Expert Report of Clement Kirabo Jackson

Expertise

1. I am the Abraham Harris Professor of Human Development and Social Policy at Northwestern University. I also have an appointment at the Institute for Policy Research and a courtesy appointment in the Economics department at Northwestern. I received tenure in 2013 and have been a full professor since 2017.

2. I am an educational economist. I have a B.A. in Ethics, Politics and Economics from Yale University and a Ph.D. in Economics from Harvard University. The focus of my postgraduate and professional research has been in the field of the economics of education. Within this, much of my research has been on the topics of teacher quality and the impacts of school spending.

3. In recognition of my standing in the field of education, I was elected to the National Academy of Education in February 2020. I have also been asked to serve as a review panelist for projects funded by the Institute for Educational Sciences, the U.S. Department of Education, the W.T. Grant Foundation, and others.

4. I was Faculty Research Fellow at the National Bureau of Economic Research between 2009 and 2013 and have been a Research Associate there since 2017. I have been a member of the American Economic Association since 2007. I am currently the First Vice President of the Midwest Economics Association.

5. Since September 2019, I have been co-editor for the American Economic Journal: Economic Policy (one of the leading journals in applied economic policy) since September 2019. Between 2014 and September 2019, I served as co-editor of the Journal of Human Resources (a leading journal in labor economics and education policy). Between 2014 and 2017, I served on the editorial boards of Education Finance and Policy and Education Next – two leading education policy journals.

6. I have been a peer review referee for several publications in my field, including the American Economic Review, Quarterly Journal of Economics, Journal of Political Economy, Review of Economic Studies, American Economic Journal: Applied Economics, American Economic Journal: Economic Policy, Review of Economics and Statistics, The Economic Journal, Journal of Labor Economics, Journal of Public Economics, Journal of Human Resources, Industrial and Labor Relations Review, Labour Economics, Journal of Policy Analysis and

Management, Economic Inquiry, Economics of Education Review, Education Evaluation and Policy Analysis, Education Finance and Policy, Journal of Health Economics, Education Next, Berkeley Electronic Journal of Economic Analysis & Policy, Scandinavian Journal of Economics, Oxford Bulletin of Economics and Statistic, Southern Economic Journal, International Tax and Public Finance.

7. I am recognized within my profession as a leading expert on education economics, with particular expertise on the effects of educational expenditures and measuring teacher quality.

8. Between 2009 and 2020, I authored or co-authored 23 articles. Of these, 18 are published (or accepted for publication) in peer-reviewed journals. These works are listed below:

Peer-Reviewed Publications

[1] "Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending" (with Rucker Johnson) **American Economic Journal: Economic Policy**, vol. 11(4) (November 2019), pages 310-349.

[2] "Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood?" NBER Working Paper 22222. *Forthcoming* **Journal of Human Resources**

[3] "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes." **Journal of Political Economy**, (October 2018) Volume 126, Issue 5:

[4] "Can Online Off-The-Shelf Lessons Improve Student Outcomes? Evidence from A Field Experiment" (with Alexey Makarin). **American Economic Journal: Economic Policy**, 10(3) (August 2018): 226-54.

[5] "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms" (with Rucker Johnson Claudia Persico). **Quarterly Journal of Economics**, (2016) 131 (1): 157-218.

[6] "Checklists and Worker Behavior: A Field Experiment," (with Henry Schneider), **American Economic Journal: Applied Economics**, 7(4) (October 2015): 136-68.

[7] "Teacher Quality at the High-School Level: The Importance of Accounting for Tracks" **Journal of Labor Economics**. Volume 32, No. 4 (October 2014).

[8] "Do College-Prep Programs Improve Long-Term Outcomes?" **Economic Inquiry**. Vol. 52, No. 1, (January 2014), 72–99.

[9] "Can Higher-Achieving Peers Explain the Benefits to Attending Selective Schools? Evidence from Trinidad and Tobago." **Journal of Public Economics.** Volume 108, (December 2013), 63–77.

[10] "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers." **Review of Economics and Statistics**. Volume 95 (October 2013), 1096-1116.

[11] "School Competition and Teacher Quality: Evidence from Charter School Entry in North Carolina." **Journal of Public Economics**. 96(5-6): 431-438 (2012).

[12] "Single-Sex Schools, Student Achievement, and Course Selection: Evidence from Rule-Based Student Assignments in Trinidad and Tobago." **Journal of Public Economics**. 96(1-2): 173-87 (2012).

[13] "Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry." (with Henry Schneider). **American Economic Journal: Applied Economics**. 3.3 (2011): 244-267.

[14] "One for the Road: Public Transportation, Alcohol Consumption, and Intoxicated Driving." (with Emily G. Owens), **Journal of Public Economics**. 95(1-2):106-121 (February 2011).

[15] "Do Students Benefit From Attending Better Schools? Evidence From Rule-based Student Assignments in Trinidad and Tobago." **The Economic Journal**, 120(549): 1399–1429 (2010).

[16] "A Little Now for A lot Later: An Evaluation of a Texas Advanced Placement Incentive Program." **Journal of Human Resources**. 45(3):591-639 (2010).

[17] "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." (with Elias Bruegmann), **American Economic Journal: Applied Economics**. 1.4 (2009): 85–108.

[18] "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation." **Journal of Labor Economics**. 27.2 (2009): 213-256.

9. Other Publications

[19] “The Full Measure of a Teacher: Using value-added to assess effects on student behavior” ***Education Next*** (Winter 2019).

[20] "Boosting Educational Attainment & Adult Earnings: Does School Spending Matter After All?" (with Rucker C. Johnson, Claudia Persico) ***Education Next*** (Fall 2015).

[21] “Teacher' Effects and Teacher-Related Policies” **Annual Review of Economics**, Vol. 6: (September 2014) (with Jonah E. Rockoff, and Douglas O. Staiger)

[22] "Recruiting, Retaining, and Creating Quality Teachers" **Nordic Economic Policy Review**. 3(1) (2012).

[23] “Cash for Test Scores: The Impact of the Texas Advanced Placement Incentive Program,” **Education Next Fall** 8.4 (2008): 70-77.

10. I have not previously served as an expert witness.

11. A true and correct copy of my current curriculum vitae is attached hereto as Exhibit A.

Retention as an Expert in This Case

12. I have been retained as an expert witness by the law firm of Arnold & Porter Kaye Scholer LLP ("A&P"), and the ACLU Foundation of Delaware, Inc. ("ACLU-DE"), attorneys for Delawareans for Educational Opportunity ("DEO") and the NAACP Delaware State Conference of Branches ("NAACP-DE"), in In re Delaware Public Schools Litigation, No. 2018-0029-VCL.

13. I was assisted in preparing this report by a graduate student researcher, Claire Mackevicius. A statement for our accrued fees is attached as Exhibit B.

14. In this report, I offer my professional opinion regarding:

a. the effects increasing public school spending on student outcomes.

b. the effects of increasing public school spending on the outcomes of low-income students.

c. how the research applies to Delaware.

15. The opinions set forth herein are based on my expertise in the field of education economics, my review of reports and academic publications in the economics, education, and public policy literature, my review and analysis of documents and data provided to me by counsel for the plaintiffs and publicly available data and other resources from the Delaware Department of Education, and my work in this litigation, as described in this report.

Summary of Opinions on the Effect of Increasing Public School Spending on Student Outcomes for all Students Generally and for Low-Income Students

[Key Claim 1] **Based on a systematic examination of all known high-quality studies (published and unpublished) that consider whether there are credible causal relationships between school spending and outcomes, there is compelling evidence that policies that increase per-pupil spending improve student outcomes.**

- The pattern of positive associations observed in these high-quality studies is statistically incompatible with the notion that spending increases (caused by policies) do not improve student outcomes.
  - Out of 33 high-quality studies, all find a positive effect of policies to increase school spending on student outcomes. Statistically, the chance of observing this if the correlation were zero is less than one in 8 million.
- Based on a meta-analysis of 26 papers that examine test score impacts (and also provide sufficient information that a school spending effect can be computed), a policy that increases per-pupil spending by $1000 (that is sustained for four years) increases test scores on average by 11.6 percent of a standard deviation. The median effect of this same spending increase is about 4.68 percent of a standard deviation.
  - While some studies report larger test score effects for low-income populations, on average, we cannot reject that the test score effects are equally positive for both high- and low-income students.
- Based on 11 studies that provide sufficient information that a school spending effect can be computed and examine the effect on educational attainment (high school completion rates or postsecondary enrollment), policies that increase school spending increase educational attainment on average. All studies find a positive relationship. Based on a meta-analysis of these 11 papers that examine educational attainment impacts, a policy that increases per-pupil spending by $1000 (that is sustained for four years) increases educational attainment on average by about 45 percent of a standard deviation. This corresponds to being about 16 percentage points more likely to graduate high school, and 20 percentage points more likely to attend college. Because this average effect is large, we also report median effects which are less susceptible to outlier estimates. The median effect of

this spending increase is about around 14 percent of a standard deviation. This corresponds to being about 5.4 percentage points more likely to graduate high school, and 6.7 percentage points more likely to attend college.

- There is compelling evidence that increases in school spending lead to better outcomes (test scores and educational attainment) for children from low-income families. Out of 14 papers that examine the effect for low-income and non-low-income students, 12 find larger <u>educational attainment</u> benefits for children from low-income families.
- Only three studies examine the effects on adult earnings (and two are overlapping). However, all three find that increased school spending during childhood leads to significantly increased adulthood earnings. These three studies find larger benefits for children from low-income families.
- Any claim that specific policies to increase school spending do not increase educational attainment is highly inconsistent with the facts.

[<u>Key Claim 2</u>] **Many of the approaches that have been used to support the claim that increased spending does not improve student outcomes are not credible and are unlikely to be informative about policy.**

- Studies based on correlations between school spending and outcomes (where the changes in school spending are not driven by well-defined and explicit policies) are unlikely to yield causal estimates.
    - <u>NOTE</u>: If the answer to the question, "why did the spending differences examined arise?" is not clear, then the school spending differences are not driven by well-defined and explicit policies.
- No matter how high-quality any individual study, credible policy predictions should be based on an evaluation of several high-quality studies.

<u>Summary of Findings Using Delaware Data and the Resulting Conclusions on How the Research Applies to Delaware</u>

- Based on the most complete measure of state-level school spending

available, the average level of state support (on a per-pupil basis) tends to be higher in schools with higher shares of low-income students. However, after accounting for the prevalence of students with disabilities (a subpopulation that receives additional funds to accommodate special needs), the state tends to spend *less* money (on a per-pupil basis) at schools with higher shares of low-income students. That is, "disability-adjusted" per-pupil state spending is lower at high-poverty schools.

- Because many students (from both low and high-income families) attend schools with a mix of high- and low-income students, there is a small difference in state per-pupil spending between the median low- and non-low-income student. While half of the low-income students attend a school with disability-adjusted state spending levels above $6813 per-pupil, half of the non-low-income students attend a school with state spending levels above $6909 per-pupil – a $ 96 gap. At higher spending levels, there are larger differences. While 10 percent of the low-income students attend a school with disability-adjusted state spending levels above $7868 per-pupil, 10 percent of the non-low-income students attend a school with state spending levels above $8025 per-pupil – a $157 gap. At the highest spending schools, this gap is even larger. While 5 percent of the low-income students attend a school with disability-adjusted state spending levels above $8050 per- pupil, 5 percent of the non-low-income students attend a school with state spending levels above $8226 per-pupil – a $177 per-pupil gap. This shows that the most well-funded schools by the state (after accounting for students with disabilities) are those disproportionately attended by non-low-income students.
- Because some school districts may spend more state funding than others (in ways that could potentially account for the spending difference between low- and non-low-income students), we also explore patterns within districts. Looking at spending disparities between schools within districts, the majority of districts examined have lower per-pupil state spending levels in schools with higher shares of low-income students. *This is true whether one uses raw state spending levels or "disability-adjusted" state spending levels*. That is, after accounting for spending associated with disability status, most districts spend less state money (on a per-pupil basis) on low-income students.

- On average, low-income students perform worse than non-low income

students in the state. Averaging across schools in our analytic sample[1], the share of low-income students scoring below proficient on the Delaware assessment tests is 58% compared to 49% of non-low-income students on ELA. For Math, the percentage of low-income students scoring below proficient on the assessment tests is 70% compared to 58% of non-low income students.

- The share of low-income students scoring below proficient on the assessment tests is not the same across all schools. Some of these differences across schools in the performance of low- income students are structural – that is, (1) schools with higher concentrations of low-income students have worse outcomes for low-income students, and (2) schools with larger shares of students with a disability have lower performance of low-income students. However, some of the differences in performance are related to the quality of instruction provided by the school.

- After accounting for the share of students with disabilities, schools that spend more on teacher salaries (on a per-pupil basis) have lower shares of low-income students scoring below proficient on the state assessment tests. This is consistent with the notion that higher-quality instruction afforded by higher spending levels on teacher salaries is associated with better performance of low- income students on the state test. This may reflect funding gaps across schools, and also may reflect the difficulties that schools with high levels of student poverty face in attracting and retaining high-quality teachers[2].

- Many low-income students attend schools that are also attended by non-low-income students. There is a distribution of the share of low-income students across schools, ranging from 4.7% to 84.6%.

- The 95th percentile of state per-pupil direct building spending (i.e., the state-funded part of the "Direct Building Expenditures" from the ESSA sheets) is $8295.025, while the 5th percentile is $5199.432- a $3096 gap. Looking at overall school-level spending (which also includes state spending that goes to districts which is then allocated to individual schools based on Division I

[1] As detailed in the report, to focus on tradiational public school that receive standard funding from the state, we do not include vocational schools, special schools, or military schools in our naylsis. As such, these averages may differ slightly form statewide averages.

[2] See Clotfelter, Ladd, Vigdor, and Wheeler 2006; Lankford, Loeb and Wyckoff 2002; Jackson 2009, and others for research on this.

units), the 95th percentile of state per-pupil spending for individual schools (direct and district-allocated) is $9527.40, while the 5th percentile is $6784.10- a $2743 gap. To provide a sense of what would happen if spending was increased at the 5th percentile school sufficiently to bring per-pupil funding to the same level of the 95th percentile school, one can use the estimates from the causal literature (Section 1). We estimate that if one increased per-pupil school spending by $2743, the percent of low-income students scoring below proficient on the exam would fall by 2.743*(11.6*0.458)=14.6 percentage points. The average of state per-pupil spending for individual schools (direct and district-allocated) is $8114.24. Based on this, if all schools had state per-pupil spending levels brought up to that of the 95th percentile school (all else equal), one would expect the percent of low-income students scoring below proficient on the exam would fall by 1.413*(11.6*0.458)=7.5 percentage points.

- The low-income student graduation rate in Delaware is 77.88 (State Board of Education 2019). To provide a sense of what would happen to graduation rates if spending was increased at the 5th percentile school sufficiently to bring per-pupil funding to the same level of the 95th percentile school, one can use the estimates from the causal literature (Section 1). At this level, using a gap of $2743, we estimate that if one increased per-pupil school spending by $2743, the percent of low-income students graduating high school would increase by roughly 2.743*(14*0.415)= 15.9 percentage points. We also compute what one might expect if one brought the spending levels of all schools up to that of the school with the 95Th percentile of state per-pupil spending. If all schools had state per-pupil spending levels brought up to that of the 95th percentile school (all else equal), one would expect the graduation rate for low-income students to increase by 1.413*(14*0.415)=8.2 percentage points.

Discussion of Analysis for Key Claim 1

To provide some guidance on the effect of school spending, I do not rely on any single study but rather evaluate the relevant body of evidence. To this aim, I compiled a list of all studies on the effect of school spending on student outcomes. I then determined if the study fit the following inclusion criteria; that it:

(a) relied on quasi-experimental or policy variation (that is, examined the effect of a particular policy that altered school spending, or an identifiable change in school spending caused by policy)[3];

(b) demonstrate that their analysis is based on policies (or policy-induced variation) that had a large enough effect on school spending to facilitate exploring the effect of school spending on student outcomes. That is, we exclude studies of policies that do not have robust effects on school spending (as they are by definition uninformative of the effects of school spending on outcomes); and[4]

(c) demonstrated that the variation in school spending examined is unrelated to other determinants of student outcomes such as other policies or demographics (that is, provided evidence that they make comparisons across entities with the different levels of school spending but for which on average *all else was equal*).

How studies are identified:

To locate studies that meet the inclusion criteria, I performed a comprehensive search for all papers on the topic published or made public since 2000. I do not look before 2000 because, based on my knowledge of the literature and an initial search, no studies that meet the inclusion criteria existed before 2000. This is because the empricial practices in this literature were not focused on causal

[3] One well known study is excluded based on this criterion. *Husted and Kenny (2000)* does not rely on an identifiable change in school spending due to a policy. As they state "Our preferred resource equalization measure […] equals the change in resource inequality since 1972 relative to the predicted change (that is, the unexplained change in inequality). A fall in this variable reflects either the adoption of state policies that have reduced districts' ability to determine how much to spend in their district or an otherwise unmeasured drop in spending inequality." (298)

[4] Some well-known studies are excluded based on this criterion. Specifically, *Van der Klaauw (2008)* states that Title I "eligibility does not necessarily lead to a statistically significant increase in average per pupil expenditures" (750) Also, while *Hoxby (2001)* reports that "A flat grant that is 10 percent of mean income, say, is estimated to reduce the drop-out rate by 0.0076 to 0.0101 (0.76 to 1.01 percentage points)" which suggests that the effect of spending is to reduce dropouts, there is no analysis showing that this policy has a strong and robust enough effect on school spending to allow for reliable inference. Some studies examine the effects of policies that influence school spending, but do not show the effect of said policies on school spending. These include *Downes, Dye, McGuire (1998)* and *Figlio (1997).*

estimation until the early 2000s. Indeed, the earliest identified study was published in 2001, the vast majority of studies meeting this creiteria were published after 2010. From Jackson (2018), I compiled this list by Google search and citation searches. When studies were located that fit the inclusion criteria, I consulted the bibliographies for other possible papers. Also, once a list was compiled, I asked active researchers in the field to locate any additional papers for possible inclusion. Using this approach, I identified 33 studies that meet this condition (summarized in Table 1 below).[5] Note that each of these 33 studies (a) use policy variation, (b) show that their policy variation lead to robust changes in spending, and (c) establish that the variation in school spending examined is unrelated to other determinants of student outcomes to facilitate an *all else was equal* analysis.

For each study, I report the outcome(s) examined. I also code whether the relationship between school spending and the outcome examined was positive on average[6] (irrespective of statistical significance), the quasi-experimental strategy employed (see Appendix A for details), and the type of spending that was studied (i.e., operational spending, capital spending, unrestricted funds, and Title I funds). Finally, *if it was tested by the study*, we also report if the study found larger effects for students from low-income backgrounds than for higher income backgrounds.

Vote Count analysis: Does School Spending Matter?

The first straightforward way to summarize these papers (on the whole) is to highlight how many of them find positive effects and negative effects (irrespective of the magnitude or precision of the estimate). I will refer to this as the **vote count approach**. If school spending has no effect on outcomes, then each study can be considered like a fair coin flip with a 50-50 chance of being positive or negative. Out of the 33 studies, one would expect about 16 or 17 to be positive and 17 or 16 to be negative. In fact, *while some effects are small and some are not statistically significant*, all 33 report that as school spending increases (decreases) due to a particular policy (while key other factors are accounted for), student outcomes improve (deteriorate). If we treat each study as an independent datapoint, the likelihood of this many positive studies or more occurring by random chance (i.e., if there were no effect) is the same as flipping a coin 33 times and getting all heads. The likelihood of this is one in 8,589,934,592. To put it plainly, these

[5] Note that I included Jackson, Johnson and Persico (2015) and Johnson and Jackson (2019) as a single study because they use the same data and study the same setting.

[6] Most studies report overall effects. In a small number of cases, studies will report effects for different populations (e.g., 4th grade and 8th grade). In such cases, we take the average of the two effects as the overall effect.

studies provide overwhelming statistical evidence of a causal link between increased school spending and improved student outcomes.

Table 1: *List of Studies Meeting the Inclusion Criteria*

Study	pos	pos. & sig.	Outcome	Strategy	$ type	Low-Income Effect > Non-Lo
Multi-State Studies						
Abott Kogan Lavertu Peskowitz, 2019	Yes	Yes	Test Scores, Grad. Rates	Regression Discontinuity	Operational	No
Biasi, 2019	Yes	Yes	Enroll Postsecondary, Income Mobility	Instrumental Variables	Any	Yes
Brunner Hyman Ju, 2019	Yes	No	Test Scores	Instrumental Variables	Any	n/a
Candelaria Shores, 2019	Yes	No	Grad. Rates	Event-Study DiD	Any	Yes
Card Payne, 2002	Yes	Yes	Test Scores	CO-SFR	Any	Yes
Cascio Gordon Reber, 2013	Yes	No	Dropout Rates	Instrumental Variables	Title I	n/a
Jackson Johnson Persico 2015, Johnson Jackson, 2019	Yes	Yes	Grad. Rates, Years of Ed., Wages	Event-Study DiD, Instrumental Variables	Any	Yes
Jackson Wigger Xiong, 2020	Yes	Yes	Test Scores, Enroll Postsecondary	Instrumental Variables	Any	Yes
Johnson, 2015	Yes	Yes	Grad. Rates, Other Ed. Outcomes, Wages	Event-Study DiD	Title I	Yes
Lafortune Rothstein Schanzenbach, 2018	Yes	No	Test Scores	Event-Study DiD	Any	Yes
Miller, 2018	Yes	Yes	Test Scores, Grad. Rates	Instrumental Variables	Any	n/a
Non-Multi-State Studies						
Baron, 2019	Yes	Yes	Test Scores	Regression Discontinuity	Operational	n/a
Carlson Lavertu, 2018	Yes	No	Test Scores	Regression Discontinuity	Any	n/a
Cellini Ferreira Rothstein, 2010	Yes	Yes	Test Scores	Regression Discontinuity	Capital	n/a
Clark, 2003	Yes	No	Test Scores	Event-Study DiD	Any	n/a
Conlin Thompson, 2017	Yes	No	Test Scores	Instrumental Variables	Capital	n/a
Gigliotti Sorensen, 2018	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Goncalves, 2015	Yes	No	Test Scores	Event-Study DiD	Capital	Yes
Guryan, 2001	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Holden, 2016	Yes	No	Test Scores	Regression Discontinuity	Operational	n/a
Hong Zimmer, 2016	Yes	No	Test Scores	Regression Discontinuity	Capital	n/a
Hyman, 2017	Yes	Yes	Enroll Postsecondary	Instrumental Variables	Any	No
Kogan Lavertu Peskowitz, 2017	Yes	Yes	Test Scores	Regression Discontinuity	Any	n/a
Kreisman Steinberg, 2019	Yes	Yes	Test Scores, Grad. Rates, Enroll Postsecondary	Instrumental Variables	Any	Yes
Lafortune Schonholzer, 2018	Yes	Yes	Test Scores	Event-Study DiD	capital	n/a
Lee Polachek, 2018	Yes	Yes	Dropout Rates	Regression Discontinuity	Any	n/a
Martorell Stange McFarlin, 2016	Yes	No	Test Scores	Regression Discontinuity	Capital	n/a
Matsudaira Hosek Walsh, 2012	Yes	No	Test Scores	Regression Discontinuity	Title I	No
Neilson Zimmerman, 2014	Yes	No	Test Scores	Event-Study DiD	Capital	n/a
Papke, 2008	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Rauscher, 2019	Yes	No	Test Scores	Regression Discontinuity	Capital	Yes
Roy, 2011	Yes	Yes	Test Scores	Instrumental Variables	Any	Yes
Weinstein Stiefel Schwartz Chalico, 2009	Yes	No	Grad. Rates	Regression Discontinuity	Title I	n/a

Does School spending matter for various outcomes?

