ELs) and hosted declining shares of courses that qualified graduates for college admission. Mean achievement climbed overall, but EL and poor students fell further behind in schools receiving greater funding."

Lee, J. H., & Fuller, B. (2022). Does progressive finance alter school organizations and raise achievement? The case of Los Angeles. Educational Policy, 36(3), 587-623.

Knight and Mendoz (2019) show LCFF led to an increase in the targeting of state and local revenue ot higher-poverty districts; however, Lafortune et al., (2023) find that "Schools with more high-need students generate more funding, but most districts do not spend these additional dollars in the same proportion" (p. 1).

Knight, D. S., & Mendoza, J. (2019). Does the measurement matter? Assessing alternate approaches to measuring state school finance equity for California's Local Control Funding Formula. AERA Open, 5(3), 2332858419877424.

Lafortune, J., Herrera, J., Gao, N., & Barton, S. (2023). Examining the reach of targeted school funding. Policy Brief. Public Policy Institute of California. https://www.ppic.org/publication/examining-the-reach-of-targeted-school-funding/

92 Candelaria, C., Sun, M., Knight, D., Almasi, P., Liu, A., & Kabourek, S. (2023, November). School finance reform implementation and impacts on resource allocation: A case study of funding formula changes in Washington State. In 2023 APPAM Fall Research Conference. APPAM.

Fujioka, K., & Knight, D. S. (forthcoming). McCleary at twelve: Examining policy designs following court-mandated school finance reform in Washington State. West Education Law Reporter.

Knight, D. S., Candelaria, C. A, Sun, M., Almasi, P., Kabourek, S., & Liu, A. (2023, March). More for the wealthy? Unintended consequences of court-mandated school finance reforms. Paper presented at the Annual Conference of the Association for Educational Finance and Policy, Denver, CO.

Sun, M., Candelaria, C. A., Knight, D., LeClair, Z., Kabourek, S. E., & Chang, K. (2024). The Effects and Local Implementation of School Finance Reforms on Teacher Salary, Hiring, and Turnover. Educational Evaluation and Policy Analysis, 01623737231213880

93 In a study examining a finance reform in the Netherlands, Leuven and colleagues used regression discontinuity to compared outcomes for districts on either side of a funding cut off, finding zero positive impacts for funding policies that mandated spending on computers.

E. Leuven, M. Lindahl, H. Oosterbeek and D. Webbink, "The Effect of Extra Funding for Disadvantaged Pupils on Achievement," *Review of Economics and Statistics* 89, no. 4 (2007): 721-736.

94 References to "diminishing returns" in various articles:

Rauscher: Evidence of diminishing returns by previous investment documents another way that schools increase equality and establishes another argument for progressive school funding: efficiency.

Abbott/Kogan: Effects on outcomes driven by increases in spending below the median.

Johnson (CA LCFF): Speculates that states that are better funded to begin with may not see similar returns as CA under LCFF.

Baron (WI): Suggests perhaps diminishing returns to facilities investments.

95 Tyner (2023) argues that school investments suffer from diminishing returns, referring several research studies. Jackson and Mackevicius provide a more comprehensive look at the impact of money across settings, demonstrating limited evidence for diminishing returns to educational investments.

Tyner, A. (2023). Think again: Is education funding in American still unequal? Thomas B. Fordham Institute. https://fordhaminstitute.org/national/research/think-again-education-funding-america- still-unequal

Jackson, C. K., & Mackevicius, C. L. (2024). What impacts can we expect from school spending policy? Evidence from evaluations in the U.S. American Economic Journal: Applied Economics. https://doi.org/10.1257/app.20220279

96 Jackson, C. K., Wigger, C., & Xiong, H. (2021). Do school spending cuts matter? Evidence from the Great Recession. American Economic Journal: Economic Policy, 13(2), 304-335.

Knight, D. S. (2017). Are high-poverty school districts disproportionately impacted by state funding cuts? School finance equity following the Great Recession. Journal of Education Finance, 169-194.

97 Baker, B. D., Di Carlo, M., & Weber, M. (2024). The adequacy and fairness of state school finance systems. 6th Edition. Albert Shanker Institute.

98 Bartanen, B., Grissom, J. A., & Rogers, L. K. (2019). The impacts of principal turnover. Educational Evaluation and Policy Analysis, 41(3), 350-374.

Brown, C. H., & Knight, D. (2024). Student-to-school counselor ratios: understanding the history and ethics behind professional staffing recommendations and realities in the United States. Ethics & Behavior, 1-19.

Brown, C. H., & Knight, D. (2023). Staffing schools to support the classroom: Examining student-to-school-counselor ratios and academic student outcomes in Texas. Professional School Counseling, 27(1), 2156759X231165497.

Reback, R. (2010). Noninstructional spending improves noncognitive outcomes: Discontinuity evidence from a unique elementary school counselor financing system. Education Finance and Policy, 5(2), 105-137.

Reback, R. (2010). Schools' mental health services and young children's emotions, behavior, and learning. Journal of Policy Analysis and Management, 29(4), 698-725.

Reback, R. (2018). Investments in Student Health and Mental Health in California's Public Schools. Technical Report. Getting Down to Facts II. Policy Analysis for California Education, PACE.

Schroeder, K., Travers, J., & Smaldone, A. (2016). Are school nurses an overlooked resource in reducing childhood obesity? A systematic review and meta-analysis. Journal of school health, 86(5), 309-321.

Yoder, C. M. (2020). School nurses and student academic outcomes: An integrative review. The Journal of School Nursing, 36(1), 49-60.

Allen, G. (2003). The impact of elementary school nurses on student attendance. The Journal of School Nursing, 19(4), 225-231.

99 Jackson, C. K., Johnson, R. C., & Persico, C. (2015). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms (No. w20847). National Bureau of Economic Research.

100 "Memorandum Opinion and Entry of Judgment," *Gannon vs. State of Kansas*, 10C1569, January 11, 2013, http://www.shawneecourt.org/DocumentCenter/View/457.

101 Baker, B. D. (2012). Rearranging deck chairs in Dallas: Contextual constraints and within-district resource allocation in urban Texas school districts. Journal of Education Finance, 287-315.

102 Hanushek (1986) explains: "Thus the basic determinants of instructional expenditures in a district are teacher experience, teacher education and class size, and most studies, regardless of what other descriptors of schools might be included, will analyze the effect of these factors on outcomes" (p. 1160).

103 E. A. Hanushek, "Teacher Characteristics and Gains in Student Achievement: Estimation Using MicroData," *Econometrica* 61, no. 2 (1971): 280-288; C. T. Clotfelter, H. F. Ladd and J. L. Vigdor, "Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects," *Economics of Education Review* 26 (2007): 673-682; D. Goldhaber and D. Brewer, "Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity," *Journal of Human Resources* 332, no. 3 (1997): 505-523; R. G. Ehrenberg and D. J. Brewer, "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," *Economics of Education Review* 13, no. 1 (1994): 1-17; R. G. Ehrenberg and D. J. Brewer, "Did Teachers' Verbal Ability and Race Matter in the 1960s?," *Economics of Education Review* 14, no. 1 (1995): 1-21; C. Jepsen, "Teacher Characteristics and Student Achievement: Evidence from Teacher Surveys," *Journal of Urban Economics* 57, no. 2 (2005): 302-319; B. A. Jacob and L. Lefgren, "The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform," *Journal of Human Resources* 39, no. 1 (2004): 50-79; S. G. Rivkin, E. A. Hanushek and J. F. Kain, "Teachers, Schools, and Academic Achievement," *Econometrica* 73, no. 2 (2005): 471; and A. J. Wayne and P. Youngs, "Teacher Characteristics and Student Achievement Gains," *Review of Educational Research* 73, no. 1 (2003): 89-122. For a recent review of studies on the returns to teacher experience, see J. K. Rice, *The Impact of Teacher Experience: Examining the Evidence and Policy Implications* (National Center for Analysis of Longitudinal Data in Educational Research, 2010).