On the whole, there is overwhelming evidence that school spending improves student outcomes. However, different studies examine different outcomes so that it is helpful to look at studies that examine different kinds of outcomes separately. Of the 33 papers, 26 examine test scores, 11 examine educational attainment, and 2 examine wages. Vote count analyses on the papers by type are strongly supportive of positive school spending effects on all three outcomes.

Table 2: *Vote Count by Outcome Examined*

	Papers	Positive	% Positive	1 in X Chance
Test Scores	26	26	100	67,108,864
Dropout/Graduate/College	11	11	100	2,048
Wages (income mobility)	2	2	100	4

Of the 26 studies that examine the effects of policy-induced changes in school spending on test scores, 100 percent of them find positive school spending effects. If there were no effect, this high number of positive studies would occur with probability 1 in 67,108,864. This is compelling evidence of a positive test score effect on test scores – that is, compelling evidence that specific policies that increase school spending improve student test scores. For the other outcomes, the number of available credible studies is more limited. As such, the level of confidence one can have about the relationship is lower. However, despite this, for the other outcomes, all the papers find positive school spending impacts. Of the 11 papers that look at educational attainment (dropout, high school completion, or college going), all 11 find that increased school spending increases educational attainment. If there were no effect, this high number of positive studies would occur with probability 1 in 2,048 -- compelling evidence that specific policies that increase school spending improve students' educational attainment. The final outcome studied is adult earnings. There are only two independent studies that credibly link school spending to adult earnings and both find positive impacts. With only 2 studies, there is the possibility that this occurred by chance. However, even so, if there were no effect this high number of positive studies would occur with probability 1 in 4. Taken together, the pattern of results is statistically incompatible with the notion that "*money does not matter*" and provides

overwhelming evidence that policies that increase school spending improve student outcomes.

How much does school spending matter?

The vote count approach provides compelling evidence that the credible evidence on the topic indicates that school spending matters and the policies that have increased school spending improve student outcomes. However, it does not speak to the size of the school spending effects. To assess this, I perform a meta-analysis. This is a formal way to combine effects from various studies into a range of plausible estimates one might expect in the future.

This involves several steps that I outline below. To allow for comparability, I focus on papers that look at school spending effects on the same outcome. That is, papers that examine effects on test scores or test score proficiency rates would be explored together, while those that examine impacts on educational attainment (such as high-school dropout, or college-going) would be examined together. Note that some papers report effects on measures of inequality that are not easily compared to other studies (2 out of the 33 studies). While we can include them in the vote count analysis we must exclude them from the meta-analysis.

- I use the estimates from the 24 studies on test score impacts to report a range of estimates that we are confident include the true effect of school spending on test scores.
- I use the estimates from the 11 studies on educational attainment impacts to report a range of estimates that we are confident include the true effect of school spending on educational attainment.

Meta-analysis requires that all the included studies are comparable to each other. Because different studies may report effects in different ways and on different scales, or define school spending differently, each paper must be standardized in the same way to allow for comparability. In most cases, this is straightforward, but in some cases, assumptions need to be made. Given the consistent pattern across studies (i.e., all the studies find that increased school spending is associated with improved student outcomes) none of the assumptions made change the final conclusions in any appreciable way.

We seek to have one estimate per study. In most cases, there is an overall estimate provided across all populations (e.g. high-income and low-income) across all subjects (e.g. Math and English) and across all grade levels (i.e., 8^{th} grade and 4^{th}

grade). In such cases, we take the estimated effect as reported in the study. However, in some cases, overall estimates are not reported and multiple effects may be reported by subject (i.e., math and English), or grade (e.g., 4th grade and 8th grade), or population (i.e., high and low-income, or high school and elementary school). When we combine effects across subjects for the same grade, we use the simple average as our single overall effect. When we combine effects across grades, we use a precision-weighted average as our single overall effect.[7] When we combine effects across populations (i.e., high and low-income, or high school elementary school), we use the population weighted average (i.e., putting greater weight on the larger population) as our overall study effect -- this ensures that our overall estimate is representative of what the effect would be for the overall population. [8]

Standardize the effect on the outcome: Different papers may report effects on test scores with different scales, and may report impacts on different outcomes (e.g. district proficiency rates or high school completion). To facilitate comparison across studies, as is common practice, we convert each study's estimated effect into standardized units if the effect is not already reported in standardized units. That is, we divide the reported effect, ΔY, by the standard deviation of the outcome (i.e. σ_Y). The standardized effect, $\Delta Y/\sigma_Y$, is the effect on the outcome in standard deviation units.[9] As a rule of thumb, increasing an outcome by 1 standard deviation (i.e., an effect size of 1) would imply moving from the median to the 85th percentile of the outcome distribution. A modest intervention would have an effect size of between 0.05 standard deviations

[7] Precision weighting is a way to aggregate multiple estimates into a single estimate with the greatest statistical precision. Instead of a simple average, this approach puts heavier weight on estimates that are more precisely estimated (i.e., placing more weight on the estimate that is the most reliable).

[8] Each study will have an associated margin of error. This margin of error represents a range of values that are consistent with the data. A key component of this margin of error is the standard error --- which quantifies the precision of the estimates for each study. When the standard error for the overall sample is reported, we use the reported standard error. To compute the standard error of the average estimates, one must know the correlation between the effects being combined. Because this is never reported, we use the Borenstein et al. (2009) method of combining the standard errors across different effects into a single weighted estimate. When combining estimates for the same population across subjects (for example, Math and ELA scores for the same students), we assume a correlation of 0.5. When combining estimates across different grades or populations, we assume a correlation of zero.

[9] To perform this standardization, we need to gather information from each paper on the standard deviation of the outcome of interest. This standard deviation is generally reported in summary statistics. In one study, this information was not reported for the test being examined so we consulted another paper that examined impacts on that same test for this information.

(equivalent to 2 percentile points) and 0.1 standard deviations (equivalent to 4 percentile points).

Standardize the treatment: When the effect of spending on outcomes is reported in a study (as is often the case), we record the estimate of the change in outcomes associated with the change in spending directly. However, many studies report the impacts of a particular spending *policy* on outcomes, rather than the impact of spending. To allow for comparison across all studies, we define each policy based on the average change in spending that resulted from that policy. That is, we record the estimated spending change associated with the policy and use that dollar amount as the estimated change in spending. *Recall that each included study has used methods that isolate the effect of the policy on spending (and that of the policy on outcomes) from other potential confounding factors and other policies*. To ensure comparability of dollar amounts across time, we adjust reported dollars into 2018 equivalent dollars using the Consumer Price Index (CPI).

Because education is a cumulative process, exposure to a spending increase for a single year would likely have a smaller effect then exposure to the same increased spending annually over several years. Many studies measure student outcomes some years after a policy change, so the duration of exposure needs to be standardized across studies. To standardize both spending and duration of exposure to said spending, for each study, we collect the estimated average change in per-pupil spending (in 2018 dollars) over the four years preceding the observed outcome, $\Delta(\$1000_{4-Year,2018})$.[10]

Standardized Spending Effect: After standardizing both the outcome and the treatment, for each study, we compute the change in the standardized outcome per

[10] For a policy that leads to a permanent shift in spending, the four-year change in spending would be 4 times the permanent shift. However, because spending can vary across years following policy enactment, the duration of exposure and duration of the policy may not be the same. In such cases, we use the average increase in spending during the preceding four years. For example, a policy may have increased per-pupil spending by $100 in the first year, and increased linearly up to a $400 increase in the 4th year. In this case, we would use the average increase in spending during the four years, which is $250. If a study does not report spending change in the four years preceding the observed outcome, we capture the change in spending and the contemporaneous measured outcome. (Note that because school spending effects are cumulative, using the contemporaneous spending change rather than the four-year average is conservative, as it will understate the effect of spending for the previous four years).

$1000 policy-induced change in school spending averaged over four years. Formally, our standardized effect is $\mu = \Delta\left(\frac{Y}{\sigma_Y}\right)/\Delta\left(\$1000_{4-Year,2018}\right)$.[11] This standardized effect is comparable across different studies and can be used to quantify the magnitude of the school spending effect on outcomes implied not just by a single study, but by the literature as a whole.

Effects of Policy Induced Changes in School Spending on Test Scores

I now provide a summary of all studies that provide credibly causal school spending effects on test scores. Because the timing of operational spending and capital spending may be somewhat different (i.e., operational spending typically takes effect right away, while capital spending purchases assets that last for up to 50 years), one must account for this difference in timing. That is, one cannot directly compare the effect of spending an additional $1000 per student to construct a building that lasts 50 years to the effect of spending an additional $1000 per student to reduce class sizes and pay for teacher training in a single year. Adjustments need to be made to the capital spending increases to make them comparable to the changes in operational spending. As such, I first examine effects on these different spending types separately.

Effect of Increases in Non-Capital Spending on Outcomes.

I now summarize the effects of school spending (i.e., the effect on increasing school spending by $1000 per pupil for four years implied by each study) of all papers that examine the effect of a quasi-experimental change in *operational* school spending on student test scores. I focus on the 16 studies that estimate the effect of increases in operational spending (i.e., spending that was not tied to capital project such as new school construction). For each study, I present the

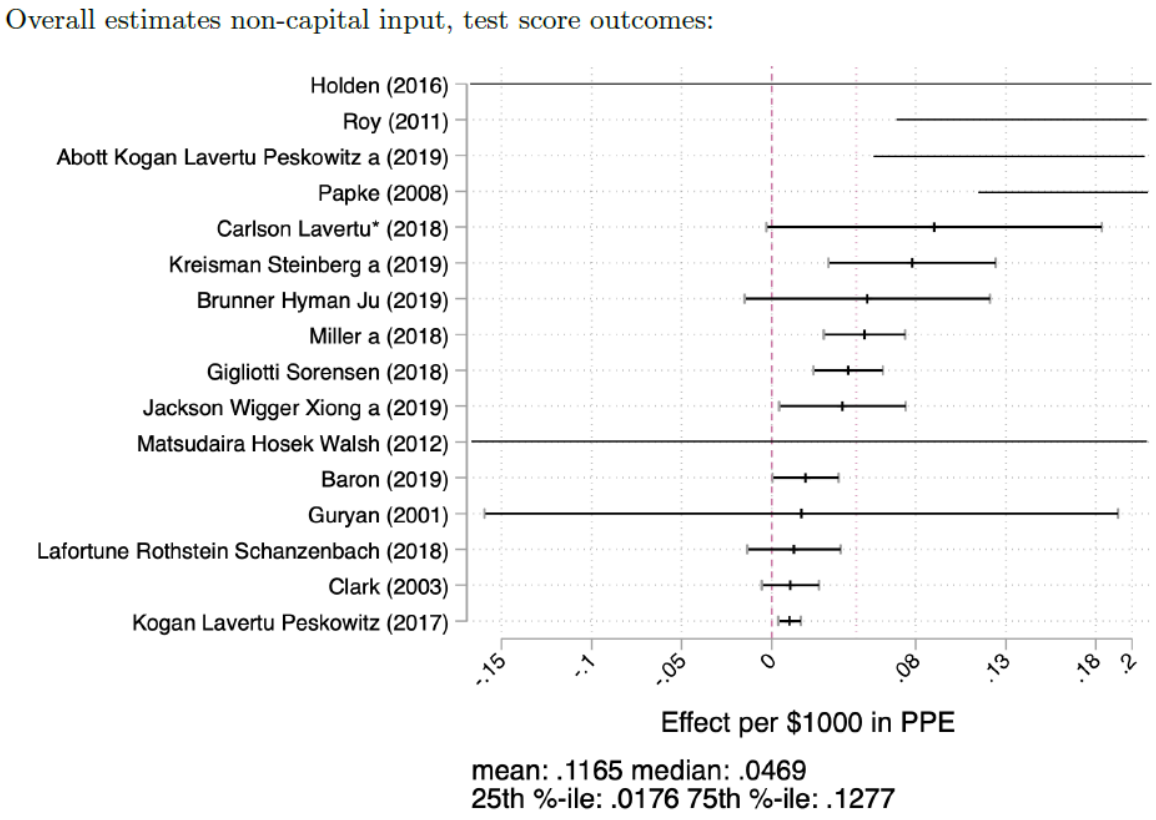
[11] When studies report effect on spending and then on outcomes, then our standardized effect, μ, is a ratio of two estimates (an estimated change in outcome divided by an estimated change in spending). When both the numerator and denominator are estimates, in the case of papers that report the effect of a *policy* and not the effect of a specific *dollar change*, one must account for this when computing the standard error of this ratio. To do this, we follow Stuart and Ord (1994) and use a Taylor expansion approximation for the variance of a ratio. If β and δ are both estimates, if $Cov(\beta, \delta) = 0$, the standard deviation of β/δ is approximately $\sqrt{(\frac{\mu_\beta^2}{\mu_\delta^2}[\frac{\sigma_\beta^2}{\mu_\beta^2} + \frac{\sigma_\delta^2}{\mu_\delta^2}])}$. This can be estimated empirically to approximate the standard error of the ratio. We take this approach.

overall estimate from the study (along with the 95 percent confidence interval for each study).[12] The 95 percent confidence interval is the interval within which there is 95% confidence that the true underlying school spending effect lies. When this confidence interval lies above zero, it means that the study finds positive school spending effects with *at least* 95% confidence.

The first takeaway is that all 16 studies have point estimates above zero. That is, all of these 16 studies indicate that policies that increase school spending improved student outcomes. Note that while 7 of the 16 studies are sufficiently imprecise that I one cannot conclude (based on that single study) that school spending has an effect, that fact all 16 studies show positive school spending effect is compelling evidence of a robust causal relationship. Of these 16 studies, 9 (56%) are statistically significant from zero at the 95% confidence level. If there were no effect of school spending on outcomes, this would be true only 5 percent of the time.

Figure 2: *Estimated School Spending Effect on Test Scores (Operational Spending)*

[12] In the case of Guryan (2001) and Matsudaira, Hosek, and Walsh (2012) the 95 CI extends beyond the range of the plot – this is a small but very imprecise estimate. In the cases of Holden (2016), Roy (2011), Abott, Korgan, Lavertu, and Peskowitz (2019), and Papke (2008), the estimate and the upper part of the CI extends past the range of the plot – these are large but imprecise estimates.



The 25th percentile of school spending effects on test scores is 0.017 and the 75th percentile is 0.127. This range of positive estimates underscores that (a) school spending effects are positive, and (b) the importance of looking at the literature as a whole to gauge magnitudes. To avoid being sensitive to outlier studies, we report both the mean and the median of these estimates. The mean effect across all the studies is 0.1165, while the median effect across all the studies is 0.0469. The average effect across all studies implies that a $1000 increase in per-pupil spending (in 2018 dollars and sustained over four years) would increase average test scores by roughly 11.65 percent of a standard deviation. To put this effect size into perspective, an increase of 11.65 percent of a standard deviation is roughly equivalent to the effect of offering universal free school breakfasts (Frisvold 2015), the difference between an average teacher and a teacher at the 84th percentile in the distribution of teacher effectiveness (Jackson, Rockoff and Staiger 2014), or the effect of reducing class size by 4 or 5 children (Schanzenbach 2014)[13]. The

[13] https://www.classsizematters.org/wp-content/uploads/2014/02/207632499-Pb-Class-Size.pdf. Note that I take the reported effect of reducing class size by 7 students of between 15 and 20 percent of a standard deviation and divide this by 17.5/11.65.

conservative median effect across all studies implies that a $1000 increase in per-pupil spending (in 2018 dollars and sustained over four years) would increase average test scores by roughly 4.69 percent of a standard deviation. Roughly half the time one would expect an effect larger than this, and the other half one would expect a smaller effect than this *but still greater than zero (as indicated above)*.

The Effect of Capital Spending on Student Test Scores

Increases in operational spending go toward educational inputs that are used in the same year (such as teacher salaries or transportation fees). In contrast, because capital spending goes toward assets that are used for several years after the initial financial outlay, it is inappropriate to relate outcomes in a given year to the financial outlays toward capital made that year. To relate capital spending to student outcomes (in a way comparable to that for operational spending) requires some additional adjustments for the life of the assets being purchased.

Before we quantify the effect of capital spending on student outcomes, we first present a simple analysis of whether policy-induced increases in capital spending are associated with improved student outcomes (irrespective of the amount). Even more than with operational spending, capital projects may take years to complete (such as school renovations), so any changes in student outcomes due to increased capital spending are likely to take time to materialize. To assess this, we first present the dynamic effects (effects over time) of the 8 studies on capital spending on student test score outcomes (see figure 3 below). Each of these studies report effects over time (relative to the year of the construction or the policy change).[14]

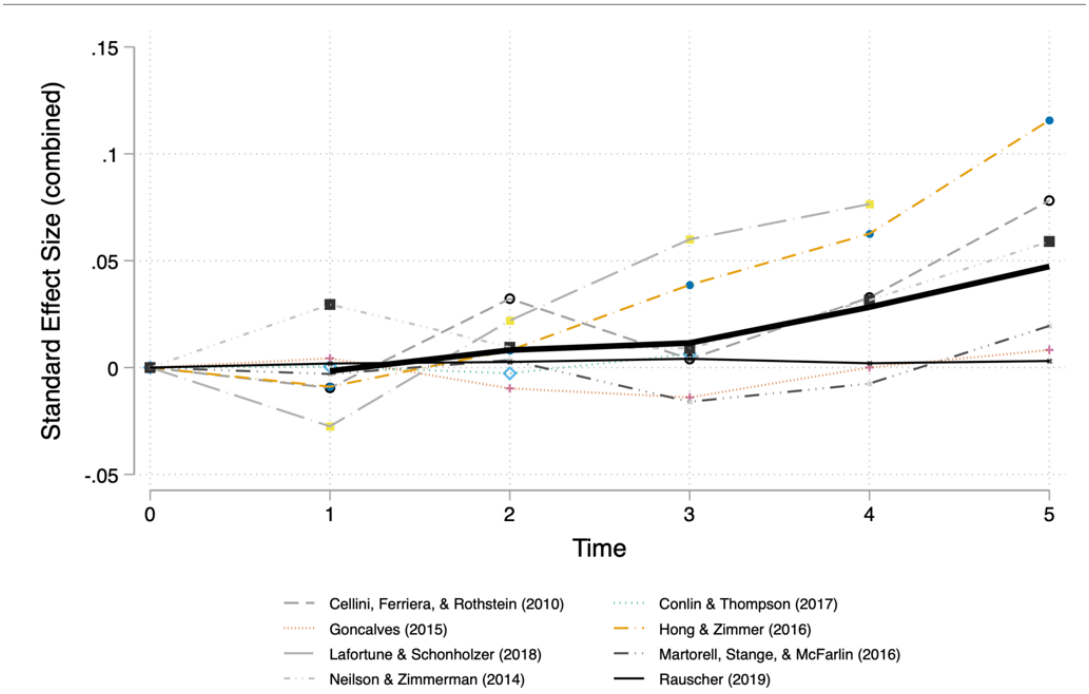
Figure 3 documents a pattern of an immediate dip in outcomes, followed by a gradual increase over time. As one can see, by about 5 years after a capital spending increase, one observes improved outcomes. Under the assumption that school capital spending has no positive effect on outcomes, the simple count of effects we observe that are positive by time t = 5 (five years later) is unlikely to occur by chance. In the first year after the money is spent, 4 out of 8 papers (0.5) are positive. This is exactly what one would expect by random chance if in fact there is no immediate effect on outcomes. By year three, however, 5 out of the 8 studies with year-three effects (62.5%) are positive. One might expect this many

[14] For Lafortune and Schonholzer (2018), Neilson and Zimmerman (2014), Goncalves (2015) year zero (t = 0) represents first year of occupancy at a new or renovated school. For all other studies, year zero (t = 0) represents year a capital bond was passed. In the case of Conlin and Thompson (2017) year zero is the year of program eligibility.

(or more) positive estimates about 14 percent of the time (i.e., 1 in 7) by random chance. By five years after the capital expense is incurred, all 6 of the studies (with 5-year effects) report positive effects on test score outcomes. If there were no effect, one would observe this 1.6 percent of the time (i.e., 1 in 64). In sum, from a statistical standpoint, while one cannot reject that there is no effect in the first year or two after a capital spending increase, one *does* reject the null hypothesis of no impact on outcomes with 95% confidence after about 5 years. The pattern indicates that (a) capital spending does improve outcomes, and (b) the benefits take about 5 years to materialize.

This pattern highlights the importance of evaluating school spending effects on outcomes a few years after the initial outlay. It also shows that if one simply correlates contemporaneous school spending on capital with contemporaneous outcomes, one may find a negative relationship. That is, the year of the financial outlay may be associated with *worse* outcomes (due to disruption effects) while the improved outcomes observed in subsequent years would be associated with lower levels of contemporaneous capital spending. This example highlights the importance of conducting a careful analysis of the timing of school spending relative to outcomes and also of accounting for the difference between when money is spent and when it may affect the classroom.

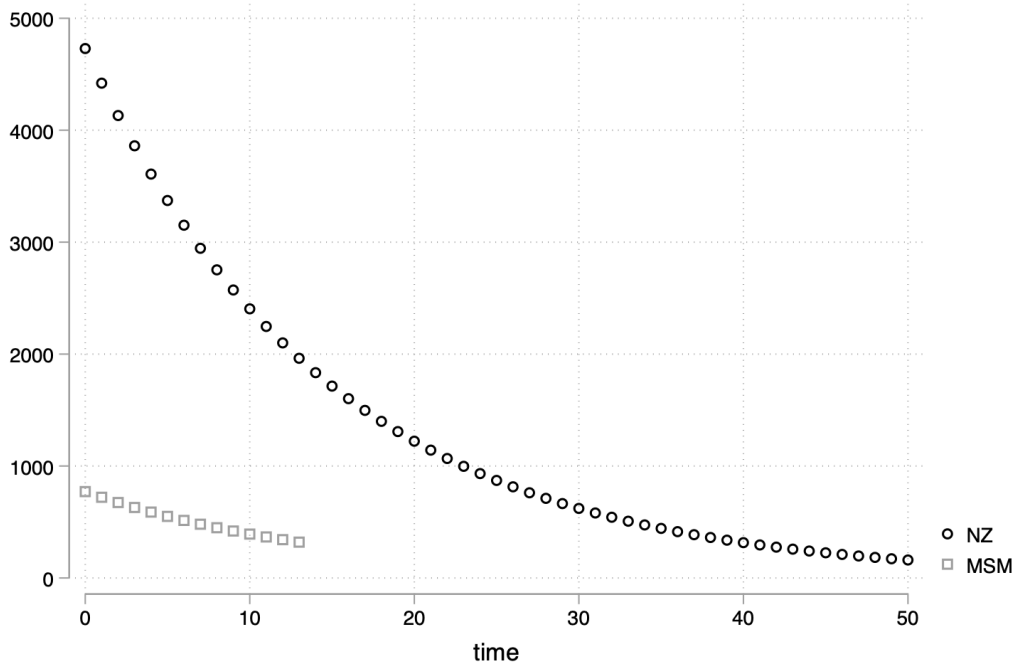
Figure 3: *Change in Student Outcomes After and Increase in Capital Spending*



How Much Does Capital Spending Matter?

Having shown compelling evidence that capital spending does improve student test scores after about 5 years, I now discuss how I adjust the capital spending to facilitate comparison with operational spending. To assess the value of $1000 in capital spending requires a few additional accounting assumptions. Specifically, it is clear that a one-time (i.e. non-permanent) $1000 increase in spending to hire an additional teacher for a single year may be reflected in outcomes in that year. In contrast, such an increase in a building should be reflected in improved outcomes for the life of the asset. In a simplistic case, where the asset does not depreciate (i.e., there is no wear and tear and the asset is equally valuable over the life of the asset), one would simply distribute the total cost of the asset over the life of an asset. For example, if the life of a building is 50 years and the building cost $1,000,000, the one-time payment of $1,000,000 would be equally distributed across the 50-year life span and be equivalent to spending 1,000,000/50=20,000 per year. However, the first year of the building is more valuable than the 50^{th} year due to wear and tear and obsolesce, so a reasonable adjustment would value the first year's value at more than $20,000 and the last year's value at less than 20,000. To make such an adjustment, I follow the convention in accounting and apply the depreciated value of a large capital spending project over the life of the asset.

Figure 4: *Annual Value of One-time Capital Expenses for two Exemplar Studies*



I assume year-to-year depreciation of 7% so that the asset losses 7% of its value each year. I depreciate the reported total expense over 15 or 50 years depending on the typical expense. The total capital expense for studies of capital spending that went primarily to new building construction or sizable renovations we depreciated over 50 years.[15] With studies that report spending on less durable assets (such as equipment or upgrading electrical wiring for technology), I depreciate the total capital spending amount over 15 years.[16] To illustrate how this relates a large one-time capital expense into a flow of education spending over the life on the assets purchased, we plot the 15-year depreciation of a $7,800 per-pupil expenditure depreciated over 15 years (as in Martorell et al. 2016) and a $70,000 per-pupil expenditure depreciated over 50 years (as in Neilson & Zimmerman 2014).[17]

[15] These studies include Neilson and Zimmerman (2014), Lafortune and Schoholzer (2018), and Goncalves (2015).
[16] For example, Martorell et al. (2016) report that most of the spending went to renovations, and Cellini et al. (2010) provide an example of specific capital projects funded by a bond referenda, including to "improve student safety conditions, upgrade electrical wiring for technology, install fire doors, replace outdated plumbing/sewer systems, repair leaky rundown roofs/bathrooms, decaying walls, drainage systems, repair, construct, acquire, equip classrooms, libraries, science labs, sites and facilities…" (220).
[17] These are the reported dollar amounts, CPI adjusted to 2018$. The reported amounts are $6773 and $54440.

Basically, a one-time capital spending outlay on new building of $70,000 per-pupil would be worth (in accounting terms) about $4,800 in year one, about $2,500 by year 10 and would fall to near zero by year 50. The fact that the value of the asset falls over time reflects the fact that the asset depreciates over time (e.g., an old building will need repairs as it ages, air conditioners will not be as efficient over time, etc.). After appropriately computing the value of the capital outlay for each year after the initial payment, we can then relate the increases in test scores observed about five years after the capital outlay to the average depreciated value of the capital outlay *during those five years*. For each paper, we compute the change in outcomes for each $1,000 in average flow value of capital spending. Because outcomes take 5 years to materialize for the capital spending papers (as shown above), we relate the year-five change in outcomes to the average spending levels (lifetime depreciated) for the previous four years.[18]

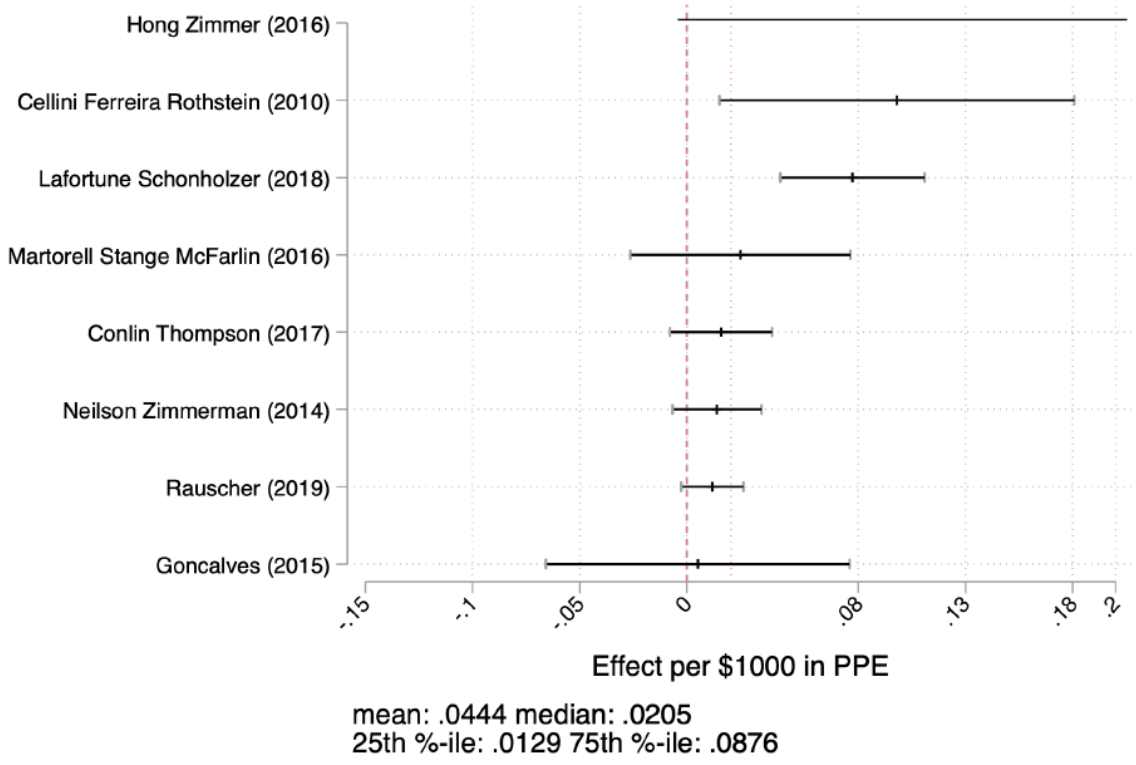
NOTE: Depreciating the asset puts more value on the early years when test scores are measured and less on the years for which outcomes are not measured (recall that most studies do not evaluate what the effect is more than 6 years after the funds are used). Because the school spending effect has the spending change in the denominator, this reduces the reported school spending effect relative to not depreciating the asset. Accordingly, my approach may be considered conservative.

In Figure 5, for each study on the effect of capital spending on test scores, I present the combined estimate of a $1000 spending increase (along with the 95 percent confidence interval for each study). The first takeaway is that all 8 studies have point estimates above zero. Estimates range from 0.0067558 to 0.2301727. However, only 2 of the 8 studies is significant at the 5 percent level and 3 at the 10 percent level. This wide range of positive estimates underscores (a) that school capital spending effects are positive, and (b) the importance of looking at the literature as a whole to gauge magnitudes. The average effect across all these 8 studies is 0.0444. That is, a $1000 increase in per-pupil spending over a four-year period increases test scores by roughly 4.4 percent of a standard deviation. The median effect is somewhat smaller at 0.0205 – indicating that a $1000 increase in per-pupil spending over a four-year period increases test scores by roughly 2.05 percent of a standard deviation. The 25th percentile of the estimated effects is 1.29 and the 75th percentile is 8.76 percent of a standard deviation. Taken together, the

[18] Note that when the five-year effect is not reported, we use the four-year effect.

estimates indicate that capital spending does improve student test scores (after about 5 years) and that the likely estimated effect would be somewhere between 1.29 and 8.76 percent of a standard deviation. These effects are somewhat smaller, indicating that operational spending may tend to have larger test score effects than capital spending. However formal statistical tests suggest that the effects are likely similar across spending types (i.e., one cannot reject that the differences by spending type are due to randomness).

Figure 5: *Estimated Effect of Capital Spending on Test Scores: 5 years after Capital Outlay*

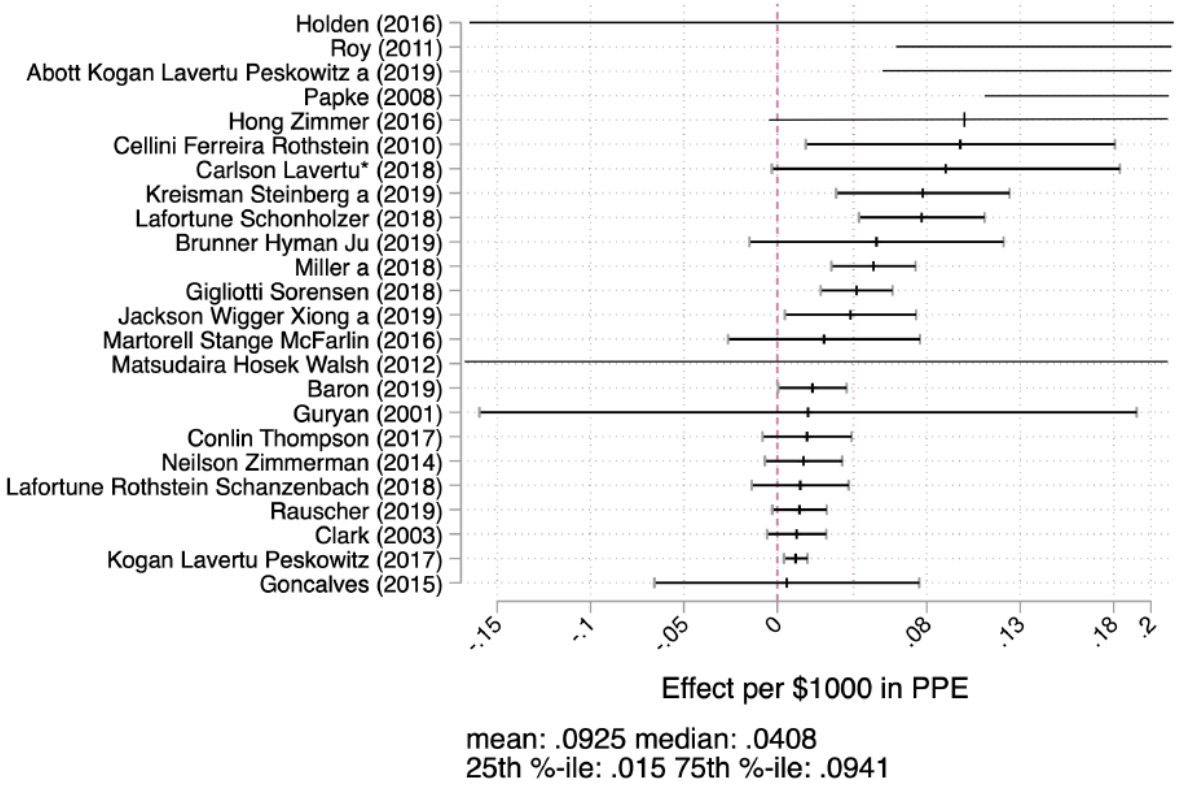


OVERALL

Having shown evidence that increases in both capital and operational spending lead to improved outcomes in the most rigorous and credible studies, I now combine both sets of studies. These studies are listed along with the estimated school spending effects in Figure 6. The first takeaway is that while not all studies are sufficiently precise to yield statistically significant effects, all 24 studies have point estimates above zero. This is, in every single credible study relating school spending to student test scores, spending increases lead to higher test scores. Estimates range from 0.00675 to 0.3. The average effect is 0.0925, indicating that, on average, increasing school spending by $1000 per pupil in 2018 dollars (and sustained over four years) would increase test scores by 9.25 percent of a standard deviation. To put this effect size into perspective, an increase of 9.25 percent of a standard deviation is roughly equivalent to the effect of having a teacher at the 82nd percentile in the distribution of teacher effectiveness instead of an average teacher (Jackson, Rockoff and Staiger 2014), or the effect of reducing class size by 4 children (Schanzenbach 2014)[19]. The 25th and 75th percentiles of the distribution suggest that one can expect a test score effects between 1.5 and 9.41 percent of a standard deviation half the time, and an effect greater than 1.5 percent of a standard deviation 75 percent of the time. Importantly, given that none of the studies yield effects of zero or below, one would likely expect positive school spending effect nearly all of the time. The conservative median effect across all studies implies that a $1000 increase in per-pupil spending (in 2018 dollars and sustained over four years) would increase average test scores by roughly 4.08 percent of a standard deviation.

[19] https://www.classsizematters.org/wp-content/uploads/2014/02/207632499-Pb-Class-Size.pdf. Note that I take the reported effect of reducing class size by 7 students of between 15 and 20 percent of a standard deviation and divide this by 17.5/11.65.

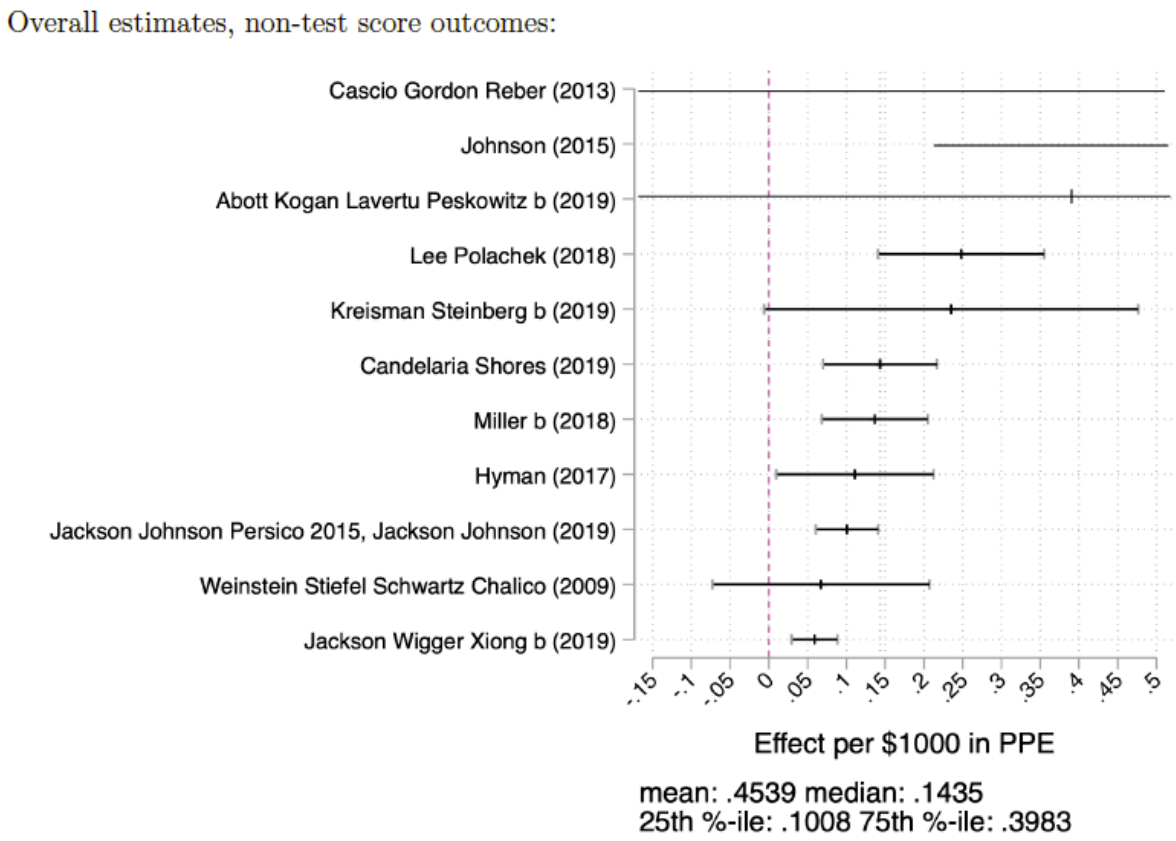
Figure 6: *Estimated School Spending Effects on Test Scores (capital and operational spending)*



Effects on Educational Attainment

While the evidence is clear that policies that lead to substantively increased school spending almost always improve student test scores, improved test scores only capture part of the overall benefit of increased school spending. Indeed, much research has shown that estimated impacts on test scores often understate meaningful impact on other key outcomes. As such, I also examine those studies that estimate effects on educational attainment.

Figure 7: *Estimated School Spending Effects on Educational Attainment*



Overall, there are 11 independent studies that relate policy-induced school spending changes to students' eventual educational attainment. Six studies examine effects on high school graduation rates, 2 studies examine effects on high school dropout rates, and 3 examine effects on postsecondary enrollment. All 11 of these studies find a positive effect of increased school spending on measures of educational attainment, and 7 out the 11 are significant with 95% confidence. If there were no school spending effects, the chances of observing all 11 positive estimates is 1 in 2048. Moreover, if there were no school spending effects, the chances of observing 7 statistically significant positive effects out of 11 is less than 1 in 170 million. Simply put, *any claim that specific policies to increase school spending do not increase educational attainment is highly inconsistent with the facts*.

How Much Does Increased School Spending Matter for Educational Attainment?

Having shown that increased school spending improved students' educational attainment, I now turn to quantifying the size of this effect. As with test scores, I standardize the effect on each educational attainment outcome by dividing by the standard deviation of that outcome. I then rescale the change in outcomes so that the reported effect represents the change in the standardized outcome from a $1000 per-pupil increase in spending sustained over the previous four years.

The first notable pattern shown is that the standardized effect sizes are larger for the educational attainment outcomes than for test scores. This reflects the fact that in many cases, the effect of increased school spending on educational attainment is much larger than what one would expect based on test score impacts alone. That is, improvements in test scores do not capture all of the benefits of increased school spending. The average educational attainment effect is 45 percent of a standard deviation. For an outcome such as high-school graduation (with a standard deviation of about 0.36) this suggests that on average, increasing school spending by $1000 (sustained for 4 years) would increase high-school graduation rates by about 0.36*0.45=16 percentage points. For college-going (with a standard deviation of about 0.45) this suggests that on average, increasing school spending by $1000 (sustained for 4 years) would increase high-school graduation rates by about 0.45*0.45=20.2 percentage points. These large average effects are potentially affected by two studies (Cascio, Gordon and Reber 2013, and Johnson 2015) that find very large effects of Title 1 spending on educational attainment after the inception of the program. As such, one may be interested in the more conservative median effects. The median across all studies is 0.143, indicating that the median educational attainment effect is 14 percent of a standard deviation. This more conservative estimate suggests that increasing school spending by $1000 (sustained for 4 years) would increase high-school graduation rates by about 0.36*0.14=5.04 percentage points, and college-going rates by about 0.45*0.14=6.3 percentage points. This is more than twice as large as the effect of being assigned to a class with 7 fewer students in kindergarten (Dynarski et al)[20], and more than six times the effect of having a teacher at the 85th percentile of test score effectiveness versus an average teacher in a single year (Chetty et. al.).[21]

[20] https://www.jstor.org/stable/42001559?seq=1#metadata_info_tab_contents
[21] https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.104.9.2633

Effects on Wages

Given the sizable effects on educational attainment, one would expect sizable effects on earnings also. While there are only two independent papers that explore this question, the evidence also suggests sizable school spending effects on earnings. Because the number of studies is small, I will discuss each study briefly. Jackson, Johnson, and Persico (2016) and Johnson and Jackson (2019) use the same data and similar approaches, so we consider them a single study. They find that for low-income children, increasing K-12 spending by 10 percent during all 12 years of a child schooling leads to about 12.5 percent higher earnings as an adult. While they do find larger effects for low-income children, when they look at non-poor children they find that increasing K-12 spending by 10 percent during all 12 years of a child's schooling leads to about 4.15 percent higher earnings as an adult. These are sizable effects that are more pronounced for low-income children. The other study is Biasi (2019). This study examines whether school finance reforms that equalized spending between those from low and high-income families resulted in a greater share of children from low-income families becoming middle and high-income as adults. She finds that this is indeed the case – largely due to improved school inputs at previously low spending schools.

Low-Income Populations (Attainment Gaps):

Having established that policies that lead to increased school spending improve test scores, educational attainment and wages, I now explore whether the effects of increased spending are more or less pronounced for children from low-income families. To examine this, I look at all the papers that test this hypothesis directly. Many papers provide estimates for low-income groups and high-income groups. If school spending effects are the same for low-income and high-income groups, then one might expect half of the papers to find larger effects for low-income students and half to find larger effects for the high-income students. Similar to the vote count tests, I examine all papers (across all outcomes) along this dimension.

As one can see in the last column of Table 1, 14 papers provide estimates for both high and low-income children. Of these 14, 12 find larger effects for low-income children. If the positive effects found in these papers were the same for low and high-income children, one would observe so many papers showing larger effects

for low-income children with probability 1 out of 1093 – that is, the statistical evidence overwhelmingly supports the hypothesis that school spending effects are larger for children for low-income families.

Table 3: *Regression Estimates of the Effect of School Spending on Standardized Educational Attainment*

Non-test score, multiple per paper:

	(1)	(2)	(3)	(4)	(5)	(6)
EstimatedEffect_lowinc		0.0809	0.0809		0.225	0.225
		(0.334)	(0.334)		(0.201)	(0.201)
capital			0			
			(.)			
_cons	0.479*	0.454	0.454	0.235**	0.143	0.143
	(0.163)	(0.215)	(0.215)	(0.0782)	(0.0979)	(0.0979)
N	16	16	16	16	16	16

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To assess the magnitude of this difference, we compare the estimated effects for low-income children to that for high-income children within the same study. We do this separately for test scores and for educational attainment. For test scores, we find little evidence that the average reported effect for low-income children differs from that of high-income children. However, for educational attainment the differences are large. Focusing on the more conservative median estimates, we find that the effect for non-low-income children is about 0.143, while that for children from low-income families is about 0.368. That is, for non-low income children, the median educational attainment effect is 14.3 percent of a standard deviation. This suggests that increasing school spending by $1000 (sustained for 4 years) would increase high-school graduation rates by about 0.36*0.143=5.14 percentage points, and college-going rates by about 0.45*0.143=6.3 percentage points. In contrast, the median educational attainment effect is 36.8 percent of a standard deviation. This suggests that increasing school spending by $1000 (sustained for 4 years) would

increase high-school graduation rates by about 0.36*0.368=13.2 percentage points, and college-going rates by about 0.45*0.368=16.6 percentage points. These are sizable effects. It is important to note that these reported effects are estimated with error so that the difference between high- and low-income children is not precisely estimated. As such, a conservative reading of the evidence is that school spending effects are larger for low-income children, but that the extent to which this is true can only be quantified with limited precision. Importantly, even the more precisely estimated overall effects are economically meaningful.