104 Add citations about salaries impacting who majors in education, applies for a teaching job, accepts a position.

105 R. J. Murnane and R. Olsen, "The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching: Evidence from Michigan," *Review of Economics and Statistics* 71, no. 2 (1989): 347-352.

106 D. N. Figlio, "Can Public Schools Buy Better-Qualified Teachers?," *Industrial and Labor Relations Review* 55 (2002): 686-699; D. N. Figlio "Teacher Salaries and Teacher Quality," *Economics Letters* 55 (1997): 267-271; and R. Ferguson, "Paying for Public Education: New Evidence on How and Why Money Matters," *Harvard Journal on Legislation* 28, no. 2 (1991): 465-498.

107 Ondrich, J., Pas, E., & Yinger, J. (2008). The determinants of

teacher attrition in upstate New York. Public Finance Review, 36(1), 112-144.

108 S. Loeb and M. Page, "Examining the Link Between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-Pecuniary Variation," *Review of Economics and Statistics* 82, no. 3 (2000): 393-408.

109 D. N. Figlio and K. Rueben, "Tax Limits and the Qualifications of New Teachers," *Journal of Public Economics* (April 2001): 49-71. See also T. A. Downes and D. N. Figlio, "Do Tax and Expenditure Limits Provide a Free Lunch? Evidence on the Link Between Limits and Public Sector Service Quality," *National Tax Journal* 52, no. 1 (1999): 113-128.

110 Kraft, M. A., & Lyon, M. A. (2024). The rise and fall of the teaching profession: Prestige, interest, preparation, and satisfaction over the last half century. (No. w32386). National Bureau of Economic Research and (EdWorkingPaper: 22-679). Annenberg Institute at Brown University: https://doi.org/10.26300/7b1a-vk92..

111 Gilmour, A. F., Theobald, R., & Nathan, D. J. (2023). Economic interventions as policy levers for attracting and retaining special education teachers. Handbook of Research on Special Education Teacher Preparation, 190-212.

112 Nguyen, T. D., Pham, L. D., Crouch, M., & Springer, M. G. (2020). The correlates of teacher turnover: An updated and expanded meta-analysis of the literature. Educational Research Review, 31, 100355.

113 Sun, M. Candelaria, C., Knight, D. S., et al. EEPA paper on teacher salaries.

114 Hendricks, M. D. (2015). Towards an optimal teacher salary schedule: Designing base salary to attract and retain effective teachers. Economics of Education Review, 47, 143-167.

Hendricks, M. D. (2014). Does it pay to pay teachers more? Evidence from Texas. Journal of Public Economics, 109, 50-63.

115 Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2005). Teacher bonuses and teacher retention in low performing schools: Evidence from the North Carolina $1,800 teacher bonus program (Working Paper). Maxwell Policy Research Symposium.

Clotfelter, C. T., Glennie, E. J., Ladd, H. F., & Vigdor, J. L. (2008a). Teacher bonuses and teacher retention in low-performing schools: Evidence from the North Carolina $1,800 Teacher Bonus Program. Public Finance Review, 36(1), 63–87. https://doi.org/10.1177/1091142106291662

Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2008b). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. Journal of Public Economics, 92, 1352–1370. doi:10.1016/j.jpubeco2007.07.003

Glazerman, S., Protik, A., Teh, B., Bruch, J., & Max, J. (2013). Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment. U.S. Department of Education. https://files.eric.ed.gov/fulltext/ED544269.pdf

Feng, L., & Sass, T. R. (2018). The impact of incentives to recruit and retain teachers in "hard- to-staff" subjects: Incentives to recruit and retain teachers. Journal of Policy Analysis and Management, 37(1), 112–135. https://doi.org/10.1002/pam.22037.

Cowan, J., & Goldhaber, D. (2018). Do bonuses affect teacher staffing and student achievement in high poverty schools? Evidence from an incentive for national board certified teachers in Washington State. Economics of Education Review, 65, 138–152. https://doi.org/10.1016/j.econedurev.2018.06.010

116 Feng, L., & Sass, T. R. (2018). The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. Journal of Policy Analysis and Management, 37(1), 112-135.

For additional evidence on the relationship between overall salary and retention, see Ondrich, Pas, and Yinger (2008), who find that "teachers in districts with higher salaries relative to non-teaching salaries in the same county are less likely to leave teaching and that a teacher is less likely to change districts when he or she teaches in a district near the top of the teacher salary distribution in that county."

Ondrich, J., Pas, E., & Yinger, J. (2008). The determinants of teacher attrition in upstate New York. Public Finance Review, 36(1), 112-144.

117 For major studies specifically on the topic of "merit pay," each of which generally finds no positive effects of merit pay on student outcomes, see S. Glazerman and A. Seifullah, *An Evaluation of the Teacher Advancement Program in Chicago: Year Two Impact Report* (Mathematica Policy Research Institute, 2010); M. G. Springer, D. Ballou, L. Hamilton, V. Le, J. R. Lockwood, D. McCaffrey, M. Pepper and B. Stecher, *Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching* (Nashville, TN: National Center on Performance Incentives at Vanderbilt University, 2010); and J. A. Marsh, M. G. Springer, D. F. McCaffrey, K. Yuan, S. Epstein, J. Koppich, N. Kalra, C. DiMartino and A. Peng, *A Big Apple for Educators: New York City's Experiment with Schoolwide Performance Bonuses, Final Evaluation Report* (RAND Corporation and Vanderbilt University, 2011). Some studies find positive effects, such as S. Dee and J. Wyckoff, "Incentives, Selection, and Teacher Performance: Evidence from IMPACT," Journal of Policy Analysis and Management 34, no. 2 (2015): 267-297 and Biasi, B. (2021). The labor market for teachers under different pay schemes. American Economic Journal: Economic Policy, 13(3), 63-102. A recent meta-analysis found a larger number of studies identifying positive effects.

118 K. Yuan, V. N. Le, D. F. McCaffrey, J. A. Marsh, L. S. Hamilton, B. M. Stecher and M. G. Springer, "Incentive Pay Programs Do Not Affect Teacher Motivation or Reported Practices: Results from Three Randomized Studies," *Educational Evaluation and Policy Analysis* 35, no. 1 (2013); M. G. Springer, D. Ballou, L. Hamilton, V. N. Le, J. R. Lockwood, D. F. McCaffrey and B. M. Stecher, "Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching (POINT)," *Society for Research on Educational Effectiveness* (2011); and M. G. Springer, J. F. Pane, V. N. Le, D. F., McCaffrey, S. F. Burns, L. S. Hamilton and B. Stecher, "Team Pay for Performance

Experimental Evidence from the Round Rock Pilot Project on Team Incentives," *Educational Evaluation and Policy Analysis* 34, no. 4 (2012): 367-390.