Different Kinds of Estimation Strategies

As I classified studies, I recorded the quasi-experimental estimation strategy employed in each study to isolate the effect of a specific policy on school spending and on outcomes. I briefly outline each estimation strategy below.

Difference in Difference Event Study:

An **event study** is a statistical method to assess the impact of a particular event on some outcome (e.g., school spending or test scores). For example, the passage of a state-wide school finance reform would be a particular event. To understand the effect of the event on outcomes, researchers may use a **difference-in-difference** approach. This involves estimating the change in outcomes caused by the event. To do this, researchers first compute the change in outcomes in areas affected by the event (say those low-income districts in school finance reform state) – this is the first difference. Because outcomes may change over time for reasons other than the event, researchers would also compute the change in outcomes for unaffected areas over the same time period (i.e., the change in outcomes for similarly low-income districts on non-reform states over the same time period) – the second difference. Under the assumption that the change in outcomes in areas unaffected by the event are similar to the change that would have been observed in the affected areas had the event not occurred, one can compare the difference over time in the affected areas to the difference in the unaffected areas to isolate the effect of the event on outcomes. By comparing outcome before versus after an event in the same location (such as a district or state), one avoids making comparisons across entities that may differ in important unobserved ways. By comparing the change in outcomes for affected areas to that of unaffected areas, one accounts for any underlying changes over time that may affect all entities. Studies of this type spend much time establishing that (a) the areas affected by the event were on similar trajectories before the event so that they would likely have had similar trajectories had the

event not occurred, and (b) there were no other coincident policies or changes that occurred at the time of the event.

Regression Discontinuity:

A **regression discontinuity design (RDD)** is a quasi-experimental design that elicits the causal effects of interventions by assigning a cutoff or threshold above or below which an intervention is assigned. For example, schools below a pre-specified poverty level (i.e., the cutoff) might receive additional state funds. By comparing observations lying closely on either side of the threshold, it is possible to estimate the effect of the treatment (in this case additional funds) on outcomes. For example, many local areas have referenda to possibly increase school spending. These motions to increase school spending pass if the vote share exceed 50 percent and fail otherwise. Because areas that vote overwhelmingly in support of increased spending may differ from those that do not, the RDD design compares the outcomes of areas that just passed (i.e. with 50 or 51 percent of the vote) to that of areas that just failed to pass (i.e. with 49 or 48 percent of the vote). Because the actual vote share is very similar in areas right around the cutoff (50 percent), *within a narrow band of the cutoff*, areas that pass referenda will tend to be very similar on average to areas that do not. If one lines up outcomes for each local area against the vote share, any sudden change **or discontinuity** in outcomes that occurs right at 50 percent can be attributed to the increased spending.

Instrumental Variables:

In many settings, one wishes to estimate the effect of some variable X (say school spending) on another variable Y (S.A.T. test scores). However, in many settings one worries that differences in X may be correlated with unobserved determines of the outcome Y. For example, if one were to simply correlate school spending with test scores, one would worry that because higher spending areas tend to have higher income students, the simple relationship between school spending and test scores may be confounded with a parental SES effect. Similarly, given that schools that enroll high-shares of special education students will receive additional funds, the simple relationship between school spending and test scores may be confounded with a special education effect.

An instrumental variable (sometimes called an "instrument" variable) is a third variable, Z, that influences X but is not correlated with unobserved determinants of the outcome. A good example comes from policies that create mechanical changes in school spending that are unrelated to underlying changes in demographics of other policy changes. For example, Michigan's Proposal A (passed in 1994) sought to equalize spending over time between high and low spending districts. It did so by setting a schedule for the state foundation allowance that varied over time and differed across districts. Because per-pupil spending reflects the revenues from state, local and federal government, the changes in this foundation allowance over time would be strong predictors of changes in school spending over time. That is, areas for which this allowance was growing more rapidly would tend to have more rapid growth in per-pupil spending. Because the growth in the allowance over time within a district was determined by a formula (and not related to other potential confounders), if areas with more rapid growth in this allowance (and also school spending) also experienced more rapid test score improvements over time, it would be evidence of a causal relationship between school spending and test scores. In an instrumental variable framework, this instrumental variable Z would be the foundation allowance. The instrumental variables estimate would divide the change in the outcome Y due to Z, by the change in the treatment X due to Z, to uncover the relationship between X and Z (as generated by Z). The key underlying assumption behind instrumental variables models is that the variable Z is unrelated to other unobserved determinants of Y. Studies in this category spend much time presenting empirical evidence supporting this assumption.

Discussion of Analysis for Key Conclusion 2:

Appropriate forms of evidence: From a policy perspective, what researchers wish to understand is the extent to which some policy or rule that increases per-pupil spending will result in a change in student outcomes. I will refer to this as the "policy effect." Observed correlations that are not driven by policy changes or policy variation, *no matter how sophisticated*, are unlikely to inform our understanding of the effect of increasing school spending on outcomes. The best and most credible way to assess the effect of enacting policy that increases school spending on outcomes is to examine whether previous policies that increased (or decreased) the level of per-pupil spending resulted in a change in student outcomes. In sum, to best inform the likely effect of future policies, one should examine the effect of past policies. Moreover, when evaluating the effect of

policies, one must always be careful to use a well-defined **counterfactual** (i.e., guess of what outcomes would have looked like without the policy). By using carefully chosen comparison groups, one is able to make like-with-like (or all else equal) comparisons. In addition to having a well-defined comparison group, credible studies must establish that other than the policy change being examined, the comparison group and the affected group are similar.

What kind of evidence is valid?

It is important to state that no evidence is perfect -- this is true in the social sciences and medical sciences and is a fact of life. However, (a) some forms of evidence are better than others, (b) the most credible studies attempt to isolate the effect of school spending on outcome and present evidence that there are no other confounding factors, and (c) evidence on how school spending policies affect outcomes speaks much more directly to the relevant question of whether a policy to increase school spending in Delaware would impact student outcomes.

There are several kinds of correlational evidence that do not examine the effect of policies or use policy-induced variation in spending. I put the studies into categories and describe them below. These correlational studies do not generally have a well-defined counterfactual, and if they do, the counterfactual is typically implausible. Moreover, many of these studies make little attempt to establish that the changes in school spending examined are unrelated to other confounding policies or demographic shifts. Fundamentally, most of the problems with these correlational studies stem from the fact that the variation in school spending (i.e., the differences in school spending being examined) is not exogenous (i.e., determined from outside the system). The differences in spending that these correlational studies use are the results of family decisions, demographic changes, and other policies that each independently may affect outcomes. When this is the case, the comparisons are rarely made *ceteris paribus*, or all else equal.

I briefly outline four kinds of correlational studies that are generally presented. For each kind I outline the underlying logic, highlight the flaws, and explain why studies of this type are not informative about the effects of a policy to increase school spending on student outcomes.

Time Series Correlations: Time series correlation evidence is evidence based on comparing spending levels within an entity (a nation, state, district, or school) over time to the change in student outcomes for that same entity over time. I discuss two time-series examples that, if naively interpreted, would lead one to different conclusions.

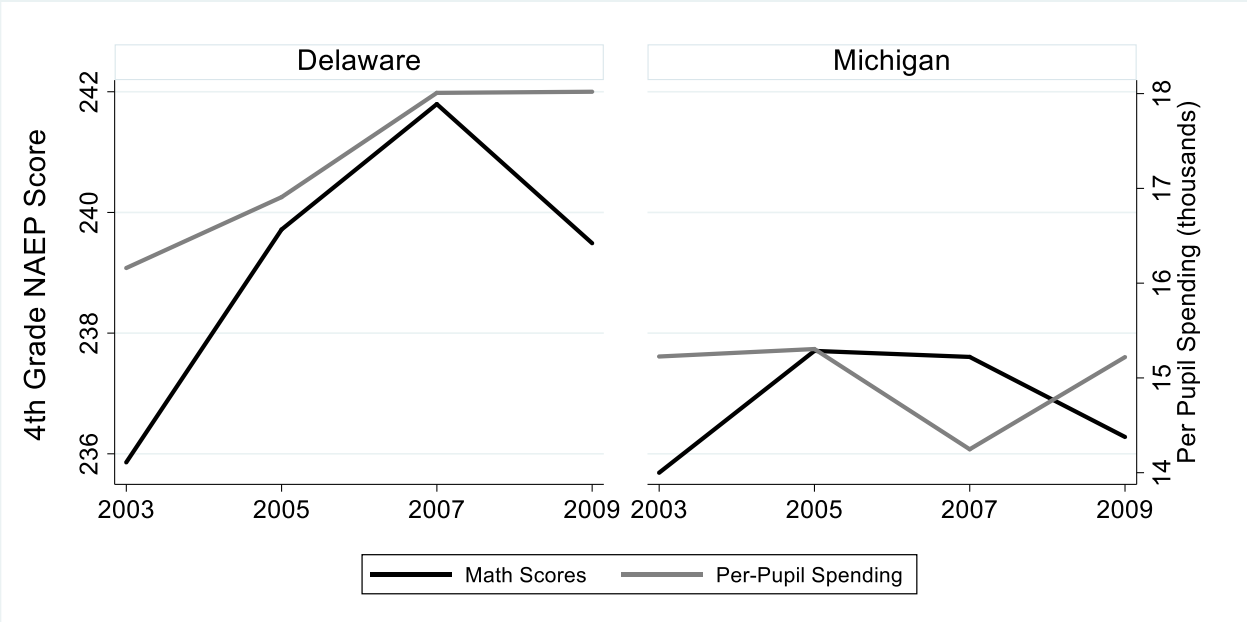
(a) A commonly used example of time-series evidence used to support the claim that school spending does not improve student outcomes is that *per-pupil U.S. education spending has more than doubled since 1970, while student achievement has not dramatically improved*.[22] In the case of this time series comparison, the model compares the United States in 1970 to the United States in 2010. To interpret the correlation between school spending and student achievement over time as reflecting a causal relationship (i.e., a relationship that reflects what would have happened if we had a policy to suddenly double per-pupil spending) requires that other than higher spending levels, all else was equal in the United States between 1970 and 2010. This is on its face implausible.

(b) To provide another example, consider changes on the National Assessment of Educational Progress (NAEP).[23] The NAEP is referred to as the Nation's Report Card as it tests students across the country on the same assessments and has remained relatively stable over time. Per-pupil spending in Delaware increased from $16,160 in 2003 to $18,020 in 2009 (all reported in 2015 dollars). During that same time period, 4th grade math scores on the NAEP increased by 3.63 points (see Figure 8 below). That is, during this time period, every $1,000 increase in per-pupil spending was *associated* with a 3.63/(18.020-16.160)=1.95 point increase in statewide average NAEP test scores. However, this does not mean that a policy that increased spending by $1,000 per pupil would increase student test scores by that amount. For this relationship to be causal (i.e., indicative of what a policy would do) requires that nothing else changed in Delaware between 2003 and 2009 that could have influenced test scores. We know that unemployment rates in Delaware fell from 4.3% in 2003 to 3.4% in 2007 and then jumped to 8.3% in 2009. Given that economic conditions could have an influence on both per-pupil spending levels and test scores, while these patterns are suggestive, it would be unwise to give these time series associations a causal interpretation.

[22] *See, e.g.*, https://www.heritage.org/education/report/does-spending-more-education-improve-academic-achievement. As a point of fact, test scores have improved for many age groups and demographic groups since 1970.
[23] Data is available at https://www.nationsreportcard.gov/ndecore/landing.

In sum, time-series correlations can be very misleading and should not be used to make inferences about the impacts of possible policies to increase school spending.

Figure 8: *4th Grade NAEP Scores and Per-Pupil Spending in Delaware and Michigan Over Time*



Cross Sectional Correlations: In the case of the cross-sectional correlations, one compares outcomes in one geographical location (say a state) to those of other locations. For example, Delaware spent more per-pupil than Michigan in 2007 ($18,009 vs. $14,247) and 4th-grade math scores on the NEAP were higher in Delaware than in Michigan in 2007 (241.8 vs. 237.6 for 4th-grade math).[24] Indeed, each additional thousand dollars spend per-pupil in Delaware compared to Michigan, was *associated* with test scores that were (241.8-237.6)/(18.009-14.247)=1.116 points higher. While the correlation between the spending levels and test scores across these two states is positive -- that is, Delaware spent more per-pupil and had higher test scores -- it would be misguided to infer that this means that a policy to increase per-pupil spending in Michigan to the levels in Delaware would lead test scores to be equal in the two states. To interpret the correlation between differences in school spending and student achievement across states as reflecting a causal relationship requires that other than higher per-pupil

[24] *Id.*

spending levels, all states are essentially the same. *To be clear, states need not be identical, but it needs to be the case that there are no systematic differences between high- and low-spending states that also could have an independent influence on student performance*. Given that high-spending states and low-spending states tend to differ along a variety of important dimensions (such as average income levels, geography, demographic composition, education policies), it is clear that high- and low-spending states do differ in important systematic ways so that cross-sectional comparisons cannot provide estimates of the effect of policies that increase school spending in isolation of anything else happening in the state. Put differently, given that high spending states tend to be located in the northeast and have lower levels of poverty, the achievement levels of low-spending states is not a plausible counterfactual for the achievement levels of high-spending states had their spending levels been lower.

To provide another example from within Delaware, consider Kent Elementary Intensive Learning Center in Caesar Rodney district. This school has the highest level of personnel spending from the state of any other school in the district ($22,659.27, compared to $10,844.24 in 2018). At the same time, it has some of the highest rates of students scoring [not proficient] on the state assessment. While a naïve analysis might conclude that more school spending *reduces* student performance, closer inspection reveals that Kent Elementary Intensive Learning Center is a school for students with special needs so that the student population at this school differs considerably from that at other schools in the district. As such, correlating spending levels and student performance across Kent Elementary Intensive Learning Center and other schools is not a *ceteris paribus* comparison, and is uninformative about what the effect would happen to student achievement in any given school if school spending were increased. NOTE: I use this example to illustrate the conceptual issue. In the analysis of schools, I do not include special schools such as Kent Elementary Intensive Learning Center to avoid this kind of confounding.

In sum, cross-sectional correlations can be very misleading and should not used to make inferences about the impacts of possible policies to increase school spending.

Difference in Difference Evidence: In a Difference-in-Difference (DiD) study, one compares the change in outcomes that occurs in one entity (state, district, or school) that experiences larger spending increases over time to the change in

outcomes that occur in other entities that experience smaller spending increases over that same time. I provide an example below.

As stated above, per-pupil spending in Delaware increased from $16,160 in 2003 to $18,020 in 2009 (all reported in 2015 dollars). By contrast, between 2003 and 2009 per-pupil spending in Michigan was virtually unchanged -- it went from $15,226 in 2003 to $15,220 in 2009. Since school spending was virtually unchanged in Michigan between 2003 and 2009, it is reasonable to expect that the trajectory of test scores in Michigan is informative of what they would have been in Delaware had school spending not increased. It turns out that 4th grade math scores on the NAEP in Michigan increased by 0.6 points (from 235.7 to 236.3). Informed by this pattern, a difference-in difference estimate would assume that had there been no change in spending in Delaware, Delaware test scores would have *also* increased by 0.6 points. In fact, during that same time period, 4th grade math NAEP scores in Delaware increased by 3.6 points. Relative to the "comparison" state of Michigan, Delaware increased by 3.6 – (0.6) =3 points. Also relative to the "comparison" state of Michigan, Delaware spending increased by ($18,020 – $16,160) – ($15,220-$15,226)=$1,866. Using the difference in difference approach, every $1,000 *relative* increase in per-pupil spending was associated with a 3/(1.866)=1.608 point *relative* increase in 4th grade math test scores.

The power of the difference in difference approach is that unlike the cross-sectional correlations or the time series correlations, it <u>does</u> have a well-defined counterfactual. From this perspective, difference-in-difference evidence is superior to the previous two types. However, whether any difference in difference approach is credible hinges on whether the counterfactual is plausible. To interpret the correlations between spending growth and changes in outcomes in one state, while using the change in outcomes for low spending growth states (or districts) as basis for comparison, requires that states with high and low spending growth also experienced a similar change in other dimensions also. Given that school spending levels may be based on changes in the population of English language learner students or special education students, and also on the health of local economies (each of which may independently affect outcomes), this "all else equal" assumption is implausible in this setting. Indeed, while the unemployment rate increased by 4.9 percentage points in Delaware, it increased by 6.7 percentage points in Michigan so that the change in Michigan is unlikely to be a plausible counterfactual. Similarly, the evolution of the prevalence of special education

students or English language learners may also not have been similar across these two states.

Another related limitation of this kind of difference-in-difference analysis is that it compares overall changes in per-pupil spending in one state to that of others, but does not focus on those changes in spending that are unrelated to other potentially confounding policies and demographic changes. Some of the changes in spending experienced in a given state reflects changes in school finance policies, while some of the changes reflect changes that occur due to changes in housing prices, changes in the economy, and even changes in student demographics (e.g., more special education students). Because we know that changes in student demographics, changes in the economy, and changes in housing prices may independently influence student outcomes, we know that the overall changes in school spending are related to factors that may independently influence outcomes – precluding a like-with-like comparison. NOTE: Most credible studies will seek to isolate the changes in school spending that are due to particular policy changes and relate those to changes in outcomes in order to address this concern. To do this requires the examination of specific policies (which can also be done within a difference-in-difference framework).

Controlling for Other Factors: In all observational studies there is the possibility that there is some unobserved factor that confounds the estimated effect. The most common of these are family background or economic conditions. A common approach to dealing with this is to obtain data on economic conditions and family background and to account for these directly in the model. While researchers often claim that this fixes the problem, it does not fix all problems, and this approach may actually introduce others.

For example, in models that control for family background, one compares the outcomes of families of similar income levels that have decided to expose their children to different levels of per-pupil spending. Note that this variation in spending may have nothing to do with policy. Given that families do not typically choose schools at random, this implies that holding family background fixed, families that choose lower resourced schools will either be those that place a lower value on education or those families that have access to extra resources to supplement their child's education – in either case, controlling for family income does not facilitate an *all else equal* comparison and certainly does not isolate

differences in spending attributable to specific policies. While it may seem counterintuitive, because families choose schools and other educational inputs as a bundle, controlling for family background itself can create the negative correlation between per-pupil spending and spending on additional educational supplements. That is, the typical high-income family will have a more expensive "education bundle" and is likely to attend a public school with higher per-pupil spending and spend more on educational supplements than a low-income family. However, among high-income families (i.e., controlling for family background) that spend the same amount on the "education bundle," the families that attend higher resourced schools will likely spend less on additional supplements.[25] That is, without controlling for family income, the differences in per-pupil spending are likely positively correlated with additional supports, but controlling for family income, the differences in per-pupil spending are likely negatively correlated with additional supports. This example makes clear how controlling for a possible confounder in correlational studies can fix one problem only to cause another. In essence, because families choose schools and other educational inputs as a bundle, when one compares families with similar income levels who attend schools with higher or lower per-pupil spending, one is likely not comparing families that are truly similar – precluding an all else equal interpretation. Moreover, even if one were able to account for family background differences, the remaining variation in spending may have nothing to do with policy.

Similarly, in models that control for economic conditions, one compares locations that face similar economic conditions, but have different levels of per-pupil spending. Given that school spending is determined, in part, by economic conditions, there are likely key differences across areas with similar economic conditions but different spending levels. For example, holding economic conditions fixed, areas with higher property values (and therefore higher wealth) will spend more on education. Also, because state budgets are finite, holding economic conditions fixed, areas that spend more on K12 schools may spend less on other things that also affect outcomes (such as pre-K or healthcare). Unless one can control for all these factors, correlational studies are unlikely to identify the all-else-equal effect of increased school spending. Moreover, even if one were able to account for economic conditions, the remaining variation in spending may have nothing to do with policy so that the resulting relationships may not represent the

[25] Note that this is the same behavior that leads some families to use extra resources instead of sending their children to private schools.

changes in outcomes that one would observe based on policies that increase school spending.

The above discussion makes clear that observational studies that do not focus on particular policies are (a) unlikely to reflect causal relationships and (b) unlikely to be informative about the effects of school spending policy on student outcomes. It also highlights the fact that because overall school spending levels are determined by family decisions, policy changes, economic conditions and demographic changes (many of which are interrelated), adding controls for particular confounders may not eliminate the bias due to those confounders and can introduce other biases. The best way to be confident that the variation in school spending is caused by school spending policy is to investigate school spending policies directly. Also, the best way to be confident that the variation in school spending is unrelated to other confounding factors is to analyze particular policies that change school spending levels for reasons unrelated to other policies, family decisions, or demographic changes. In short, instead of using variation in school spending that is confounded and applying imperfect fixes to it, a much better approach is to use variation in school spending that is not confounded in the first place.

A Better Approach

Because it is generally impossible to account for all possible confounding factors in an observational study, most credible research relies on natural experiments, or policy changes that lead to changes in per-pupil spending that *are not* the direct result of demographic changes (such as the number of students with disabilities or low-income students), individual family decisions, or other confounding policies. In the most credible studies, researchers examine the effect of specific changes in per-pupil spending (driven by a specific policy) that are plausibly unrelated to other confounding factors. By examining the effect of specific policies on outcomes, these studies speak directly to the question of how a policy to increase school spending may affect student outcomes. Also, because these studies examine specific policies, the reasons why one state or district has higher spending levels (or faster per-pupil spending growth) is known. As such, the set of possible confounders is also known and understood. Importantly, not only do these studies rely on exogenous variation, they also spend much time presenting evidence that the variation used is unrelated to other factors. Only those studies that both (a) rely on policy-driven differences in school spending that are plausibly unrelated to

other determinants of outcomes, and (b) demonstrate empirically that the variation in school spending used is indeed unrelated to other determinants of outcomes (i.e., that the comparisons are truly all else equal), provide credible evidence on the policy-relevant causal effect of increasing (or decreasing) school spending on outcomes.

The New Literature on School Spending

In the past ten years, a new more credible literature relating school spending to student outcomes has emerged. The more recent studies rely on empirical models that employ exogenous variation in school spending in order to disentangle the influence of school spending from that of family background and other influences. What distinguishes the new research from the old research is that is it design based. That is, the new studies lay out a clear comparison group to which some treated group will be compared. These studies do not simply use any variation in school spending, but rather rely *only* on changes in school spending that are based on specific policies and are therefore known and understood. These studies examine changes in school spending that are generated by policy (this includes, the passage of legislation regarding school spending levels, the passage of school spending bonds via close elections, rules in school spending formulas, and others). By examining variation in school spending due to particular policies, researchers need not speculate, but know exactly why one family is exposed to more school spending than another. This specificity is critical to assessing the extent to which the results in each study can be interpreted causally. In each of these newer studies, because they rely on policy variation, the source of variation in school spending is transparent, and that source can credibly be argued to be unrelated to family background and other attributes. Although none of these studies is perfect, each study is clear about the possible sources of any bias, and each study conducts considerable sensitivity analysis on the main results. The meta-analysis of studies relies only on studies of this type.

An Exemplar Study[26]

To illustrate how these policy-based studies seek to isolate the casual impact on increasing school spending, I detail one such study of my own. It is helpful to start

[26] This text is taken from Chapter 7 of *Confronting Inequality: How Policies and Practices Shape Children's Opportunities*. *See* https://www.apa.org/pubs/books/confronting-inequality.

out with what an ideal experiment would look like and then show how one can use a policy to approximate this ideal. Ideally, one would want to run an experiment in which money was randomly dropped on some school districts but not on others. To ensure that one makes comparisons among similar populations, one could then compare the outcomes of cohorts that were in school during and after the money drop to the outcomes of cohorts from the same school district before the money drop. Note that the fact that the money is dropped from above ensures that it is not driven by the decisions of the individual school districts or parents. The fact that the timing and location of the money drop is random makes it such that the places that received the money drop were not also areas in which families were becoming richer or poorer, etc. Although such a money drop does not exist in reality, School Finance Reforms (SFRs) provide a context that approximates this idealized experiment.

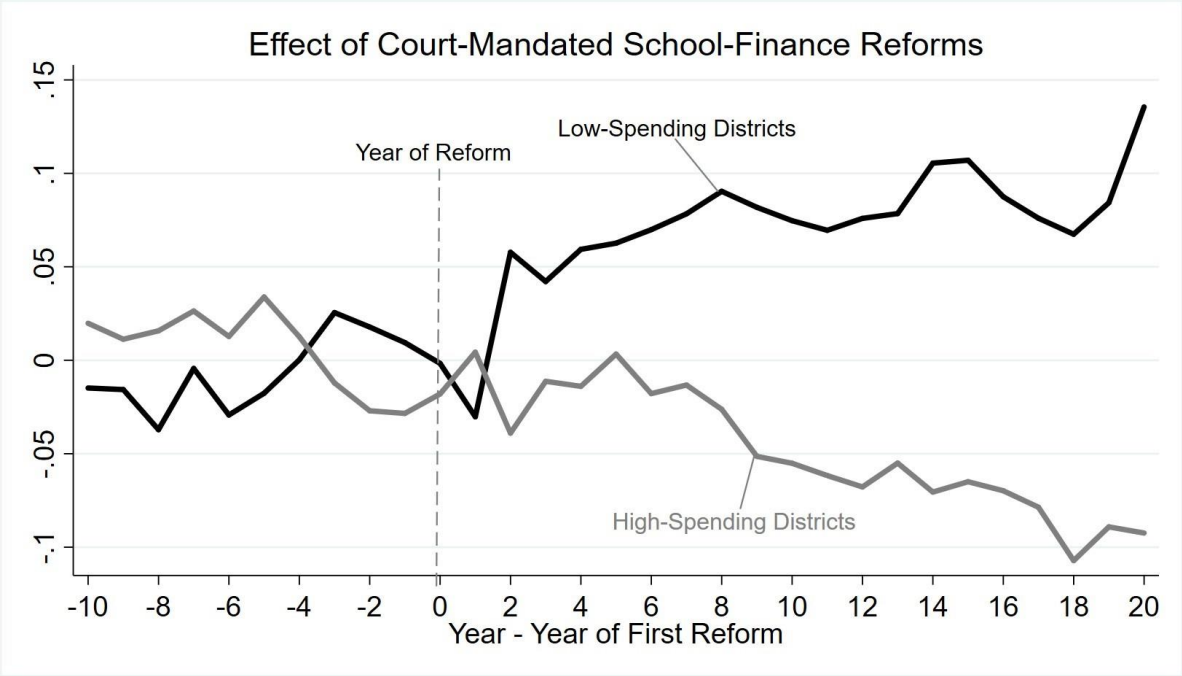
In most states, before the 1970s, local property taxes accounted for most resources spent on K-12 schooling. Because the local property tax base is typically higher in areas with higher home values, and there are high levels of residential segregation by socioeconomic status, heavy reliance on local financing contributed to affluent districts' ability to spend more per student. In response to large within-state differences in per-pupil spending across wealthy/high-income and poor districts, state supreme courts overturned school-finance systems in over 25 states between 1971 and 2010. Because of these court decisions, many states implemented SFRs that led to important changes in public education funding. Most of these court-ordered SFRs changed the parameters of spending formulas to reduce inequality in public-school spending and weaken the relationship between per-pupil school spending and the wealth and income level of the district.

Jackson, Johnson, and Persico (2015) examined SFRs that occurred between 1971 and 1990. Each court-ordered SFR during these years facilitated a mini "policy experiment" for individual districts within a state. We compared the changes in spending in previously low-spending and high-spending districts in years before and after a court-mandated SFR. We classified districts as low- or high-spending based on whether their average per-pupil spending levels were in the bottom or top 25% of districts in their state as of 1972, before any reforms were implemented.

To show how outcomes evolved within districts over time before and after the passage of a SFR, in Figure 9 we show districts' per-pupil school spending levels relative to the year of the passage of the first SFR in the state. As such, for districts

in California year zero would be 1971 (corresponding to Serrano v. Priest, 1971), while for districts in Arkansas year zero would be 1983 (corresponding to Dupree v. Alma school Districts No. 30, 1983). The Y-axis represents the percentage change in spending levels (relative to the average spending levels 10 pears prior to the first SFR). Figure 9 below shows that, in states that passed SFRs, low-spending districts experienced greater increases in per-pupil spending than similar districts in non-reform states, while high-spending districts experienced relative decreases -- reducing spending gaps between previously low- and high-spending districts in reform states. Importantly, we establish that the high and low spending districts were on a similar trajectory of spending (and also outcomes) prior to the passage of a SFR, and we show that the changes in spending that occurred within districts due to the passage of a court-ordered SFR were unrelated to other demographic changes or policies.

Figure 9. *SFR Event Study on Percent Change in Spending by Pre-Reform Spending*



Notes: The Y-Axis represents the percent change in per-pupil spending relative to the average spending levels in the 10 years prior to the court order. Negative values on the X-axis indicate years prior to the court order, whereas positive values indicate years after the court order. Low- and high-spending districs are those in the bottom and top 25% of the per-pupil spending distribution within the state in 1972, respectively. Adapted from "Boosting Educational Attainment and Adult

Earnings," by C. K. Jackson, R. C. Johnson, and C. Persico, 2015, *Education Next, 15*.

Figure 9 shows that (relative to pre-SFR levels), in the lowest spending districts, K12 per-pupil spending increased by about 7 percent 10 years after the court ordered SFR. In contrast, in the highest spending districts, K12 per-pupil spending decreased by about 5 percent 10 years after the court ordered SFR (relative to the pre-SFR levels). However, the pre-reform spending levels are not the only characteristics of school districts that moderate the effect of the SFR on school spending. Many SFRs also allocated additional funds to low-income districts. As such, the pre-reform income levels of districts can also be used to predict the effect of a SFR on the spending levels of a districts. We also noted that some SFRs led to different kinds of funding formulas being enacted, which affect how the SFR would impact school spending. For example, SFRs that led to revenue limits that prohibited local revenues from being above a certain amount reduced per-pupil spending levels much more dramatically in previously high-spending districts than similar districts in other states than did not impose these revenue limits. Following this logic, we used detailed information about both (1) the specific reforms enacted in each state, and (2) the pre-reform income levels and pre-reform spending levels to predict how much of a spending increase each district would receive after the passage of an SFR based on the behaviors of similar districts <u>in other states</u> passing similar kinds of reforms.[27] The above example illustrate that because certain kinds of reforms have systematically led to predictable changes in per-pupil spending on certain kinds of school districts, then one can predict district-level changes in school spending after a court-ordered SFR based only on factors that are unrelated to potentially confounding changes in unobserved determinants (e.g., local commitment to education or the state of the local economy) of school spending and student achievement. This "*predicted spending increase*" reflects the change in school spending within a district that can be directly attributed to the state-level SFR but is unrelated to the choices of parents, demographics changes within the district, or the results of any other policies within the district (recall that this

[27] This prediction is complicated so it may bear some additional explanation. For each district, we create a "doppelganger district" which is the average of similar districts in other states facing similar reforms. Using the experience of similar districts in other states after an SFR, we compute the increase in school spending that these similar districts (with the same pre-reform income level and the same pre-reform spending level) experienced in other states after the passage of an SFR which lead to the same formula changes (e.g., spending limit, or other formula changes). This is our prediction for each district.

prediction is based on the experiences of similar districts facing similar reforms in other states).

If the predicted spending increases have captured the systematic changes in school spending that can be directly attributed to the passage of an SFR, then districts that are predicted to experience larger reform-induced spending increases should actually experience larger school spending increases. We show exactly this pattern in the left panel if Figure 10. We show the change in actual per-pupil spending for cohorts based on the year they turned 17 relative to the year of a court-ordered SFR in their state. As such, a cohort with relative year 0 was 17 years old at the passage of an SFR and is therefore likely just old enough to not be affected by any ensuing reforms. Cohorts with relative years -4 would have been 21 at the passage of an SFR and would therefore be unaffected. In contrast, cohorts with relative year 4 would have been exposed to the SFR for 4 years if they remained in school and would have been 13 at the passage of a SFR. For each cohort within each district, we compute the average level of per-pupil spending in their home district while they were school going age (ages 5 through 17). We then plot the percentage change in school-age per-pupil spending for exposed cohorts (age 17 or younger at the passage of an SFR) compared to unexposed cohorts (older than 17 at the passage of an SFR) for districts with large predicted spending increases (as described above) and those districts with small predicted spending increases. Consistent with the notion that the predicted spending increases (based on the experiences of similar districts facing SFRs in other states) capture real changes in spending, we find that areas with larger predicted spending increases, experience much larger spending increases after the passage of an SFR. Specifically, relative to the cohorts that were unexposed to the reforms (relative years 0 through -8), cohorts that were exposed to 10 to 12 years of reforms in the "*high predicted spending increase*" districts had about 14 percent higher level of per-pupil spending during their school-age years. In contrast, relative to the unexposed cohorts, cohorts that were exposed to 10 years of reforms in the "*low predicted spending increase*" districts had about 4 percent higher actual per-pupil spending during their school-age years. Importantly, in our paper we present several tests indicating that these spending increases (as predicted by the passage of an SFR), are unrelated to underlying economic conditions, demographic changes, and other policies that could affect student achievement.

The increased financial resources that suddenly become available to some cohorts, but not others, within some districts but not available to other districts due to the

statewide passage of an SFR, approximates the money drop analogy laid out above. By relating outcomes with only the reform-induced variation in school spending (rather than all variation in spending), one removes the confounding effect of unobserved factors that might influence both school spending and student outcomes (e.g., other local policies or changes in family background).

Having established that the predicted spending increases due to an SFR, are associated with real spending increases due to the passage of an SFR that are unrelated to confounding factors, we than explore the extent to which the predicted spending increases due to an SFR, are associated with real improvements in student educational outcomes due to the passage of an SFR. To show this, on the right panel of Figure 10, we show the years of educational attainment for the same cohorts as on the left. The right panel shows that exposed cohorts in reform districts predicted to experience larger per-pupil school spending increases during their school-age years did experience larger spending increases, whereas exposed cohorts in reform districts predicted to experience smaller spending increases saw little change in school spending.

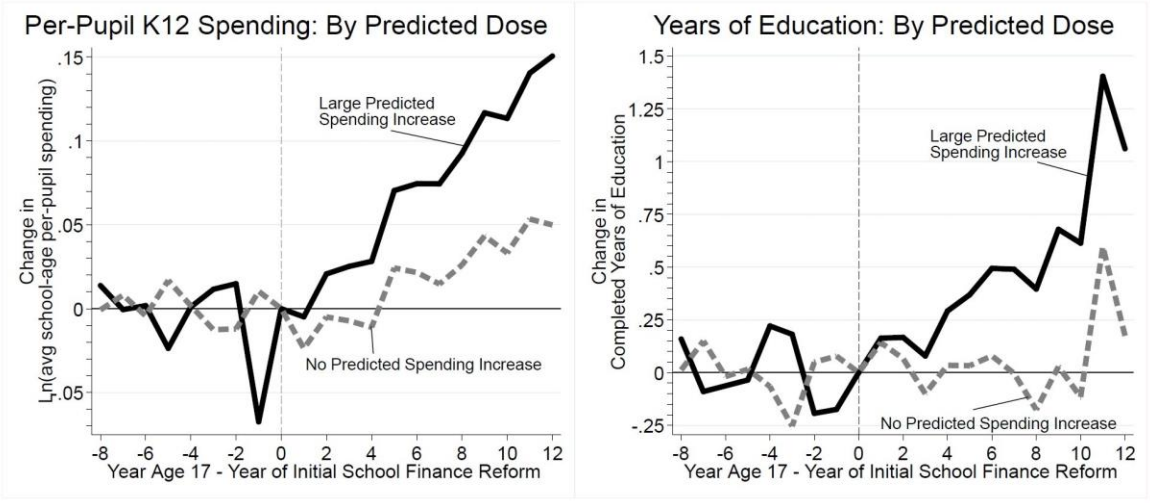
Specifically, relative to the cohorts that were unexposed to the reforms (relative years 0 through -8), cohorts that were exposed to 10 to 12 years of reforms in the "*high predicted spending increase*" districts had about 0.8 more years of completed education. In contrast, relative to the unexposed cohorts, cohorts that were exposed to 10 years of reforms in the "*low predicted spending increase*" districts had about 0.2 more years of completed education. Because these spending changes are not confounded by other factors, one can relate the improvement in years of education associated with the passage of an SFR to the increase in per-pupil spending associated with the passage of an SFR in order to obtain the effect of policy-induced increases in school spending on years of education.

Using these patterns in an instrumental variables framework, we found that a 10% increase in per pupil spending each year for all 12 years of public school led to 0.31 more completed years of education, about 7% higher wages, and a 3.2 percentage-point reduction in the annual incidence of adult poverty. We also found that the effects were more pronounced for children from low-income families. For children from low-income families, a 10% increase in per pupil spending each year for all 12 years of public school is associated with 0.46 additional years of completed education, 9.6% higher earnings, and a 6.1 percentage point reduction in the annual incidence of adult poverty. In follow up work, we also examined effects

on the likelihood of having ever been incarcerated. Here we find that for those from low-income families, a 10% increase in per-pupil spending each year for all 12 years of public school is associated with an 8 percentage point reduction in the likelihood of ever being incarcerated.

Unlike an observational study, this was a design-based study in which the source of the variation in school spending was well-defined and understood. Because the school-spending changes used in this study were driven by state-level legislative action, they did not reflect individual famililies' decisions. Also, because the changes induced by the SFRs were outside of the control of local policymakers (because it was a statewide policy change), they were unrelated to other polices that may have been implemented by local authorities. Importantly, Jackson et al. (2016) show that the policy-induced spending changes were unrelated to underlying economic conditions, demographic changes in local areas, and other potentially co-incident policies -- testing for and ruling-out most plausible sources of bias. Indeed, in one of the most compelling tests, Johnson and Jackson (2018) show that all of these results are similar if one compares siblings from the same family who attend the same school but we exposed to SFRs for different amounts of time due to their ages. Although this empirical approach lends itself to credible causal inferences, this is but a single study. To better understand if school spending matters, it is important to look at a range of different studies, each of which is similarly credible, but that use different data, different methods, and different samples.

Figure 10. *School Finance Reform Event Study by "Predicted Spending Increase"*



Notes: Areas with no predicted spending increase and those with large predicted spending increase doses are those in the bottom and top 25% of the distribution of predicted school finance reform impacts (i.e. dose), respectively. Adapted from "The effects of school spending on educational and economic outcomes: Evidence from school finance reforms," by C. K. Jackson, R. C. Johnson, and C. Persico, 2016, *The Quarterly Journal of Economics*, *131*.

Discussion of Analysis of Key Conclusion 3

Analysis of Delaware Data

Using data provided by the state, we explore specific empirical patterns relating to state spending levels at individual schools, how they varied by student low-income status, and how this may relate to the quality of instruction that low-income students receive. Due to some missing data entries in the 2019 financial reporting, we base our analysis on the spending and performance data for 2018 – the most recent year for which all the data are complete. We also show the extent of underachievement among low-income students in Delaware and explore whether low-income students perform very differently at different schools. We then explore correlations in data between the financial support for instruction in schools and the level of underperformance among low- income students at that school.[28] Finally, we present a description of the schools attended by low- income students in Delaware and relate these characteristics to research on predictors of student success.

Executive Summary of findings

[28] Note that we DO NOT make claims that any of these correlational relationships between spending levels and achievement are causal, and we caution readers that any association between student achievement and state spending level (or lack therefore) has no bearing on how much student outcomes would improve by increasing school spending. To answer that policy question requires an examination of high-quality credible evidence (as conducted in Section 1 of this report).

- Based on the most complete measure of state-level school spending available, the average level of state support (on a per-pupil basis) tends to be higher in schools with higher shares of low-income students. However, after accounting for the prevalence of students with disabilities (a subpopulation that receives additional funds to accommodate special needs), the state tends to spend *less* money (on a per-pupil basis) at schools with higher shares of low-income students. That is, "disability-adjusted" per-pupil state spending is lower at high-poverty schools.

- Because many students (from both low and high-income families) attend schools with a mix of high- and low-income students, there is a small difference in state per-pupil spending between the median low- and non-low-income student. While half of the low-income students attend a school with disability-adjusted state spending levels above $6813 per-pupil, half of the non-low-income students attend a school with state spending levels above $6909 per-pupil – a $ 96 gap. At higher spending levels, there are larger differences. While 10 percent of the low-income students attend a school with disability-adjusted state spending levels above $7868 per-pupil, 10 percent of the non-low-income students attend a school with state spending levels above $8025 per-pupil – a $157 gap. At the highest spending schools, this gap is even larger. While 5 percent of the low-income students attend a school with disability-adjusted state spending levels above $8050 per- pupil, 5 percent of the non-low-income students attend a school with state spending levels above $8226 per-pupil – a $177 per-pupil gap. This shows that the most well-funded schools by the state (after accounting for students with disabilities) are those disproportionately attended by non-low-income students.

- Because some school districts may spend more state funding than others (in ways that could potentially account for the spending difference between low- and non-low-income students), we also explore patterns within districts. Looking at spending disparities between schools within districts, the majority of districts examined have lower per-pupil state spending levels in schools with higher shares of low-income students. *This is true whether one uses raw state spending levels or "disability-adjusted" state spending levels*. That is, after accounting for spending associated with disability status, most districts spend less state money (on a per-pupil basis) on low-income students.

- On average, low-income students perform worse than non-low income students in the state. The share of low-income students scoring below proficient on the Delaware assessment tests is 58% compared to 49% of non-low-income students on ELA. On Math, the percentage of low-income students scoring below proficient on the assessment tests is 70% compared to 58% of non-low income students.

- The share of low-income students scoring below proficient on the assessment tests is not the same across all schools. Some of these differences across schools in the performance of low- income students is structural – that is, (1) schools with higher concentrations of low-income students have worse outcomes for low-income students, and (2) schools with larger shares of students with a disability have lower performance of low-income students. However, some of the differences in performance are related to the quality of instruction provided by the school.

- After accounting for the share of students with disabilities, schools that spend more on teacher salaries (on a per-pupil basis) have lower shares of low-income students scoring below proficient on the state assessment tests. This is consistent with the notion that higher-quality instruction afforded by higher spending levels on teacher salaries is associated with better performance of low- income students on the state test. This also may reflect the difficulties that schools with high levels of student poverty face in attracting and retaining high-quality teachers[29].

- After accounting for students with disabilities, in 10 out of 16 school districts the direct school per-pupil spending from the state is lower in schools that enroll larger shares of low-income students. Even without accounting for students with disabilities, the direct school per-pupil spending from the state in half of the 16 school districts is lower in schools that enroll larger shares of low-income students than non-low-income students.

- Many low-income students attend schools that are also attended by non-low-income students. There is a distribution of the share of low-income students across schools, ranging from 4.7% to 84.6%.

[29] See Clotfelter, Ladd, Vigdor, and Wheeler 2006; Lankford, Loeb and Wyckoff 2002; Jackson 2009, and others for research on this.

- The 95th percentile of state per-pupil direct building spending (i.e., the state-funded part of the "Direct Building Expenditures" from the ESSA sheets) is $8295.025, while the 5th percentile is $5199.432- a $3096 gap. Looking at overall school-level spending (which also includes state spending that goes to districts which is then allocated to individual schools based on Division I units), the 95th percentile of state per-pupil spending for individual schools (direct and district-allocated) is $9527.40, while the 5th percentile is $6784.10- a $2743 gap. To provide a sense of what would happen if spending was increased at the 5th percentile school sufficiently to bring per-pupil funding to the same level of the 95th percentile school, one can use the estimates from the causal literature (Section 1).

- We estimate that if one increased per-pupil school spending by $2743, the percent of low-income students scoring below proficient on the exam would fall by 2.743*(11.6*0.458)=14.6 percentage points. The average of state per-pupil spending for individual schools (direct and district-allocated) is $8114.24. Based on this, if all schools had state per-pupil spending levels brought up to that of the 95th percentile school (all else equal), one would expect the percent of low-income students scoring below proficient on the exam would fall by 1.413*(11.6*0.458)=7.5 percentage points.

- The low-income student graduation rate in Delaware is 77.88, 2017-2018 Delaware Graudation Summary Statistics, June 1019.[30] To provide a sense of what would happen to graduation rates if spending was increased at the 5th percentile school sufficiently to bring per-pupil funding to the same level of the 95th percentile school, one can use the estimates from the causal literature (Section 1).

- At this level, using a gap of $2743, we estimate that if one increased per-pupil school spending by $2743, the percent of low-income students graduating high school would increase by roughly 2.743*(14*0.415)= 15.9 percentage points. We also compute what one might expect if one brought the spending levels of all schools up to that of the school with the 95Th percentile of state per-pupil spending. If all schools had state per-pupil spending levels brought up to that of the 95th percentile school (all else

[30] https://www.doe.k12.de.us/cms/lib/DE01922744/Centricity/Domain/467/GraduationSummaryReport.2017-18.pdf.

equal), one would expect the graduation rate for low-income students to increase by 1.413*(14*0.415)=8.2 percentage points.

SECTION 1: *How State Spending is Distributed Across Schools*

School Demographic Measures

To classify school spending levels on a per-pupil basis, and to make comparisons across schools by the share of low-income students enrolled, we need to use data on student enrollment and demographics at each school. To this aim, we use data on student demographics from Student_Enrollment.csv available at https://data.delaware.gov/Education/Student-Enrollment/6i7v-xnmf. This sheet contains data from 2015-2019. Because the school spending data are incomplete for 2019, for our analysis, we focus on the 2017-2018 school year (coded as SchoolYear = 2018). The sheet includes counts of students enrolled by grade and sub-category of demographic (gender, race, ELL status, low-income status, disability status).