119 T. S. Dee and J. Wyckoff, "Incentives, Selection, and Teacher Performance: Evidence from IMPACT," *Journal of Policy Analysis and Management* 34, no. 2 (2015): 267-297.

120 R. Balch and M. G. Springer, "Performance Pay, Test Scores, and Student Learning Objectives," *Economics of Education Review* 44 (2015): 114-125.

121 A. J. Sojourner, E. Mykerezi and K. L. West, "Teacher Pay Reform and Productivity Panel Data Evidence from Adoptions of Q-Comp in Minnesota," *Journal of Human Resources* 49, no. 4 (2014): 945-981.

122 R. G. Fryer Jr., S. D. Levitt, J. List and S. Sadoff, *Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment* (no. w18237) (National Bureau of Economic Research, 2012).

123 Goodman, S. F., & Turner, L. J. (2013). The design of teacher incentive pay and educational outcomes: Evidence from the New York City bonus program. Journal of Labor Economics, 31(2), 409-420.

Imberman, S. A., & Lovenheim, M. F. (2015). Incentive strength and teacher productivity: Evidence from a group-based teacher incentive pay system. Review of Economics and Statistics, 97(2), 364-386.

124 Gilpin and Kaganovich (2011) propose a general equilibrium model of teacher quantity and quality adopting the premise that teachers' relative wages (to other sectors) are critical to maintaining a quality teaching workforce. Additionally, compression of salaries (at the high end) may reduce retention and recruitment of talented teachers. Illustrated in their data, the long-term increase in teacher quantity has led to lagging wage competitiveness, thus potentially compromising labor quality. But so too has growth in wages of competing sectors. They explain that a rise in premium for high ability will outpace that for the average. These increase costs are not offset by technological change, hence an additional downward pressure on the "real" quality of education inputs. The cost of high-quality labor is increasing faster than resources required to pay the premium to simply maintain quality, and there exist no viable technological substitutes to offset those increases. As such, we can expect the real quality of educational inputs to decline. Because of the rise in premium for high ability, Gilpin and Kaganovich assert that "countering this trend would therefore require an increase in the share of GDP spent on basic education, assuming that the institutional setup of the school system remains unchanged" (428). In other words, because talent is becoming more expensive more rapidly in other sectors, more investment, as a share of GDP, may be required merely to maintain education quality. That said, this theoretical exposition, while built on much the same research base as we review herein and previously, is not fully vetted in the present article.

Gilpin, G., & Kaganovich, M. (2012). The quantity and quality of teachers: Dynamics of the trade-off. *Journal of Public Economics*, *96*(3-4), 417-429.

125 J. Rothstein, "Teacher Quality Policy When Supply Matters," *Document de treball de IEB* 35 (2012): 1-65; and J. Rothstein, *Teacher Quality Policy When Supply Matters* (no. w18419) (National Bureau of Economic Research, 2012).

126 Dee and Wyckoffs found higher IMPACT scores at the highly effective/effective threshold but no retention effect.

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

127 Arguments that money doesn't matter in education because teachers are not paid according to performance especially emerged during the 2000s and 2010s, thanks in part to the efforts of major education reform organizations such as the Walton Foundation and Heritage Foundation.

128 Some go so far as to argue that half or more of teacher pay is allocated to "non-productive" teacher attributes, and so it follows that that entire amount of funding could be reallocated toward making schools more productive. See, for example, a recent presentation to the New York State Board of Regents from Sept. 13, 2011 (page 32), slides by Stephen Frank of Education Resource Strategies: www.p12.nysed.gov/mgtserv/docs/SchoolFinanceForHighAchievement.pdf.

129 H. Lankford, S. Loeb and J. Wyckoff, "Teacher Sorting and the Plight of Urban Schools," *Educational Evaluation and Policy Analysis* 24 (2002): 37-62.

130 S. A. Allegretto, S. P. Corcoran and L. R. Mishel, *The Teaching Penalty: Teacher Pay Losing Ground* (Washington, DC: Economic Policy Institute, 2008).

131 For a review of much the same literature, see www.shankerinstitute.org/blog/recent-evidence-teacher-experience-and-productivity.

132 M. Wiswall, "The Dynamics of Teacher Quality," *Journal of Public Economics* 100 (2013): 61-78.

133 J. P. Papay and M. A. Kraft, "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-Term Career Improvement," *Journal of Public Economics* (2015).

134 H. F. Ladd and L. C. Sorensen, *Returns to Teacher Experience: Student Achievement and Motivation in Middle School*, Working Paper 112 (National Center for Analysis of Longitudinal Data in Education Research, 2014).

135 M. D. Hendricks, *Public Schools Are Hemorrhaging Talented Teachers: Can Higher Salaries Function as a Tourniquet?*, March 24, 2015, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2564703.

136 M. D. Hendricks, "Towards an Optimal Teacher Salary Schedule: Designing Base Salary to Attract and Retain Effective Teachers," *Economics of Education Review* (2015), doi: 10.1016/j.econedurev.2015.05.008.

137 The ideal or "adequate" level of teacher compensation changes over time with changes in student needs, community demands, and the broader economy. Gilpin and Kaganovich (2011) propose a general equilibrium model of teacher quantity and quality adopting the premise that teachers' relative wages (to other sectors) are critical to maintaining a quality teaching workforce. Additionally, compression of salaries (at the high end) may reduce retention and recruitment of talented teachers. Illustrated in their data, the long-term increase in teacher quantity has led to lagging wage competitiveness, thus potentially compromising labor quality. But so too has growth in wages of competing sectors. They explain that a rise in premium for high ability will outpace that for the average. These increase costs are not offset by technological change, hence an additional downward pressure on the "real" quality of education inputs. The cost of high quality labor is increasing faster than resources required to pay the premium to simply maintain quality, and there exist no viable technological substitutes to offset those increases. As such, we can expect the real quality of educational inputs to decline. Because of the rise in premium for high ability, Gilpin and Kaganovich assert that "countering this trend would therefore require an increase in the share of GDP spent on basic education, assuming that the institutional setup of the school system remains unchanged" (428). In other words, because talent is becoming more expensive more rapidly in other sectors, more investment, as a share of GDP, may be required merely to maintain education quality. That said, this theoretical exposition, while built on much the same research base as we review herein and previously, is not fully vetted in the present article.

138 www.americanprogress.org/issues/2011/04/pdf/class_size.pdf.

139 See, for example, D. J. Brewer, C. K. Kropp, B. P. Gill and R. Reichardt, "Estimating the Cost of National Class Size Reductions Under Different Policy Alternatives," *Educational Evaluation and Policy Analysis* 21, no. 2 (1999): 171-192. While this article provides insights into the cumulative costs of adding large numbers of teachers, it makes no comparisons to other strategies that might be employed for the same dollar. The article acknowledges the research on positive effects of class size and then estimates large-scale implementation costs, seemingly implying either that achieving these positive effects is simply too expensive or that there might be more cost-effective uses of the same dollar.