For all of our analysis, we exclude special schools identified for us by plaintiffs' counsel.[31] These excluded schools are "special education schools" for students

[31] These schools are: Douglass School (District 33, School 537), Major Goerge S. Welch Elementary School (District 10, School 612), Kent Elementary Intensive Learning Center (District 10, School 615), Dover Air Force Base Middle School (District 10, School 628), John S. Charlton School (District 10, School 630), Kent County Community School (District 13, School 650), Kent County Secondary ILC (District 13, School 655), Delaware Early Childhood Center (District 15, School 663), North Laurel Elementary School (District 16, School 752), Sussex Consortium (District 17, School 728), Morris (Evelyn I.) Early Childhood (District 18, School 670), Cedar Lane Early Childhood Center (District 29, School 22), Townsend Early Childhood Center (District 29, School 27), Spring Meadow Early Childhood Center (District 29, School 29), Brandywine SITE (District 31), Bush (Charles W.) Pre-School (District 31, School 510), Meadowood Program (District 32, School 516), Richardson Park Learning Center (District 32, School 526), First State School 32 (District , School 530), Brennen School (The) 33 (District , School 538), Delaware School for the Deaf School (DSD) (District 33, School 540), Christina Early Education Center (District 33, School 545), The Colwyck Center (District 34, School 450), Leach (John G.) School (District 34, School 514), The Wallace Wallin School (District 34, School 522), Woodbridge Early Childhood Education Center (District 35 , School 776), Carver (G.W.) Educational Center (District 36, School 689), Ennis (Howard T.) School (District 36, School 750), Kent County Alt Program (District 10, School 633), Appo Adult Ed (District 29, School 654), Early Years (District 32, School 529), Christina Adult Ed (District 33, School 654), Kingswood Academy (District 38, School 501), NCC Adult Education (District 38, School 654), Parkway (District 38, School 792), Delaware Skills Center (District 38, School 4565), POLYTECH Adult Ed (District 39, School 656), Sussex Tech Adult Ed (District 40, School 654).

with particular disabilities, schools that are "disciplinary alternative schools" that are staffed differently than typical public schools, and have a more transient student body. For example, Kent Elementary Intensive Learning Center (District 10, School 615) provides a 100-day program, based on the student learning coping skills and replacement behaviors. The program goal is to integrate the student, as much as possible, into a setting that mirrors the setting the student will return to – this is not a traditional public school. We also exclude pre-school programs, which are also staffed and funded differently from K-12 schools, and we exclude vocational schools. When we refer to all schools in the state, we are referring to all schools except these excluded schools. For our analyses of all schools in the state, we include charter schools (all of which except Charter School of Wilmington and Delaware Military Academy in Red Clay are treated as single-school districts).[32]

Spending measure

One of the key components of our analysis is a measure of state funding at each school. That is, we focus on the level of financial support provided **by the state of Delaware** to individual schools. The spending data we use come from individual district-level sheets created under the federal Every Student Succeeds Act (ESSA) (Title I, section 1111) and provided to us. ESSA requires each school district to report expenditures to their state education agency (SEA), which in turn reports that information to the U.S. Department of Education. Under this reporting requirement, expenditures are reported for each school, by funding source (either with federal funds or with state or local funds), and per-pupil. Using district reports of school-level spending for 2017-2018, we capture two measures of State spending at the school level[33]:

(1) The first measure captures the per-pupil direct building expenditures from state sources (excluding capital and debt payments). This first measure captures spending (both personnel and non-personnel) that occurred in (or on behalf of) each school.
(2) The second measure captures the per-pupil direct building expenditures from state sources **plus** state spending that goes to the district, which is

[32] Note, however, that for analyses that require identifying within-district differences across schools we exclude charter schools that are not part of a particular district.
[33]The names of the individual district files are as follows: SD_0040909, SD_0040914, SD_0040919, SD_0040923, SD_0040928, SD_0040932, SD_0040936, SD_0040940, SD_0040944, SD_0101099, SD_0101103, SD_0101107, SD_0101111, SD_0101115, SD_0101119, SD_0101123, SD_0101127, SD_0101131, SD_0101138, SD_0115656, SD_0115658, SD_0115662, SD_0115663, SD_0115667, SD_0115670

Charter Schools: SD_0101250, SD_0101148, SD_0101152, SD_0101156, SD_0101160, SD_0101164, SD_0101168, SD_0101172, SD_0101176, SD_0101184, SD_0101188, SD_0101192, SD_0101196, SD_0101200, SD_0101209, SD_0101213, SD_0101217, SD_0101221, SD_0101225, SD_0101229, SD_0101235, SD_0101246

then attributed to individual schools in the district by Division I Units (excluding capital and debt payments). The second measure captures state spending (both personnel and non-personnel) that occurred in (or on behalf of) each school in addition to each school's share of any state spending *at the district-level* that may benefit students at the school.

These are two measures of how much State spending supported activities specific to each school. *Note that both these measures exclude capital spending, debt services, and other exclusions.* To facilitate an understanding of these data, we show a sample district sheet below. We have sheets similar to this for 2018 and 2019. Each sheet is separated into District Expenditures (on the left) and School Expenditures (on the right). Figure 11 provides an annotated exemplar from 2018:

The left panel of each sheet (we have one per district) reports aggregate district-level expenditures. There are three major components of district spending:

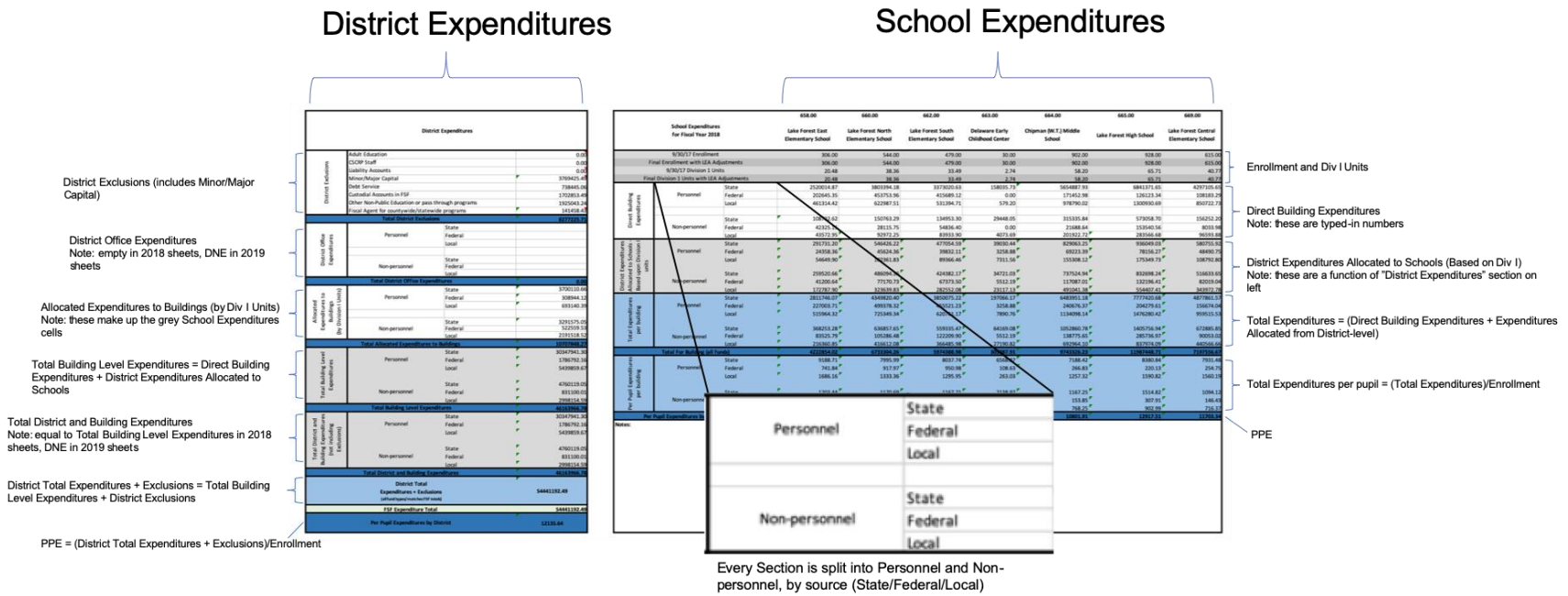
- Direct spending at individual schools (*Total Building Level Expenditures*): This represents school-level spending for operations that occur in the school. This would include teacher salaries, staff salaries, and operations that occur in the individual school.
- District level spending (*Allocated Expenditures to Buildings*): This represents district-level spending that is not specific to any individual school. Note that this includes district-level spending that is funded by the state. As such, this includes state-funded district-level employees and state-funded district level services (in many states, this would include services such as running a school bus) that cannot be solely attributed to a single school.
- District exclusions: (*District Exclusions*): This category is largely comprised of capital spending and debt service. Capital and debt account for 78% of the expenditure in this category. This represents non-operational spending that the ESSA sheets exclude from school-level spending calculations.

The combination of these three represents the total dollar amount spent *at the district level*. This total district-level spending divided by the total student enrollment in the district is reported at the bottom of the left panel and represents

the total dollar amount *at the district level* spent per-pupil in the district.[34] This variable is called "Per Pupil Expenditures by District."

[34] Note: the PPE level reported in the "District Expenditures" portion is a measure of "District Total Expenditures + Exclusions" divided by enrollment. This is *different from* school-level PPE measures, which do not include exclusions. The spending measure we use for each school is from the School Expenditures section of the spreadsheet, and is a measure of Personnel and Non-personnel spending by the state. In 2018 sheets, a section "District Office Expenditures" is blank and "Total District and Building Expenditures" is equal to "Total Building Level Expenditures." Those two sections are not included in 2019 sheets.

Figure 11: *Annotated ESSA sheet for a Sample District in 2018*[35]



[35] Some ESSA sheets contain notes, 2018 sheets include:

SD_0040909 (Caesar Rodney): DAFB Choice - units and enrollment allocated to actual buildings 628 and 612
Funding combined for 628 and 891
Funding combined for 612 and 892
Funding combined for 626 and 893
District 14 Federal enrollment spread to 628,612,626

SD_0101123 (Delmar): • State and local funds are primarily allocated based on enrollment and the entitlement of State Division I units unless prohibited by code and/or regulations. For example, Division II vocational funds must be allocated to the school in which they were earned; therefore, the Delmar High School is allotted more per student than the Delmar Middle School because the High School earns more.
• Additional factors that influence the allocation of state and local funds are:
o Changes in student population, such as Special Education and English Language Learners
o Professional development needs due to changes in curriculum, staff turnover etc.
o Grade span differences, such as additional funds budgeted to the High School for graduation expenses, guidance counseling needs, etc

SD_0040928 (Cape Henlopen): District Office includes, Superintendent, Asst. Super, HR, BOE, Transportation, CNS and Facilities staff and expenditures. Used DOE provided data and sort option

SD_0101107 (Christina): DOE Spread - REACH (512), ILC (535), Douglass (537)

Vocational Units spread between high schools ******OVERSTATEMENT OF EXPENSES****AGENCY 953300 HAS EXPENSE OF ~ $15 Million for Tuition Payments to Agencies 955100,955600,955900,and 956000. These agencies then expend funds for Payroll and Non-Payroll. Detailed information previously provided.

The focus of our analysis is on per-pupil spending at the individual school level. The right panel of each sheet reports expenditures for each school. Note that these school-level expenditures *do not* include the exclusions reported in the district-level spending. The school-level expenses include "Direct Building Expenditure" (which reflects the direct expenditure that occurred in each school) and "District Expenditures Allocated to Schools," which reflects district-level expenses that are divvied-up across schools. For each of these two spending categories, the summary school-level spending measure is disaggregated in subcategories (i.e., Personnel and Non-personnel), where each subcategory reports the revenue source (i.e., State, Federal, and Local sources).

Because we are interested in how the State of Delaware supports individual schools, we aggregate the State contribution to each spending category for each school. Specifically, our measure of Direct Building Expenditures funded by the State is [Direct Building Expenditures on personnel funded by the state] + [Direct Building Expenditures on Non-personnel funded by the state]. Similarly, our measure of District-level spending Allocated Expenditures to Buildings funded by the State is [District Expenditures on personnel Allocated to Schools funded by the state] + [District Expenditures on Non-personnel Allocated to Schools funded by the state]. Our overall operational spending funded by the state measure is the sum of the Direct Building Expenditures funded by the State and District-level spending Allocated Expenditures to Buildings funded by the State. We then divide our spending measures for each school by the student enrollment. We use end-of-year enrollment which represents the total number of students being educated within the school within a school year. This comes from the Delaware Open Data Portal. This results in the following two spending measures:

- Per-pupil school-level operational spending.
- Per-pupil school-level operational spending + Per-pupil school allocation of District-level operational spending.

These two measures reflect the best assessment of the level of state funding provided to each school. The first measure provides an accurate measure of the State's support for activities that can be directly related to the school. If one were to find differences in this measure across schools, it would be suggestive that the

state supports students in some schools more than others. However, if the State were to funds programs at the district level that affect students, it may not be apparent based on this first measure. As such, to ensure that we do not understate state support for students across schools, we also examine the second measure, which includes state spending on district-level expenditures. We show that the patterns are similar for both measures.

a. Distribution of State per-pupil support

To gain a sense of per-pupil support from the state across schools, we first show the overall distribution of state per-pupil school spending across all schools in the state. The histogram of both spending measures is in Figure 12. We also report specific percentiles of the distribution of these two measures in Table 4.

While the average school receives $6831 per pupil from the state using the per-pupil direct expenditures at the school, not all schools receive the same level of support from the state. The 90th percentile of state per-pupil spending is $7881, while the 10th percentile is $5825 - a $2,056 gap. The 95th percentile of state per-pupil spending is $8295, while the 5th percentile is $5199 - a $3,096 difference between the top and bottom 5% of school in terms of support from the state on direct school-level expenditures.

Looking at our second measure, there is also considerable heterogeneity. While the average school receives $8114 per pupil from the state using the per-pupil direct expenditures at the school plus the district-level spending allocated across schools, not all schools receive the same level of support from the state. The 90th percentile of state per-pupil spending using this broader measure is $9315, while the 10th percentile is $7114- a $2201 gap. The 95th percentile of state per-pupil spending is $9527, while the 5th percentile is $6784 - a $2743 gap between the top and bottom 5% of school in terms of support from the state on direct school-level expenditures. The fact that the state support gaps are larger when using the direct school spending plus the district-level spending allocated to schools than using the school-level spending alone indicates that the direct school-level spending gaps that exist

across schools are not offset by state support for district-level expenditures that may go toward low-spending schools.

Figure 12: *Distribution of Per-Pupil Spending by the State*

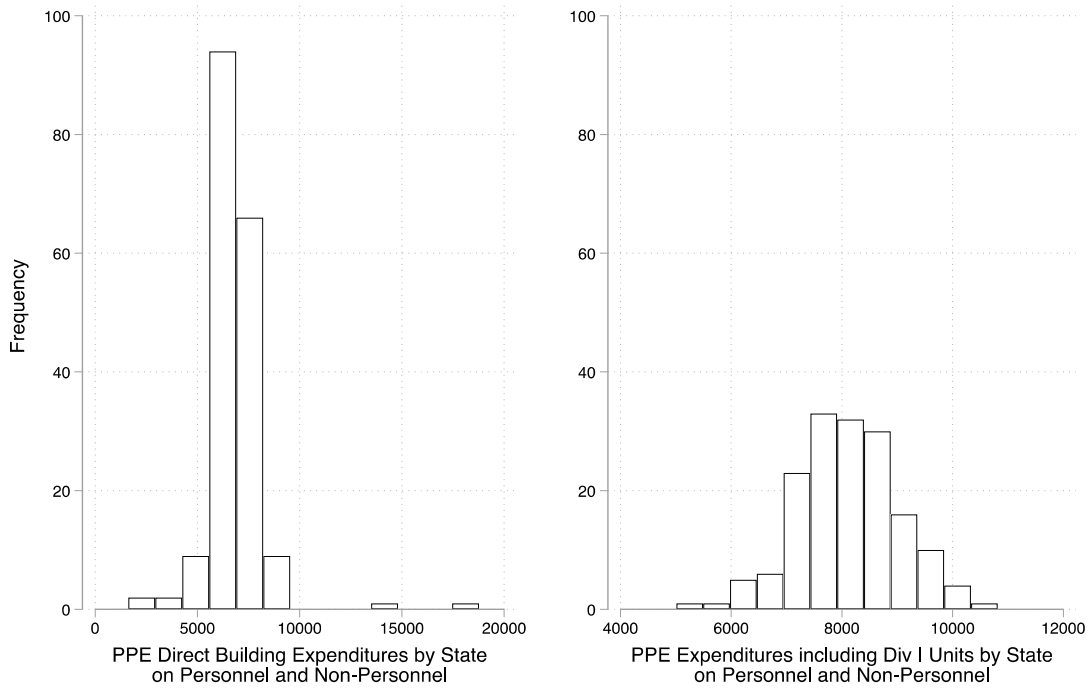


Table 4: *Percentiles of the (School-level distribution) of School-Level Spending by the State*

	5%-ile	10%-ile	25%-ile	50%-ile	75%-ile	90%-ile	95%-ile
Measure 1	$5199.43	$5824.66	$6295.15	$6757.07	$7300.50	$7880.64	$8295.03
Measure 2	$6784.10	$7113.65	$7448.98	$8104.89	$8753.46	$9315.29	$9527.40

Measure 1: Per-Pupil Direct Building Expenditures by State on Personnel and Non-Personnel
Measure 2: Per-Pupil Expenditures by State on Personnel and Non-Personnel (including Division I units)

b. Distribution of State per-pupil support by Share Low-Income

Having shown that there is a distribution of state support across schools, we now explore the extent to which this varies by students' low-income status. Schools are attended by a combination of low-income and non-low income students—schools do not only serve low-income students or only non-low income students. Thus, in any given school, state-funded school-level spending per-pupil represents state support for the education of both low-income and non-low income students. As such, to gain a sense of how state support for school spending may vary by low-income status, one must account for the schools that low income and non-low-income students attend.

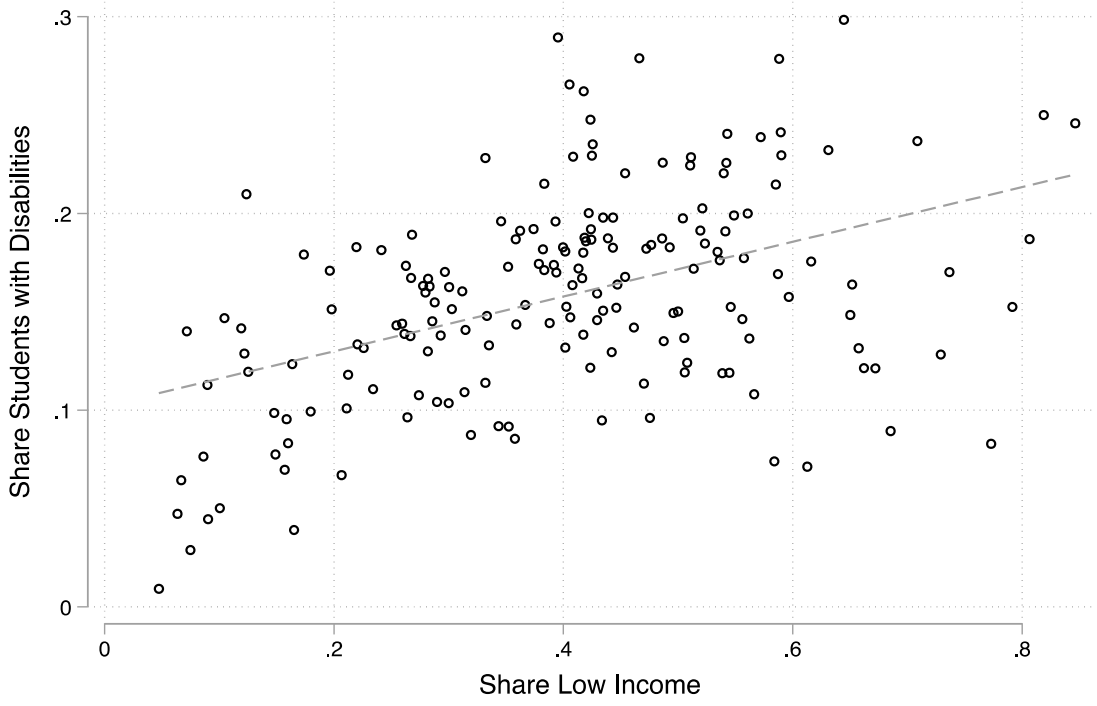
To make this point clear, consider a hypothetical district with one large school (1000 students). All students attend this large school which has 20 percent low income and spends $5000 per pupil. If suddenly a new school was created and it enrolled two students, both low-income, and had a funding level of only $2000 per pupil, there would be a strong relationship between per-pupil spending and the low-income rate. However, as this example makes clear, the experience of the vast majority of low-income students is similar to that of the not-low-income students. In contrast, if the new school were similarly sized (also 1000 low-income students), then a larger share of the low-income students would be exposed to lower levels of spending than the non-low income students in that district. This example highlights the importance of not just relating the share of low-income students to spending levels, but of accounting for the difference in the exposure of students to different levels of school spending by income status.

To conduct an analysis of the spending levels experienced by students, we weight each school's measure of per-pupil spending from the state by the number of low-income students in the school to capture the spending a typical low-income student receives. Similarly, we weight each school's measure of per-pupil spending by the number of non-low income students to capture the spending the typical non-low income student receives. We also compute percentiles of spending for low-income and not-low-income students, to shed light on whether any differences in state support on average come from differences among low-spending schools or high-spending schools.

Accounting for Student with Disabilities

Before presenting the patterns in the data, it is impotent to discuss how we deal with students with disabilities. Under the Individuals with Disabilities Act, 20 U.S.C. § 1400, *et. seq.*, the state must provide all children with disabilities a free appropriate public education that emphasizes special education and related services designed to meet their unique needs and prepare them for further education, employment, and independent living. As such, schools that enroll larger shares of students with disabilities will, and should, have *more* per-pupil funding from the state (all else equal). As such, if two schools have similar overall per-pupil spending levels, but one school enrolls a larger share of students with disabilities, the students at the school with more students with disabilities have lower levels of effective core operational spending (since a greater share of the money is going to special education services). In reality, schools that have high shares of students with disabilities are disproportionately attended by low-income students (see Figure 13). A failure to account for this would lead one to conclude that the schools attended by low-income students are better funded by the state (for core operational spending) than they are in reality.