140 See http://www2.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf; J. D. Finn and C. M. Achilles, "Tennessee's Class Size Study: Findings, Implications, Misconceptions," *Educational Evaluation and Policy Analysis* 21, no. 2 (Summer 2009): 97-109; J. Finn et al., "The Enduring Effects of Small Classes," *Teachers College Record* 103, no. 2 (April 2001): 145-183, www.tcrecord.org/pdf/10725.pdf; A. Krueger, "Would Smaller Class Sizes Help Close the Black-White Achievement Gap," Working Paper 451 (Princeton, NJ: Industrial Relations Section, Department of Economics, Princeton University, 2001), www.irs.princeton.edu/pubs/working_papers.html; H. M. Levin, "The Public Returns to Public Educational Investments in African American Males" (Dijon Conference, University of Bourgogne, France, May 2006), http://www.u-bourgogne.fr/colloque-iredu/posterscom/communications/LEVIN.pdf; and S. Konstantopoulos and V. Chun, "What Are the Long-Term Effects of Small Classes on the Achievement Gap? Evidence from the Lasting Benefits Study," *American Journal of Education* 116, no. 1 (November 2009): 125-154.

141 A. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114, no. 2 (1999): 497-532.

142 S. Konstantopoulos and V. Chun, "What Are the Long-Term Effects of Small Classes on the Achievement Gap? Evidence from the Lasting Benefits Study," *American Journal of Education* 116, no. 1 (November 2009): 125-154.

143 S. Dynarski, J. Hyman and D. W. Schanzenbach, "Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion," *Journal of Policy Analysis and Management* 32, no. 4 (2013): 692-717.

See also: Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2010. How does your kindergarten classroom affect your earnings? Evidence from Project STAR. Quarterly Journal of Economics 126:1593–660.

144 Another relevant study showing positive effects of pupil-to-teacher ratio reduction (different from class size) is the Wisconsin SAGE study. See A. Molnar, P. Smith, J. Zahorik, A. Palmer, A. Halbach and K. Ehrle, "Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin," *Educational Evaluation and Policy Analysis* 21, no. 2 (1999): 165-177. Unlike STAR, which was a true randomized experiment in Tennessee, SAGE in Wisconsin was designed as "a 5-year K-3 pilot project that began in the 1996-97 school year. The program required that participating schools implement 4 interventions including reducing the pupil-teacher ratio within classrooms to 15 students per teacher" (p. 165). Molnar and colleagues found: "Results of the 1996-97 and 1997-98 first grade data reveal findings consistent with the Tennessee STAR class size experiment" (p. 165). For an example of a study based on natural variation, finding no positive effects of smaller class size, see C. M. Hoxby, "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," *Quarterly Journal of Economics* 115, no. 4 (2000): 1239-1285. Hoxby uses grade-level, not student-level, data on 649 elementary schools in Connecticut, concluding "class size does not have a statistically significant effect on student achievement" (p. 1239).

145 M. M. Chingos, "The False Promise of Class-Size Reduction" (Center for American Progress, 2011), https://cdn.americanprogress.org/wp-content/uploads/issues/2011/04/pdf/class_size.pdf.

146 M. M. Chingos, "Class Size and Student Outcomes: Research and Policy Implications," *Journal of Policy Analysis and Management* 32, no. 2 (2013): 411-438.

147 Referring to estimates provided by A. B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly*

Journal of Economics 115 (1999): 497-532. See also H. Cho, P. Glewwe and M. Whitler, "Do Reductions in Class Size Raise Students' Test Scores? Evidence from Population Variation in Minnesota's Elementary Schools," *Economics of Education Review* 31 (2012): 77-95.

148 Head Start costs $8,000 per child. Given the 6 percentage-point effect noted above, the amount spent by Head Start to induce a single child into college is therefore $133,333 ($8,000/0.06). For Abecedarian, the figure is $410,000 ($90,000/0.22). The cost of reduced class size is $12,000 per student, larger than that of Head Start but considerably smaller than that of Abecedarian. The amount spent in STAR to induce a single child into college is $400,000 ($12,000/0.03). If the program could be focused on students in the poorest third of schools (the subpopulation that most closely matches that of the preschool interventions), then the cost drops to $171,000 per student induced into college. Upward Bound costs $5,620 per student. If the program could be targeted to students with low educational aspirations, the implied cost of inducing a single student into college is $93,667 ($5,620/0.06). Dynarski (2003) examines the effect of the elimination of the Social Security Student Benefit Program, which paid college scholarships to the dependents of deceased, disabled, and retired Social Security beneficiaries. Eligible students were disproportionately Black and low-income. The estimates from that paper indicate that about two-thirds of the treated students who attended college were inframarginal, while the other third was induced into the college by the $7,000 scholarship. These estimates imply that three students are paid a scholarship in order to induce one into college. The cost per student induced into college is therefore $21,000. Finally, the cost per treated subject in the FAFSA experiment (Bettinger et al., 2012) was $88, for an implied cost per student induced into college of $1,100 ($88/0.08). S. Dynarski, J. M. Hyman and D. W. Schanzenbach, *Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion* (no. w17533) (National Bureau of Economic Research, 2011).

149 Including recent work linking participation in smaller class sizes with postsecondary degree attainment: S. Dynarski, J. M. Hyman and D. W. Schazenbach, "Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion," NBER Working Paper 17533 (Cambridge, MA: National Bureau of Economic Research, 2011), www.nber.org/papers/w17533.

150 For other relatively recent studies on class size reduction, see R. Chetty, J. N. Friedman, N. Hilger, E. Saez, D. W. Schazenbach and D. Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," NBER Working Paper 16381 (Cambridge, MA: National Bureau of Economic Research, 2010), www.nber.org/papers/w16381; P. Blatchford, P. Bassett and P. Brown, "Teachers' and Pupils' Behavior in Large and Small Classes: A Systematic Observation Study of Pupils Aged 10 and 11 Years," *Journal of Educational Psychology* 97, no. 3 (August 2005): 454-467, doi: 10.1037/0022-0663.97.3.454; P. Babcock and J. Betts, "Reduced Class Size Distinctions: Effort, Ability and the Education Production Function," NBER Working Paper 14777 (Cambridge, MA: National Bureau of Economic Research, 2009), www.nber.org/papers/w14777; and S. T. Lubienski, C. Lubienski and C. Crawford-Crane, "Achievement Differences and School Type: The Role of School Climate, Teacher Certification, and Instruction," *American Journal of Education* 115 (2008): 97-138.

151 Gilraine, M., Macartney, H., & McMillan, R. (2020). Estimating the Direct and Indirect Effects of Major Education Reforms (No. w24191). National Bureau of Economic Research.

152 Gilraine, M. (2020). A method for disentangling multiple treatments from a regression discontinuity design. Journal of Labor Economics, 38(4), 1267-1311.

153 C. Jepsen and S. Rivkin, "What Is the Tradeoff between Smaller Classes and Teacher Quality?," NBER Working Paper 9205 (Cambridge, MA: National Bureau of Economic Research, 2002), www.nber.org/papers/w9205. "The results show that, all else being equal, smaller classes raise third-grade mathematics and reading achievement, particularly for lower-income students. However, the expansion of the teaching force required to staff the additional classrooms appears to have led to a deterioration in average teacher quality in schools serving a predominantly black student body. This deterioration partially or, in some cases, fully offset the benefits of smaller classes, demonstrating the importance of considering all implications of any policy change" (p. 1). For further discussion of the complexities of evaluating class-size reduction in a dynamic policy context, see D. Sims, "A Strategic Response to Class Size Reduction: Combination Classes and Student Achievement in California," *Journal of Policy Analysis and Management* 27, no. 3 (2008): 457-478; D. Sims, "Crowding Peter to Educate Paul: Lessons from a Class Size Reduction Externality," *Economics of Education Review* 28 (2009): 465-473; and M. M. Chingos, "The Impact of a Universal Class-Size Reduction Policy: Evidence from Florida's Statewide Mandate," Program on Education Policy and Governance Working Paper 10-03 (2010).