Figure 13: *Relationship Between %Disability and % Low Income*



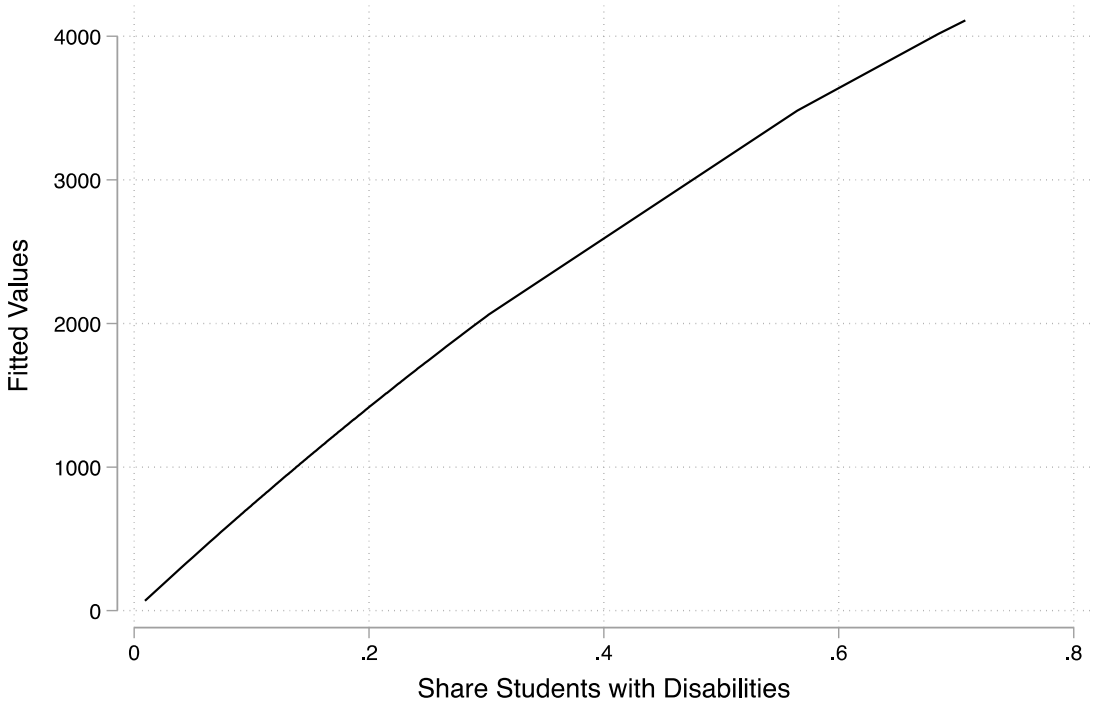
Note: This illustrative plot restricted to the 95 percent of schools with fewer than 30 percent of students with a disability.

The discussion above highlights that it is essential to account for the share of students with disabilities at each school and to adjust the per-pupil spending levels accordingly. To this aim, we construct a "disability-adjusted" measure of school spending for each school. To do this, we compare the spending levels at similar schools (i.e., from the same districts and with similar shares of low-income students) to estimate the increase in spending associated higher levels of students with disabilities. Figure 14 below plots the estimated relationship between the share of students with a disability in a school and the per-pupil spending coming from the state (both directly at the school and also to the districts which is attributed to the school). Based on our estimates, each percentage-point increase in the share of students with a disability is associated with roughly 75 additional dollars per-pupil from the state – but this relationship is slightly nonlinear.

With this measure, we use the share of students with a disability at each school and subtract the predicted per-pupil spending amount due to the share of students with a disability at each school. The result is a "disability-adjusted" per-pupil spending level for each school.[36] This "disability-adjusted" per-pupil spending level represents our estimate of the per-pupil spending level one would observe at a school if it had no students with disabilities.

Figure 14: *Effect of Share with Disability on State-funded School Spending (Measure 2)*

[36] More specifically, we do the following: We run a linear regression predicting each school's school-level state spending measure on the share of students with disabilities (and its square), the share of low-income students and district-fixed effects. This regression isolates the change in school-level state-supported spending associated with higher shares of students with disabilities, while also accounting for the share of low-income students, and the average spending level in the same school district. Our adjusted school spending measure is the spending level net of the estimated contribution of students with disabilities.



Distribution of spending among Low-income and non-low-income students.

In Table 5 we report the percentiles of the student-level distribution of state per-pupil spending for low-income students and non-low-income students in Delaware. The state per-pupil spending levels reported at any percentile X represents the state per-pupil spending level, below which X percent of students are exposed to. For example, if the 75th percentile of spending (across all students) is $8753, then 75 percent of all students have per-pupil spending levels below $8753. Similarly, if the 25th percentile of spending (across low-income students) is $7449, then 25 percent of all low-income students have per-pupil spending levels below $7449.

Table 5: *Percentiles of the student-level distribution of school-level spending by the state: by low-income status*

	5%-ile	10%-ile	25%-ile	50%-ile	75%-ile	90%-ile	95%-ile
	Not-Low Income						
Measure 1	$5239.13	$5864.16	$6358.07	$6759.92	$7265.14	$7715.18	$8142.87
Measure 2	$6793.82	$7181.56	$7491.05	$8130.47	$8727.25	$9315.29	$9477.93
	Low Income						
Measure 1	$4926.79	$5705.92	$6193.17	$6720.25	$7259.88	$7724.97	$8135.65
Measure 2	$6578.16	$7113.65	$7448.97	$8105.42	$8727.25	$9315.29	$9527.39

Measure 1: Per-Pupil Direct Building Expenditures by State on Personnel and Non-Personnel
Measure 2: Per-Pupil Expenditures by State on Personnel and Non-Personnel (including Division I units)

Table 5 reports the percentiles of the raw unadjusted state-supported spending at the school level based on our two measures by income status. Recall that measure 1 is direct state per-pupil spending at the school, and measure 2 is direct state per-pupil spending at the school **plus** state-funded district spending allocated to the school. Table 2 reports the distribution of spending for each of these measures BEFORE making any adjustment for students with disabilities (we present the distribution of the disability-adjusted measures later on in Table 6). As discussed above, these raw estimates will *understate* the effective spending for those schools that enroll high shares of students with disabilities, which also tend to be schools that enroll larger shares of low-income students. Using the direct-building spending funded by the state (measure 1), the 5th percentile of spending is $5239.13 for non-low-income students and $4926.79 for low-income students. This indicates that schools attended by the bottom 5 percent of low-income students receive about $312 less per-pupil from the state then schools attended by the bottom 5 percent of non-low-income students. Looking at the middle of the distribution, the 50th percentile of spending is $6759.92 for non-low-income students and $6720.25 for low-income students. While the unadjusted spending levels are similar, they are about $39 higher for non-low-income students. Looking at the top of the distribution, the 90th percentile of spending is $7715.18 for non-low-income students and $7724.97 for low-income students. That is, at the top of the distribution, state support for the low-income students is about $9 more than for the not-low-income students. Finally, at the very top of the distribution (the 95th percentile), state support for the low-income students is about $7 less than for the not-low-income students. In sum, based on the unadjusted measure 1, low-income students attend schools with slightly lower state support across the distribution until the very top of the distribution, and the difference is most pronounced at the lower end of the distribution.

We also report the distributions for the unadjusted measure 2. Generally, the patterns are similar for state support for individual schools, including the district-level spending allocated to individual schools. That is, at the bottom of the distribution (the 5th percentile), state support for the non-low-income students is about $215 more than for the low-income students. At the 50th percentile, state support for low-income students is about $25 more than for the not-low-income students. At the 90th percentile, state support for the low-income students is the same as for the non-low-income students. At the very top of the distribution (the 95th percentile), state support for the low-income students is about $50 more than for the not-low-income students.

Because none of this analysis accounts for students with disabilities, the patterns presented in Table 5 understate the effective difference in state per-pupil spending between the schools attended by low- and not-low-income students – i.e., the difference in state per-pupil spending that is not attributable to higher shares of students with disabilities at schools that have larger shares of low-income students. Table 6 presents the comparable figures using disability-adjusted school spending measures. Based on adjusted measure 1 (i.e., direct state per-pupil spending at the school), the low-income students attend schools with lower per-pupil state support at all percentiles. That is, at the bottom of the distribution (the 5th percentile), state support for the low-income students is about $300 less than for the not-low-income students. At the 50th percentile, state support for low-income students is about $131 less than for the not-low-income students. At the 90th percentile, state support for the low-income students is about $114 less than for the not-low-income students. At the very top of the distribution (the 95th percentile), state support for the low-income students is about $26 less than for the not-low-income students.

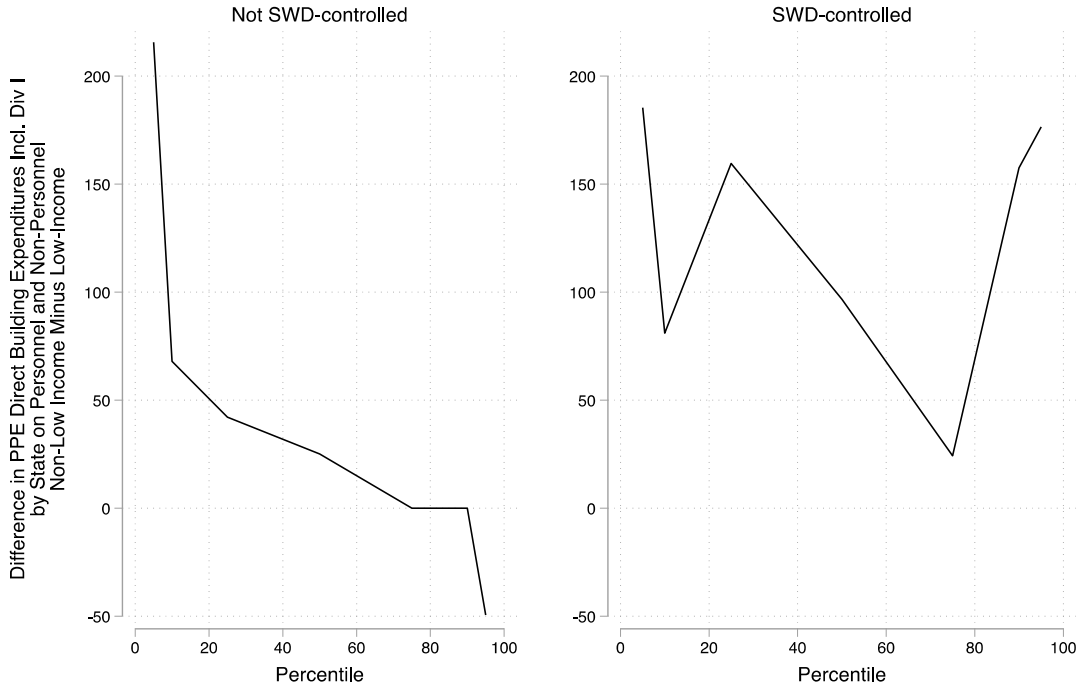
Table 6: *Percentiles of Disability Adjusted School-Level Spending by the State by Low-Income Status*

	5%-ile	10%-ile	25%-ile	50%-ile	75%-ile	90%-ile	95%-ile

	Not-Low Income						
Measure 1	$4648.14	$5042.59	$5597.62	$6076.46	$6551.41	$7027.72	$7326.72
Measure 2	$5657.09	$6035.23	$6485.84	$6909.48	$7446.26	$8025.18	$8226.30
	Low Income						
Measure 1	$4343.00	$4918.59	$5468.09	$5945.21	$6425.12	$6913.90	$7300.40
Measure 2	$5471.72	$5954.20	$6326.34	$6812.87	$7421.99	$7867.72	$8049.84

Measure 1: Per-Pupil Direct Building Expenditures by State on Personnel and Non-Personnel
Measure 2: Per-Pupil Expenditures by State on Personnel and Non-Personnel (including Division I units)

Figure 15: *Gap in the Spending at Different Percentiles of the Distribution of State Per-Pupil spending by Income Status (spending measure 2)*



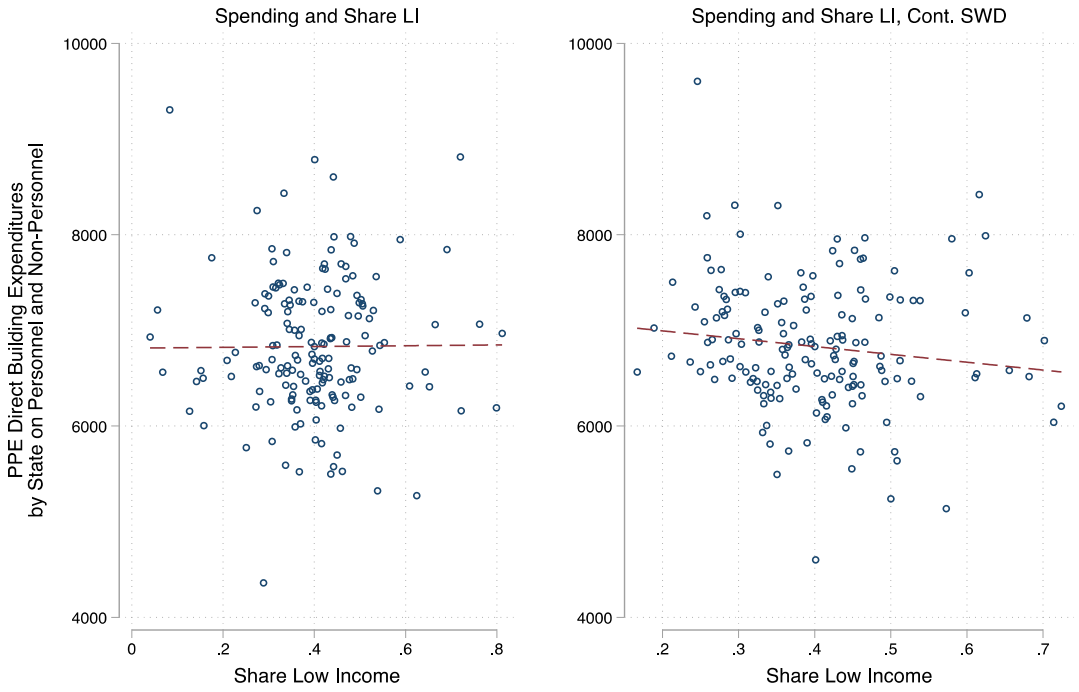
Using the overall measure, measure 2, that combines the direct school spending coming from the state and the state-funded district-level spending ascribed to the individual school, the adjusted spending gaps between low income and non-low income are even larger. For ease of presentation, we plot the state spending gap between non-low income and low-income students for the different percentiles in Figure 15. The state per-pupil spending gaps by income using the unadjusted measure 2 are on the left and the analogous gaps for the disability adjusted

spending gaps is on the right. At all percentiles, the level of state support (adjusted for disability) for schools is lower for low-income students. That is, at the bottom of the distribution (the 5th percentile), state support for the low-income students is about $185 less than for the not-low-income students and at the 10th percentile, state support for low-income students is about $81 less than for not-low-income students. At the 50th percentile, state support for low-income students is about $97 less than for not-low-income students. At the 90th percentile, state support for the low-income students is about $157 less than for the not-low-income students. At the very top of the distribution (the 95th percentile), state support for the low-income students is about $176 less than that of the not-low-income students. The fact that the adjusted spending gaps were large at the top end of the distribution indicates that (a) the schools that were well funded by the state were also those schools with high shares of students with disabilities, and (b) after accounting for spending associated with students with disabilities, the school that received the most per-pupil support from the state were schools that were disproportionately attended by higher-income students.

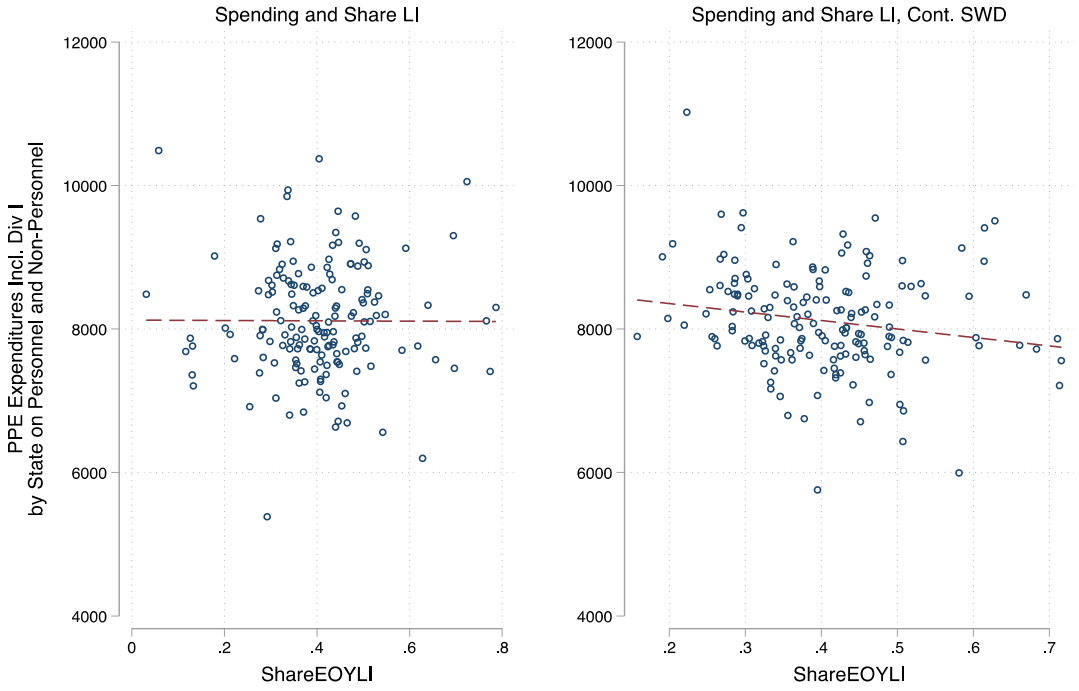
Relationship Between State Spending at Schools and Share Low-Income

The analysis of the percentiles of state per-pupil spending by students' income status suggests that schools that enroll larger shares of low-income students tend to receive lower levels of financial support from the state (after accounting for students with disabilities). We test this directly with linear regression. To show these relationships visually, we plot the unadjusted spending measure 1 against the share of low-income students at the school in the left panel of Figure 16 part a. One can see a tendency for measure 1 to be larger in schools with larger shares of low-income students. However, after accounting for the share of students with disabilities, this pattern is reversed. That is, the adjusted state support for schools (excluding the state-funded district-level spending) is lower in schools with higher shares of low-income students.

Figure 16: *Relationship Between State-Per-Pupil spending and Share Low Income at the School*


Spending and Share LI
Spending and Share LI, Cont. SWD
PPE Direct Building Expenditures
by State on Personnel and Non-Personnel
10000
8000
6000
4000
0
.2
.4
.6
.8
.2
.3
.4
.5
.6
.7
Share Low Income
Share Low Income

a.


Spending and Share LI
Spending and Share LI, Cont. SWD
PPE Expenditures Incl. Div I
by State on Personnel and Non-Personnel
12000
10000
8000
6000
4000
0
.2
.4
.6
.8
.2
.3
.4
.5
.6
.7
ShareEOYLI
ShareEOYLI

b.

Section b of Figure 16 shows the same plots for the spending measure 2, which includes both direct spending at the school and district-level spending attributed to the school. As with measure 1, unadjusted spending is higher at schools with larger low-income shares, but adjusted state support is lower in schools with large shares of low-income students. In sum, once one accounts for the share of students with disabilities, the state spends less on low-income students on a per-pupil basis. Put differently, among schools with the same share of students with disabilities, schools with larger shares of low-income students receive less money (per-pupil) from the state.

We estimate these linear relationships using multiple regression and report the results in Table 7. Using the complete spending measure (i.e., measure 2) we find that without accounting for the share of students with disabilities, on average, a 10 percentage point increase in the share low-income is associated with about $16 *less* per-pupil support from the state (column 1). After accounting for the share of students with disabilities, a ten percentage point increase in the share low-income is associated with about $107 *less* per-pupil support from the state (column 2). If we also compare schools within the same districts, we find that a ten percentage point increase in the share of low-income students is associated with about $118 *less* per-pupil support from the state (column 3). In sum, after accounting for students with disabilities, both overall and within individual districts, state spending is lower in schools with higher shares of low-income students.

Table 7: *Regression of State Per-Pupil Spending Against % Low income and % with Disability*

	(1)	(2)	(3)
	State Per-Pupil Spending Incl. Div I		
Share Low-Income	-160.0	-1074.1	-1180.5*
	(-0.29)	(-1.68)	(-2.43)
Share with Disability		8661.9*	7588.3
		(2.53)	(2.05)
Share with Disability Squared		-3629.9	-2517.4
		(-0.77)	(-0.53)
Constant	8178.9***	7195.3***	7384.1***
	(31.53)	(18.38)	(14.31)
Number of Schools Included	162	162	162

t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Patterns of Spending by Low-Income Status Within Districts

The previous analysis shows that on average (across the state of Delaware), both overall and within individual districts, state spending is lower in schools with higher shares of low-income students. To assess whether this pattern is robust, we also present the relationship between state spending and the share of low-income students across schools within each individual district. For this, we regress the measures of school spending from the state on the share of low-income students for each district, and report the slope in Table 8. Note that Table 8 reports the

relationships for measure 1 (direct state per-pupil spending at a school). We report these slopes both for the unadjusted spending measure and the disability-adjusted measures. Within each district, we also test whether the average low-income student attends a school with less state support than the average not-low-income student in that same district. Recall that this measure may not necessarily indicate the same pattern of the slope because school sizes may vary by the proportion of low-income students and the pattern that holds across *schools* on average may not hold across *students* on average. When both tests reveal the same pattern, it shows that the pattern is robust.

First, we focus on measure 1. Even without adjusting for the share of students with disabilities, the slope in 8 of the 16 traditional districts is negative (column Slope), indicating that in half of the districts, as the share of low-income students increases at a school the per-pupil spending from the state falls. Also, in 9 of these 16 districts, the average low-income student receives less per-pupil funding from the state than the typical non-low income student (column Avg. Non-LI > Avg. LI). After accounting for the share of students with disabilities, 10 of 16 districts have a pattern of the state spending being less in schools with a higher share of low-income students (column Slope (Cont. SWD)), and in 10 of 16 districts the average low-income student receives less spending by the state on education than the average non-low income student (column Avg. Non-LI > Avg. LI (Cont. SWD)). In sum, even without accounting for students with disabilities, in half of those school districts in Delaware, the direct school per-pupil spending from the state is lower in schools that enroll larger shares of low-income students. After accounting for students with disabilities, in 10 out of 16 (62.5 percent) of the traditional districts in Delaware, the direct school per-pupil spending from the state is lower in schools that enroll larger shares of low-income students. To assess the extent to which this pattern may be offset by district-level spending funded by the state that is distributed to individual school, we also examine the relationships for spending measure 2.