154 R. G. Ehrenberg, D. Brewer, A. Gamoran and J. D. Willms, "Class Size and Student Achievement," *Psychological Science in the Public Interest* 2, no. 1 (2001): 1-30.

155 Kraft, M. A., & Falken, G. T. (2021). A blueprint for scaling tutoring and mentoring across public schools. AERA Open, 7, 23328584211042858.

156 Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence.

157 Guryan, J., Ludwig, J., Bhatt, M. P., Cook, P. J., Davis, J. M., Dodge, K., ... & Stoddard, G. (2023). Not too late: Improving academic outcomes among adolescents. American Economic Review, 113(3), 738-765.

Some of the evidence on high-dosage tutoring programs would seem contrary to earlier findings from the STAR class-size studies—specifically, that the STAR studies and re-analyses of STAR data found class-size reduction to have strong positive effects but no effects of using paraprofessionals in classrooms. There are two key differences here. First, the use of

paraprofessionals suggested by Nickow and colleagues, based on their meta-analysis, is focused on individual tutoring, ideally where the paraprofessionals have been trained in tutoring. Second, these studies of high-dose tutoring involve both early grades and high school students, whereas STAR focused only on early grades. STAR has validated the long-run effects of smaller classes in early grades, while these studies have now validated the importance of trained professional or paraprofessional staff—as supplements to classroom programs—toward improving both early grades and high school students' academic achievement and attainment.

158 The authors note:

First, we only included studies that (a) employed quasi-random or quasi-experimental variation in school spending to estimate impacts on student outcomes, (b) demonstrated the spending variation was plausibly exogenous,1 and (c) demonstrated meaningful policy-induced variation in school spending. Next, to facilitate direct comparison, for each study we constructed an estimate of the marginal policy-induced impact on standardized outcomes of exposure to a $1000 per-pupil spending increase (in 2018 dollars) over four years. We identified 31 studies that met our conditions as of December 1, 2020.

159 Jackson, C. K., & Mackevicius, C. L. (2023). What impacts can we expect from school spending policy? Evidence from evaluations in the US. American Economic Journal: Applied Economics.

160 Rauscher, E. (2019). Delayed Benefits: Effects of California School District Bond Elections on Achievement by Socioeconomic Status (EdWorkingPaper No.19-18). Retrieved from Annenberg Institute at Brown University: http://edworkingpapers.com/ai19-18

Neilson and Zimmerman using data on New Haven, CT public schools, also found:

"Taking advantage of the staggered implementation of a comprehensive school construction project in a poor urban district, we find that, by six years after building occupancy, $10,000 of per-student investment in school construction raised reading scores for elementary and middle school students by 0.027 standard deviations. For a student receiving the average treatment intensity this corresponds to a 0.21 standard deviation increase. School construction also raised home prices and public-school enrollment in zoned neighborhoods." (p. 18)

Neilson, C. A., & Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, *120*, 18-31.

161 Biasi, B., Lafortune, J., & Schönholzer, D. (2024). What Works and for Whom? Effectiveness and Efficiency of School Capital Investments across the US (No. 16713). Institute of Labor Economics (IZA).

162 Goodman, J., Hurwitz, M., Park, R.J., & Smith, J. (2019). Heat and Learning (EdWorkingPaper No.19-30). Retrieved from Annenberg Institute at Brown University: http://edworkingpapers.com/ai19-30

We demonstrate that heat inhibits learning and that school air-conditioning may mitigate this effect. Student fixed effects models using 10 million PSAT-retakers show hotter school days in years before the test reduce scores, with extreme heat being particularly damaging. Weekend and summer temperature has little impact, suggesting heat directly disrupts learning time. New nationwide, school-level measures of air-conditioning penetration suggest patterns consistent with such infrastructure largely offsetting heat's effects. Without air-conditioning, a 1°F hotter school year reduces that year's learning by one percent. Hot school days disproportionately impact minority students, accounting for roughly five percent of the racial achievement gap.

163 Holden, K. L. (2016). Buy the book? Evidence on the effect of textbook funding on school-level achievement. American Economic Journal: Applied Economics, 100-127.

164 Uline, C., & Tschannen-Moran, M. (2008). The walls speak: The interplay of quality facilities, school climate, and student achievement. *Journal of Educational Administration*, *46*(1), 55-73, at 66.

Results confirmed a link between the quality of school facilities and student achievement in English and mathematics. As well, quality facilities were significantly positively related to three school climate variables. Finally, results confirmed the hypothesis that school climate plays a mediating role in the relationship between facility quality and student achievement.

Our results revealed that when learning is taking place in inadequate facilities, there tends not to be as clear a focus on academics, and the learning environment is less likely to be perceived as orderly and serious. Where school buildings are shabby and inadequate, there is less likely to be the kind of community engagement that supports teaching and learning. Teacher attitudes and behaviors are related as well, as teachers are less likely to show enthusiasm for their jobs and to go the extra mile with students to support their learning when they teach in buildings they judge to be of poor quality.

165 Ladd, H. F. (2011). Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement?. *Educational Evaluation and Policy Analysis*, *33*(2), 235-261.

166 Maxwell, L. E. (2016). School building condition, social climate, student attendance and academic achievement: A mediation model. *Journal of Environmental Psychology*, *46*, 206-216, at 206.

See also: Whipple, S. S., Evans, G. W., Barry, R. L., & Maxwell, L. E. (2010). An ecological perspective on cumulative school and neighborhood risk factors related to achievement. *Journal of Applied Developmental Psychology*, *31*(6), 422-427. (& https://www.schools.nyc.gov/)

167 D. Atchison, B.D. Baker, J. Levin, S. Fatima, A. Trauth, A. Srikanth, C. Herberle, N. Gannon-Slater, L. Junk, K.: Wallace (2023) Assessment of Delaware Public School Funding. https://education.delaware.gov/wp-content/uploads/2023/12/23-22933_1_Delaware_Full_Report-FMT-ed103023-Version-2.pdf

Levin, J., Baker, B., Atchison, D., Brodziak, I., Boyle, A., Hall, A., & Becker, J. (2016). Study of funding provided to public schools and public charter schools in Maryland. American Institutes for Research.

168 P. J. Wolf, A. Cheng, M. Batdorff, L. Maloney, J. F. May and S. T. Speakman, *The Productivity of Public Charter Schools* (Fayetteville, AR: Department of Education Reform, University of Arkansas, 2014), retrieved Aug. 5, 2014, from www.uaedreform.org/downloads/2014/07/the-productivity-of-public-charter-schools.pdf.

169 Some, however, find that charter school expansion may exert negative effects through student/peer sorting: "Results show that charters induce modest but statistically significant drops in math and language test scores, particularly for elementary students. However, results for middle and high school students show improvements in discipline." S. A. Imberman, "The Effect of Charter Schools on Achievement and Behavior of Public School Students," *Journal of Public Economics* 95, no. 7 (2011): 850-863. Epple, Romano and Zimmer (2015) discuss the much larger literature on this topic: D. Epple, R. Romano and R. Zimmer, *Charter Schools: A Survey of Research on Their Characteristics and Effectiveness* (no. w21256) (National Bureau of Economic Research, 2015).

170 D. Epple, R. Romano and R. Zimmer, *Charter Schools: A Survey of Research on Their Characteristics and Effectiveness* (no. w21256) (National Bureau of Economic Research, 2015).

More recent studies have affirmed the weak outcomes of Online Charter schools:

Center for Research on Education Outcomes (2015) Online Charter School Study. https://credo.stanford.edu/pdfs/Online%20Charter%20Study%20Final.pdf

Gill, B., Walsh, L., Wulsin, C. S., Matulewicz, H., Severn, V., Grau, E., ... & Kerwin, T. (2015). *Inside Online Charter Schools* (No. 9c4fae4409c44298a559a14ea948fb2c). Mathematica Policy Research.

171 P. J. Wolf, A. Cheng, M. Batdorff, L. Maloney, J. F. May and S. T. Speakman, *The Productivity of Public*

Charter Schools (Fayetteville, AR: Department of Education Reform, University of Arkansas, 2014), retrieved Aug. 5, 2014, from www.uaedreform.org/downloads/2014/07/the-productivity-of-public-charter-schools.pdf; and G. V. Glass, *Review of "The Productivity of Public Charter Schools"* (Boulder, CO: National Education Policy Center, 2014), http://nepc.colorado.edu/thinktank/review-productivity-public-charter.

172 B. D. Baker, *Review of "Charter Funding: Inequity Expands"* (Boulder, CO: National Education Policy Center, 2014), http://nepc.colorado.edu/thinktank/review-charter-funding-inequity.

173 B. D. Baker, K. Libby and K. Wiley, "Charter School Expansion and Within-District Equity: Confluence or Conflict?," *Education Finance and Policy* (2015).

Others report similar resource differentials (and more) for KIPP schools: T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Relative Efficiency of Charter Schools: A Cost Frontier Approach," *Economics of Education Review* 31, no. 2 (2012): 302-317.

174 Knight, D.S., & Toenjes, L.A. (2020). Do charter schools receive their fair share of funding? School finance equity for charter and traditional public schools. *Education Policy Analysis Archives*, *28*(51), 1-40. Retrieved August 5, 2020, from https://doi.org/10.14507/epaa.28.4438

175 Levin, J., Baker, B., Atchison, D., Brodziak, I., Boyle, A., Hall, A., & Becker, J. (2016). *Study of funding provided to public schools and public charter schools in Maryland.* American Institutes for Research. Retrieved August 10, 2021, from http://marylandpublicschools.org/stateboard/Documents/01242017/TabG-CharterPublicSchoolFundingStudy.pdf

176 Atchison, D., Levin, J., & de los Reyes, I.B. (2018). *Study of spending in public charter and traditional schools in California. Making research relevant* (p. vii). American Institutes for Research. Retrieved August 10, 2021, from https://gettingdowntofacts.com/publications/study-spending-public-charter-and-traditional-schools-california

"The conditional analyses, accounting for student needs and grade configuration, show that average traditional and charter spending within our sample were not substantially different in 2014-15 and 2015-16. In 2016-17, Aspire schools were expected to spend $1,000 or more than traditional schools in both LAUSD and OUSD when controlling for student needs and grade configuration (Exhibit B). When special education spending was excluded, Aspire and Green Dot schools in Los Angeles spent more than otherwise similar traditional schools in Los Angeles."

177 D. Epple, R. Romano and R. Zimmer, *Charter Schools: A Survey of Research on Their Characteristics and Effectiveness* (no. w21256) (National Bureau of Economic Research, 2015).

178 C. C. Tuttle, B. Gill, P. Gleason, V. Knechtel, I. Nichols-Barrer and A. Resch, *KIPP Middle Schools: Impacts on Achievement and Other Outcomes, Final Report* (Mathematica Policy Research, 2013).

Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). Inputs and impacts in charter schools: KIPP Lynn. *American Economic Review*, *100*(2), 239-243.

179 B. D. Baker, K. Libby and K. Wiley, *Spending by the Major Charter Management Organizations: Comparing Charter School and Local Public District Financial Resources in New York, Ohio, and Texas* (National Education Policy Center, 2012); and B. D. Baker, K. Libby and K. Wiley, "Charter School Expansion and Within-District Equity: Confluence or Conflict?," *Education Finance and Policy* (2015). Others report similar resource differentials (and more) for KIPP schools: T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Relative Efficiency of Charter Schools: A Cost Frontier Approach," *Economics of Education Review* 31, no. 2 (2012): 302-317.

180 W. Dobbie and R. G. Fryer Jr., *Getting Beneath the Veil of Effective Schools: Evidence from New York City* (no. w17632) (National Bureau of Economic Research, 2011). The authors note: "We find that traditionally collected input measures—class size, per pupil expenditure, the fraction of teachers with

no certification, and the fraction of teachers with an advanced degree—are not correlated with school effectiveness."

181 B. D. Baker, K. Libby and K. Wiley, *Spending by the Major Charter Management Organizations: Comparing Charter School and Local Public District Financial Resources in New York, Ohio, and Texas* (National Education Policy Center, 2012); and B. D. Baker, K. Libby and K. Wiley, "Charter School Expansion and Within-District Equity: Confluence or Conflict?," *Education Finance and Policy* (2015). Others report similar resource differentials (and more) for KIPP schools: T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Relative Efficiency of Charter Schools: A Cost Frontier Approach," *Economics of Education Review* 31, no. 2 (2012): 302-317.

182 Others also explain that many charter schools may achieve positive outcome effects by increasing the length of school days, providing additional weekend instruction, and extending school years into the summer months. The additional costs associated with all of this additional time are kept in check by employing inexperienced teachers and leveraging the lack of collective bargaining (in many charters) to require the longer hours and years, with less regard for "burnout" or turnover because turnover keeps staffing costs low. Again, scalability of these strategies is questionable. In all likelihood, systemwide expansion of school days and school years would require additional resources.

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *The Quarterly Journal of Economics*, *126*(2), 699-748.

Hoxby, C. M., Murarka, S., & Kang, J. (2009). How New York City's charter schools affect achievement. *Cambridge, MA: New York City Charter Schools Evaluation Project*, 1-85.

183 T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Relative Efficiency of Charter Schools: A Cost Frontier Approach," *Economics of Education Review* 31, no. 2 (2012): 302-317.

184 T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Relative Efficiency of Charter Schools: A Cost Frontier Approach," *Economics of Education Review* 31, no. 2 (2012): 302-317.

185 L. L. Taylor, B. L. Alford, K. G. Rollins, D. B. Brown, J. R. Stillisano and H. C. Waxman, "Evaluation of Texas Charter Schools 2009-2010" (Revised Draft) (Texas Education Research Center, Texas A&M University, College Station, 2011).

186 D. Epple, R. Romano and R. Zimmer, *Charter Schools: A Survey of Research on Their Characteristics and Effectiveness* (no. w21256) (National Bureau of Economic Research, 2015).

187 National Center for Education Statistics. Private School Survey. https://nces.ed.gov/surveys/pss/

188 Lubienski, C. A., & Lubienski, S. T. (2013). The public school advantage: Why public schools outperform private schools. University of Chicago Press.

See also: Lubienski, S. T., & Lubienski, C. (2006). School sector and academic achievement: A multi-level analysis of NAEP mathematics data. American Educational Research Journal, 43(4), 651-698.

189 For a recent review and critique of a report comparing earlier voucher studies with other reforms, see:

Baker, B.D. (2020). NEPC Review: "Comparing Ed Reforms: Assessing the Experimental Research on Nine K-12 Education Reforms." Boulder, CO: National Education Policy Center. Retrieved [date] from http://nepc.colorado.edu/thinktank/charter-research

See also: EdChoice (2020), Comparing ed reforms: Assessing the experimental research on nine K–12 education reforms, retrieved April 27, 2020 from https://www.edchoice.org/wp-content/uploads/2020/04/comparing-ed-reforms.pdf

Waddington and Berends found that in Indiana:

Overall, voucher students experienced an average achievement loss of 0.15 SDs in mathematics during their first year of attending a private school compared with matched students who remained in a public school. This loss persisted regardless of the length of time spent in a private school. In English/Language Arts, we did not observe statistically meaningful effects. (p. 783)

Waddington, R. J., & Berends, M. (2018). Impact of the Indiana Choice Scholarship Program: Achievement effects for students in upper elementary and middle school. *Journal of Policy Analysis and Management*, *37*(4), 783-808.

And Figlio and Karbownik found in Ohio:

The students who use vouchers to attend private schools have fared worse academically compared to their closely matched peers attending public schools. The study finds negative effects that are greater in math than in English language arts. Such impacts also appear to persist over time, suggesting that the results are not driven simply by the setbacks that typically accompany any change of school. (p. 2).

Figlio, D., & Karbownik, K. (2016). Evaluation of Ohio's EdChoice Scholarship Program: Selection, Competition, and Performance Effects. *Thomas B. Fordham Institute*.

Similarly, in 2017, Dynarski and Nichols summarized the state of research on school voucher programs as generally pointing to negative effects on measures of student achievement.

Four recent rigorous studies—in the District of Columbia, Louisiana, Indiana, and Ohio—used different research designs and reached the same result: on average, students that use vouchers to attend private schools do less well on tests than similar students that do not attend private schools. (p. 1)

Dynarski, M., & Nichols, A. (2017). More findings about school vouchers and test scores, and they are still negative. Evidence Speaks, 2, 18.

On average, effects on attainment are neutral to positive, BUT… most of these findings do not come from studies of statewide voucher systems, and instead come from studies of voucher programs in cities like New York, Washington, D.C. and Milwaukee, where voucher recipients in significant numbers

attended urban Catholic schools which were then compared against outcomes of students in urban district schools.

NYC study found gains for black and Hispanic students in college enrollment.

Chingos, M. M., & Peterson, P. E. (2015). Experimentally estimated impacts of school vouchers on college enrollment and degree attainment. *Journal of Public Economics*, *122*, 1-12.

DC study found increased graduation but no difference in college attendance.

Chingos, M. M. (2018). The effect of the DC school voucher program on college enrollment. *Washington DC: Urban Institute. Retrieved September*, *25*, 2019.

Wolf, P., Gutmann, B., Puma, M., Kisida, B., Rizzo, L., Eissa, N., ... & Silverberg, M. (2010). Evaluation of the DC Opportunity Scholarship Program. *US Department of Education, Institute of Education Sciences*.

Milwaukee studies show increased college attendance but no difference in completion.

Cowen, J. M., Fleming, D. J., Witte, J. F., Wolf, P. J., & Kisida, B. (2013). School vouchers and student attainment: Evidence from a state-mandated study of Milwaukee's parental choice program. *Policy Studies Journal*, *41*(1), 147-168.

Wolf, P. J., Witte, J. F., & Kisida, B. (2018). Do voucher students attain higher levels of education? Extended evidence from the Milwaukee Parental Choice Program. *Urban Institute Report*.

Joe Waddington, a University of Kentucky professor who has studied Indiana's voucher program, characterizes the state of private school voucher research as follows:

While the early research was somewhat mixed … it is striking how consistent these recent results are. We've started to see persistent negative effects of receiving a voucher on student math achievement.

Barnum, M. (2020) Do School Vouchers Work? As the debate heats up, here's what research really says. Chalkbeat https://www.chalkbeat.org/2017/7/12/21108235/do-school-vouchers-work-as-the-debate-heats-up-here-s-what-research-really-says

190 IRS findings were compiled for schools serving nearly 33% of all children attending Christian Association Schools in the 24 states studied, and 75% of children attending independent day schools in the states included in the study.

Baker, B. (2009). Private schooling in the U.S.: Expenditures, supply, and policy implications. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved [date] from http://epicpolicy.org/publication/private-schooling-US

191 Baker, B. (2009). Private schooling in the U.S.: Expenditures, supply, and policy implications. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved [date] from http://epicpolicy.org/publication/private-schooling-US

192 W. Duncombe and J. Yinger, "Financing Higher Student Performance Standards: The Case of New York State," *Economics of Education Review* 19, no. 4 (2000): 363-386; A. Reschovsky and J. Imazeki, "Achieving Educational Adequacy through School Finance Reform," *Journal of Education Finance* (2001): 373-396; J. Imazeki and A. Reschovsky, "Is No Child Left Behind an Un (or Under) Funded Federal Mandate? Evidence from Texas," *National Tax Journal* (2004): 571-588; J. Imazeki and A. Reschovsky, "Does No Child Left Behind Place a Fiscal Burden on States? Evidence from Texas," *Education Finance and Policy* 1, no. 2 (2006): 217-246; and J. Imazeki and A. Reschovsky, "Assessing the Use of Econometric Analysis in Estimating the Costs of Meeting State Education Accountability Standards: Lessons from Texas," *Peabody Journal of Education* 80, no. 3 (2005): 96-125.

193 T. A. Downes and T. F. Pogue, "Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students," *National Tax Journal* (1994): 89-110; W. Duncombe and J. Yinger, "School Finance Reform: Aid Formulas and Equity Objectives," *National Tax Journal* (1998): 239-262; W. Duncombe and J. Yinger, "Why Is It So Hard to Help Central City Schools?," *Journal of Policy Analysis and Management* 16, no. 1 (1997): 85-113; and W. Duncombe and J. Yinger, "How Much More Does a Disadvantaged Student Cost?," *Economics of Education Review* 24, no. 5 (2005): 513-532.

194 For a discussion, see B. D. Baker, "The Emerging Shape of Educational Adequacy: From Theoretical Assumptions to Empirical Evidence," *Journal of Education Finance* (2005): 259-287. See also M. Andrews, W. Duncombe and J. Yinger, "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?," *Economics of Education Review* 21, no. 3 (2002): 245-262; W. Duncombe, J. Miner and J. Ruggiero, "Potential Cost Savings from School District Consolidation: A Case Study of New York," *Economics of Education Review* 14, no. 3 (1995): 265-284; J. Imazeki and A. Reschovsky, "Financing Adequate Education in Rural Settings," *Journal of Education Finance* (2003): 137-156; and T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Impact of Facilities on the Cost of Education," *National Tax Journal* 64, no. 1 (2011): 193-218.

195 T. Downes, *What Is Adequate? Operationalizing the Concept of Adequacy for New York State* (2004), http://www.albany.edu/edfin/Downes%20EFRC%20Symp%2004%20Single.pdf.

196 Peer Reviewed Studies

Levin, J., Baker, B., Lee, J., Atchison, D., & Kelchen, R. (2022). An Examination of the Costs of Texas Community Colleges. REL 2023-142. Regional Educational Laboratory Southwest.

Baker, B. D., Weber, M., & Srikanth, A. (2021). Informing Federal School Finance Policy with Empirical Evidence. Journal of Education Finance, 47(1), 1-25.

Kolbe, T., Baker, B. D., Atchison, D., Levin, J., & Harris, P. (2021). The additional cost of operating rural schools: Evidence from Vermont. AERA Open, 7, 2332858420988868.

Zhao, B. (2022). Estimating the cost function of Connecticut public K–12 education: Implications for inequity and inadequacy in school spending. Education Economics, 1-32.

Gronberg, T. J., Jansen, D. W., & Taylor, L. L. (2017). Are charters the best alternative? A cost frontier analysis of alternative education campuses in Texas. Southern Economic Journal, 83(3), 721-743.*

*Previously, Gronberg and colleagues had been estimating single stage models, not treating the outcomes to spending relationship as endogenous. This approach also did not include indirect predictors of spending variation that might not be associated with outcomes (efficiency controls) but instead estimated a frontier model that assumed a normal-half-normally distributed portion of the error term represented inefficiency. As of this 2017 piece, the authors had modified their methods on both fronts, invoking a two-stage approach with outcomes as endogenous and including indirect controls for efficiency as in each of the other publications above, which follow a Duncombe-Yinger style model. The Kansas study below, by Taylor, adopts the two-stage frontier modeling approach with indirect efficiency controls as well.

State Cost Studies

Taylor, L., Willis, J., Berg-Jacobson, A., Jaquet, K., & Caparas, R. (2018). Estimating the costs associated with reaching student achievement expectations for Kansas public education students: A cost function approach. San Francisco, CA: WestEd. Retrieved from https://probstforprogress.com/wp-content/uploads/2018/03/kansas_adequacy_study_cost_function_approach_20180315_final.pdf*

Kolbe, T., Baker, B.D., Atchison, D., Levin, J. (2019) Pupil Weighting Factors Report. State of Vermont, House and Senate Committees on Education. https://legislature.vermont.gov/assets/Legislative-Reports/edu-legislative-report-pupil-weighting-factors-2019.pdf

Baker, B.D., Atchison, D., Levin, J., Kearns, C. (2020) New Hampshire Commission to Study School Funding, Final Report: https://carsey.unh.edu/sites/default/files/media/2020/09/20-12685_nh_final_report_version_v5_draft_1.pdf

D. Atchison, B.D. Baker, J. Levin, S. Fatima, A. Trauth, A. Srikanth, C. Herberle, N. Gannon-Slater, L. Junk, K.: Wallace (2023) Assessment of Delaware Public School Funding. https://education.delaware.gov/wp-content/uploads/2023/12/23-22933_1_Delaware_Full_Report-FMT-ed103023-Version-2.pdf

197 Baker, B.D., Atchison, D., Levin, J., Kearns, C. (2020) New Hampshire Commission to Study School Funding, Final Report: https://carsey.unh.edu/sites/default/files/media/2020/09/20-12685_nh_final_report_version_v5_draft_1.pdf

198 D. Atchison, B.D. Baker, J. Levin, S. Fatima, A. Trauth, A. Srikanth, C. Herberle, N. Gannon-Slater, L. Junk, K.: Wallace (2023) Assessment of Delaware Public School Funding. https://education.delaware.gov/wp-content/uploads/2023/12/23-22933_1_Delaware_Full_Report-FMT-ed103023-Version-2.pdf

199 Robert Costrell, Eric Hanushek and Susanna Loeb (CHL) refer to this falsification test in a non-peer-reviewed special issue of the Peabody Journal of Education, based on testimony provided in the state of Missouri and funded by the conservative Missouri-based Show-Me Institute.

See the acknowledgements at http://files.eric.ed.gov/fulltext/ED508961.pdf. Final published version: R. Costrell, E. Hanushek and S. Loeb, "What Do Cost Functions Tell Us about the Cost of an Adequate Education?," Peabody Journal of Education 83, no. 2 (2008): 198-223.

The critique asserts that if, as it would appear conceptually, the cost function is merely the flip side of the production function, then the magnitude of the spending-to-outcomes relationship should be identical between the cost and production functions. But, in Imazeki's attempt to reconcile cost and production functions using California data, the results differed dramatically. That is, if one uses a production function to identify the spending associated with certain outcome levels, and then the cost function, the results differ dramatically. CHL use this finding to assert the failure of cost functions as a method and, more generally, the uncertainty of the spending-to-outcomes relationship.

Completed and released in 2006, eventually published as J. Imazeki, "Assessing the Costs of Adequacy in California Public Schools: A Cost Function Approach," *Education* 3, no. 1 (2008): 90-108.

200 W. Duncombe and J. Yinger, "Are Education Cost Functions Ready for Prime Time? An Examination of Their Validity and Reliability," *Peabody Journal of Education* 86, no. 1 (2011): 28-57. See also W. Duncombe and J. M. Yinger, "A Comment on School District Level Production Functions Estimated Using Spending Data" (Maxwell School of Public Affairs, Syracuse University, 2007); and W. Duncombe and J. Yinger, "Making Do: State Constraints and Local Responses in California's Education Finance System," *International Tax and Public Finance* 18, no. 3 (2011): 337-368. For an alternative approach, see T. J. Gronberg, D. W. Jansen and L. L. Taylor, "The Adequacy of Educational Cost Functions: Lessons from Texas," *Peabody Journal of Education* 86, no. 1 (2011): 3-27.

In an education production function studying the effect of spending on outcomes, the dependent variable is predicted as a function of (a) a measure of relevant per-pupil spending; (b) characteristics of the student population served; and (c) contextual factors that might affect the value of the dollar toward achieving outcomes (economies of scale, regional wage variation).

Outcomes = f(Spending, Students, Context)

The cost model starts out similarly, switching the position of the spending and outcomes measures, and predicting spending levels as a function of outcomes, students and context factors.

Spending = f(Outcomes, Students, Context)

If it was this simple, then one would expect the statistical relationship between outcomes and spending to be the same

from one equation to the next. But there's an additional piece to the cost function that, in fact, adds important precision to the estimation of the input to outcome relationship. The above equation is a spending function, whereas the cost function attempts to distill "cost" from spending by addressing the share of spending that may be "inefficient." That is:

Cost = Spending – Inefficiency, or Spending = Cost + Inefficiency

That is, some of the variation in spending is variation that does not lead to variations in the outcome measure. While we don't really know exactly what the inefficiency is (which dollars are being spent in ways that don't improve outcomes), Duncombe and Yinger suggest that we do know some of the indirect predictors of the likelihood that school districts spend more than would be needed to minimally achieve current outcomes, and that one can include in the cost model characteristics of districts that explain a portion of the inefficient spending. This can be done when the spending measure is the dependent variable, as in the cost function, but not when the spending variable is an independent measure, as in the production function.

Spending = f(Outcomes, Students, Context, Inefficiency Factors)

When inefficiency factors are accounted for in the spending function, the relationship between outcomes and spending more accurately represents a relationship between outcomes and costs. This relationship would be expected to be different from the relationship between spending and outcomes (without addressing inefficiency) in a typical production function.

201 R. Costrell, E. Hanushek and S. Loeb, "What Do Cost Functions Tell Us about the Cost of an Adequate Education?," Peabody Journal of Education 83, no. 2 (2008): 198-223.

202 See, for example, B. D. Baker, "Exploring the Sensitivity of Education Costs to Racial Composition of Schools and Race-Neutral Alternative Measures: A Cost Function Application to Missouri," *Peabody Journal of Education* 86, no. 1 (2011): 58-83.

ALBERT SHANKER INSTITUTE
1998-2023