Table 8: *Relationship Between State Per-Pupil Direct Spending (Measure 1) and Share Low Income: by District*

	PPE State on Personnel and Non-Personnel			
	Slope	Avg. Non-LI > Avg. LI	Slope (Cont. SWD)	Avg. Non-LI > Avg. LI (Cont. SWD)
District 10	847.64	Yes	-583.33	Yes
District 13	1502.11	No	-1082.9	No
District 15	-3132.21	Yes	-548.51	Yes
District 16	-737.74	Yes	428.76	Yes
District 17	4729.95	No	5593.84	No
District 18	-2115.62	Yes	-722.02	Yes
District 23	-4316.49	Yes	2875.26	Yes
District 24	1575.39	No	-1413.53	Yes
District 29	3739.02	No	1760.54	No
District 31	2163.77	No	1299.45	No
District 32	-603.88	Yes	-1882.46	Yes
District 33	934.46	No	-709.12	No
District 34	3507.46	No	463.15	No
District 35	-7884.05	Yes	-9065.82	Yes
District 36	-4663.35	Yes	-9196.07	Yes
District 37	-38975.65	Yes	-40722.67	Yes

The results for Measure 2, which is an overall measure of state per-pupil spending at the school, are even more pronounced (See Table 9). Even without adjusting for the share of students with disabilities, the slope in 8 of 16 districts is negative (column Slope), and in 9 of 16 districts the typical low-income student receives less spending by the state on education than the average non-low income student (column Avg. Non-LI > Avg. LI). Controlling for the share of students with disabilities, 10 of 16 districts have a pattern of the state spending less in schools with a higher share of low-income students (column Slope (Cont. SWD)), and in 11 of 16 districts the typical low-income student receives less spending by the state on education than the average non-low income student (column Avg. Non-LI > Avg. LI (Cont. SWD)). The pattern for both spending measures is consistent with most districts spending less state money on low-income students (after accounting for students with disabilities). In sum, even without accounting for students with disabilities, in more than half of the non-special districts in Delaware, the per-pupil

state spending (both direct spending and district allocated) is lower for the average low-income student than the average not-low-income student. After accounting for students with disabilities, in 11 out of 16 (68.75 percent) of the non-special districts in Delaware, the per-pupil state spending (both direct spending and district allocated) is lower for the average low-income student than the average not-low-income student.

Table 9: *Relationship Between State Per-Pupil Overall Spending (Measure 2) and Share Low Income: by District*

	PPE State on Personnel and Non-Personnel Incl. Div I			
	Slope	Avg. Non-LI > Avg. LI	Slope (Cont. SWD)	Avg. Non-LI > Avg. LI (Cont. SWD)
District 10	1907.43	Yes	381.58	Yes
District 13	1805.38	No	-705.97	No
District 15	-3677.23	Yes	-803.93	Yes
District 16	-757.81	Yes	637.58	Yes
District 17	4432.76	No	5814.87	No
District 18	-2010.1	Yes	-297.78	Yes
District 23	-5043.74	Yes	3821.96	Yes
District 24	4489.97	No	-2314.66	Yes
District 29	3585.14	No	1621.92	No
District 31	2498.1	No	1244.53	No
District 32	-713.72	Yes	-2092.85	Yes
District 33	517.99	No	-1387.99	Yes
District 34	2478.81	No	-1126.22	No
District 35	-8999.78	Yes	-10434.49	Yes
District 36	-4560.6	Yes	-9862.32	Yes
District 37	-47012.67	Yes	-49156.11	Yes

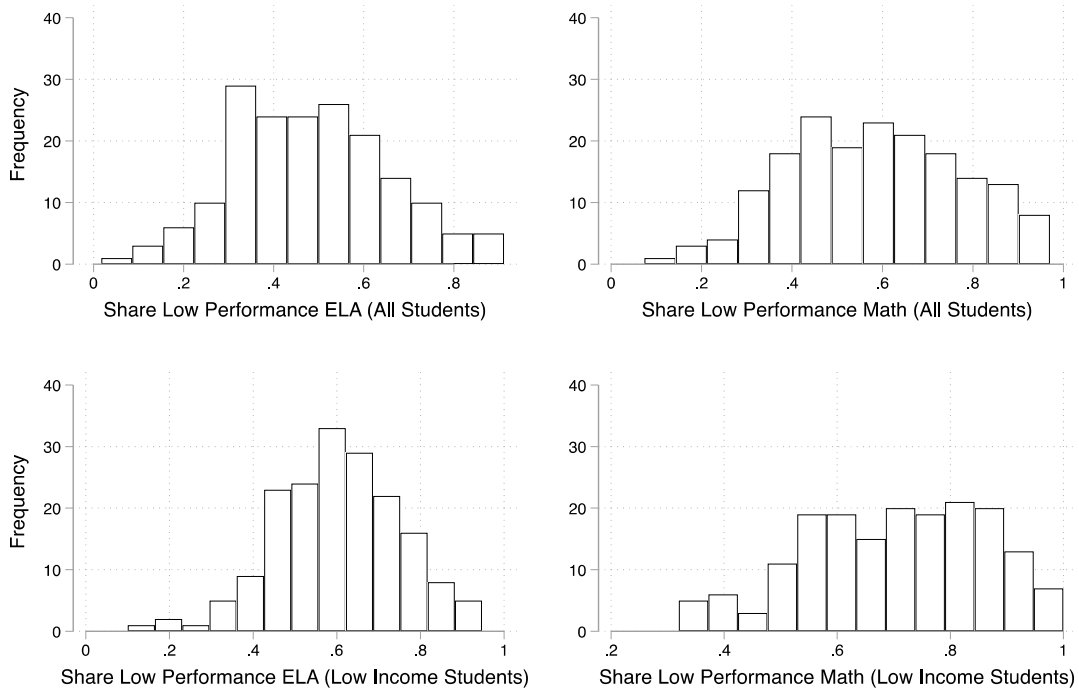
c. Student Performance

In this section of the report, we analyze student achievement data provided by the state (SD_0015691_UR_HIGHLY CONFIDENTIAL) to show patterns of student achievement (focused on low-income students) in Delaware. The meta-analysis in

Section I of the report shows that increased school spending improves student outcomes. The aim of this section is to assess the extent to which there is a particular need for improvement among low-income students in Delaware.

To provide a broad overview of student underachievement in Delaware, in Figure 17, we plot the distribution of the proportion of students in each school scoring in Levels 1 and 2 on the Delaware state assessments tests in ELA (left) and Math (right) in 2018. We plot the distributions for all students (top) and also only for low-income students at each school (bottom).

Figure 17: *Distribution of the proportion of students in each school scoring in Levels 1 and 2 on the Delaware state assessments: by Subject*



These histograms reveal three broad patterns: First, the share of students scoring not proficient is higher in Math than in English Language Arts (ELA). Looking at all students, the average school has a not-proficient rate of 58.4 percent compared to 48.8 percent for ELA. Second, low-income students are much more likely to

score not proficient on the state assessment than the population as a whole. That is, while the average percentage of all students scoring not proficient on the state ELA assessment is 48.7 percent, that for low-income students is 60.1 percent. Similarly, while the average percentage of all students scoring not proficient on the state Math assessment is 58.4, that for low-income students is 70.1. These patterns indicate a need for improvement, particularly for low-income students.

The third key pattern is that there is considerable variation across schools in how students perform. Looking only at low-income students, 10 percent of schools have a Math not-proficient rate of over 91 percent for low-income students, while 10 percent of schools have a Math not-proficient rate below 50 percent for low-income students. This suggests that some schools are getting relatively better results among their low-income students than other schools. Indeed, over one-quarter of schools have a Math not-proficient rate below 60 percent for low-income students – which is the average for all students. As evidenced by the histograms above, there are also substantial differences in the performance of low-income students across schools on the ELA assessment.

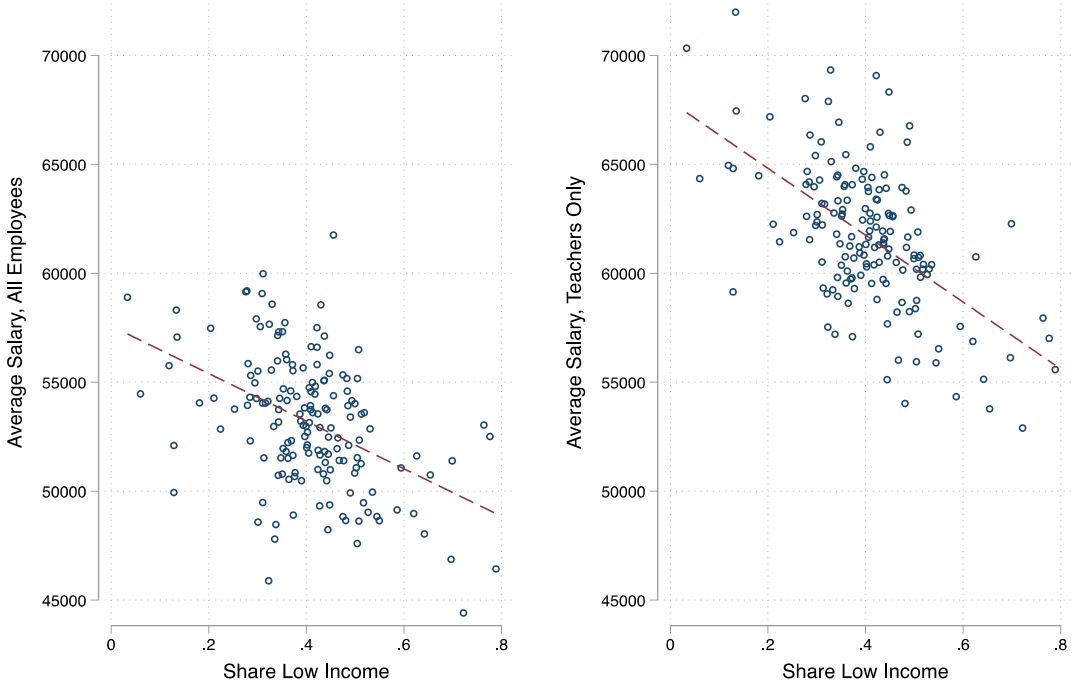
The fact that low-income students have very different levels of performance across schools is important. It suggests that school-level factors may play a vital role in the performance of low-income students. In the next section, we explore school-level factors associated with school spending that are related (in a descriptive rather than a causal sense) to some schools having better outcomes for low-income students than others.

d. Student Performance and Quality of Inputs

Most money spent by the state is spent on personnel versus non-personnel, and while we do not have a precise measure of the quality of educational inputs, one proxy for quality is teacher salary. Note that the majority portion of teacher salaries is from state sources (65%). As such, showing the relationships between teacher salaries and student outcomes provides some insight into the mechanisms through which increases in state support may influence students. Our measure of teacher salaries comes from the file SD_0012441, which includes the salaries paid to all

employees of school districts and charters in the State of Delaware in 2018 (from Peoples Dep. Tr. 15-53).

Figure 18: *Scatterplot of The Average Salary Paid to Employees of the Schools against the share of Low-Income students enrolled at the School*

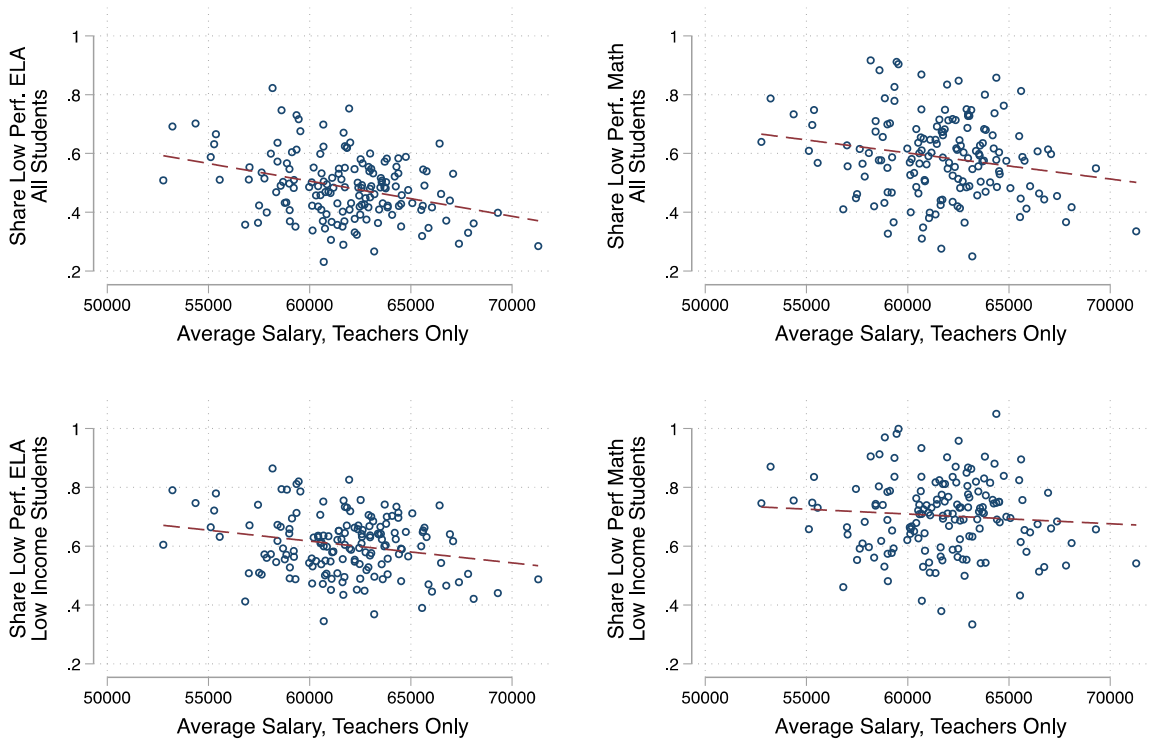


First, to gain a sense of how this particular input varies across schools, we present a scatterplot of the average salary paid to teachers at a school against the share of low-income students at that school in Figure 18. In the left panel, the average salary paid to all employees is lower in schools with higher shares of low-income students. The right panel shows the average salaries paid to teachers. Again, this indicates that schools that enroll higher shares of low-income students have teachers with lower average salaries.

Given that teachers are mostly paid on a salary scale (mainly based on years of experience, advanced degrees, and additional credentials), this is direct evidence that schools with higher shares of low-income students tend to have teachers with

fewer years of experience and other credentials. There is a strong evidence base having a more experienced teacher improves student achievement (Clotfelter, et. al. 2006; Clotfelter, et. al. 2007; Wisall 2013; Jackson and Bruegmann 2009, and others) so that this pattern alone would predict lower student achievement among low-income students on average, and also lower achievement of low-income students at schools that enroll higher shares of low-income students.

Figure 19: *Scatterplot of Share of students Scoring Not-Proficient on the Delaware state assessment by Average Teacher salary at the School: by income status*

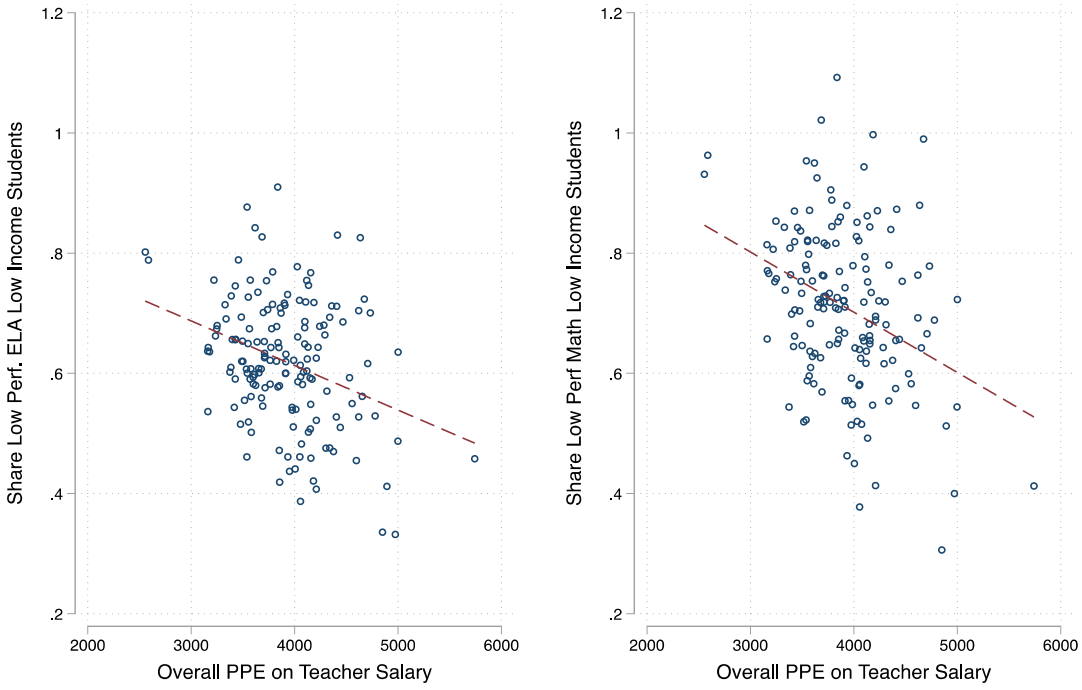


With the caveat that we are not saying that these relationships are causal (see the meta-analysis for Key Conclusion 1 for that), we assess the extent to which average teacher salaries are associated with better outcomes for students in general and low-income students in particular. Figure 19 presents a scatterplot of the share of students scoring non-proficient in Math and ELA for low-income students and overall. These models show the relationships after accounting for the share of

students with disabilities. This control is particularly important when comparing achievement levels because students with disabilities tend to have lower performance on state tests, which is also correlated with low-income status and spending levels. The first key pattern is that the share of students scoring non-proficient in Math and ELA is lower in schools with higher average teacher salaries.

Importantly, the positive relationship between student performance and average teacher salaries is directly related to the per-pupil spending on teacher salaries. To show this, we present a scatterplot of the share of low-income students scoring not-proficient against the per-pupil spending on salaries (after accounting for students with disabilities). Consistent with the salary patterns, schools that spend more on teacher salaries per pupil have fewer low-income students who are scoring not proficient on the state Math and ELA assessments. To quantify this relationship, we regress the share not-proficient low-income students against the salaries paid per-pupil at the school. The results indicate that schools with $1000 higher teacher salaries paid per pupil have 8.4 percent lower non-proficiency rates on the Math tests and 5 percent lower non-proficiency rates on the ELA tests. Combining both estimates (Math and ELA), these regressions suggest that a $1000 higher per-pupil spending on salaries would be associated with a 6.7 percentage point lower non-proficient for low-income students. Coincidentally, the causal studies indicate that a $1000 higher per-pupil spending would *cause* a 5.31 percentage-point decline in the non-proficiency rate.

Figure 20: *Relationship Between the Share of Low-Income Students Scoring Not-Proficient on the Delaware state assessment and Per-Pupil Salary Expenditure at the School (irrespective of the funding source)*



In terms of the share of low-income students performing poorly on the state assessment tests, we see that a much lower share of low-income students performs poorly in schools with a higher average salary and schools that spend more per-pupil on teacher salaries. While these are merely correlations, the data show that better-resourced schools—where those resources go toward teacher salaries—have better student outcomes. These patterns are consistent with the casual literature.

In sum, the patterns show that low-income students have lower levels of achievement than non- low-income students. The patterns also show that many schools with higher shares of low-income students have lower levels of school inputs, due in part to lower levels of support from the state (as established in the previous subsection). We show that some schools have good outcomes for their low-income students while others do not, and we show that the per-pupil spending on teacher salaries is correlated with better low-income student outcomes across

schools. Interestingly, the correlational relationships between per-pupil teacher salaries are similar to those implied by existing high-quality causal research showing that low-income students' outcomes could be improved considerably with increases in per-pupil spending (which can be supported by the state).

Table 10: *Student Performance By Per-Pupil Spending on Teacher Salaries (controlling for the share of students with disabilities)*

	(1)	(2)	(3)	(4)
	Across Districts		Within-District	
	Share Low Perf. ELA, Low Income	Share Low Perf. Math, Low Income	Share Low Perf. ELA, Low Income	Share Low Perf. Math, Low Income
PPE On Salary, Teachers Only	-0.0000225	-0.0000396	-0.0000495	-0.0000839*
	(-1.00)	(-1.51)	(-1.23)	(-2.25)
Share with Disability	1.282**	1.221**	4.260***	4.512***
	(3.33)	(3.03)	(5.24)	(5.28)
Share with Disability Squared	-0.721	-0.761	-9.425***	-9.824**
	(-1.38)	(-1.41)	(-4.21)	(-3.53)
Constant	0.500***	0.675***	0.368	0.570**
	(5.40)	(6.67)	(2.00)	(3.50)
Number of Schools Included	176	176	156	156

t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

e. Distribution of Low-income students (correlates)

To assess the extent to which improving schools can be targeted to low-income students in Delaware, it is important to document the distribution of low-income students across schools. If schools were very highly segregated by income, then most schools would have either very low shares or very high shares of low-income students. In such a scenario, the outcomes of low-income students could conceivably be improved by promoting improved outcomes—potentially through increased spending—in a small set of schools that enroll a large share of low-income students. However, if low-income students are somewhat distributed across all schools (i.e., all schools had equal shares of low-income students), then improving the outcomes of low-income students would entail promoting better, more-resourced schools across the board. The reality falls somewhere between these two extreme cases. To show this, we construct a histogram of the percent low-income across all schools in the state:

Figure 21: *Distribution of the share of Low-Income Students Across Schools*

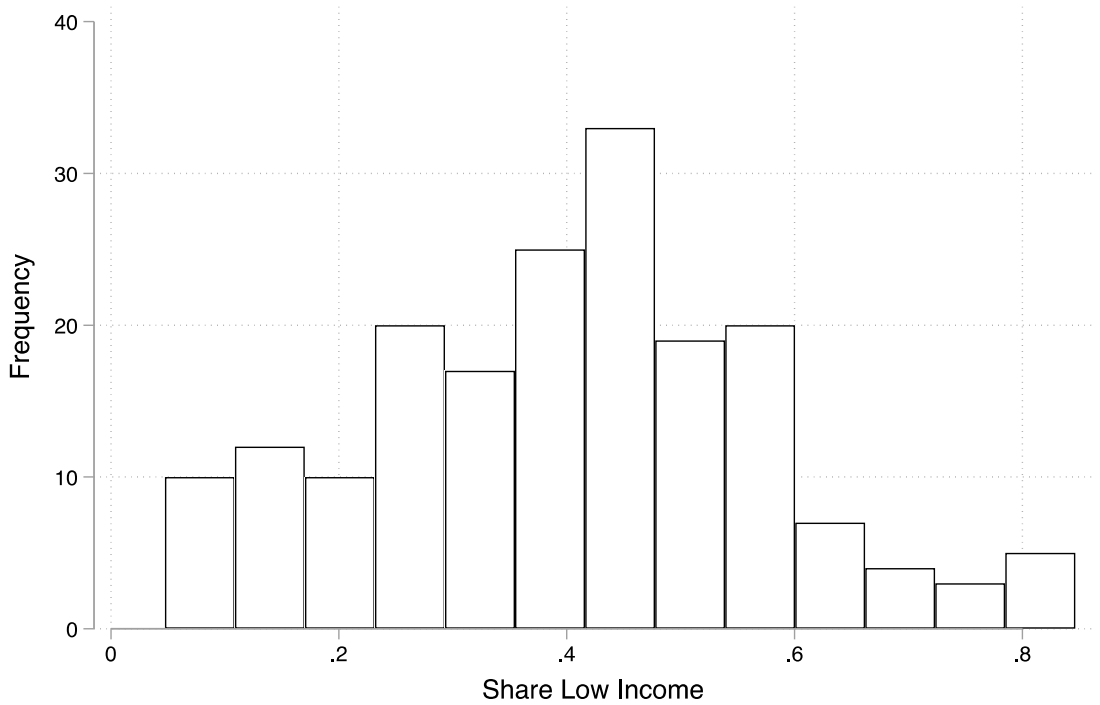


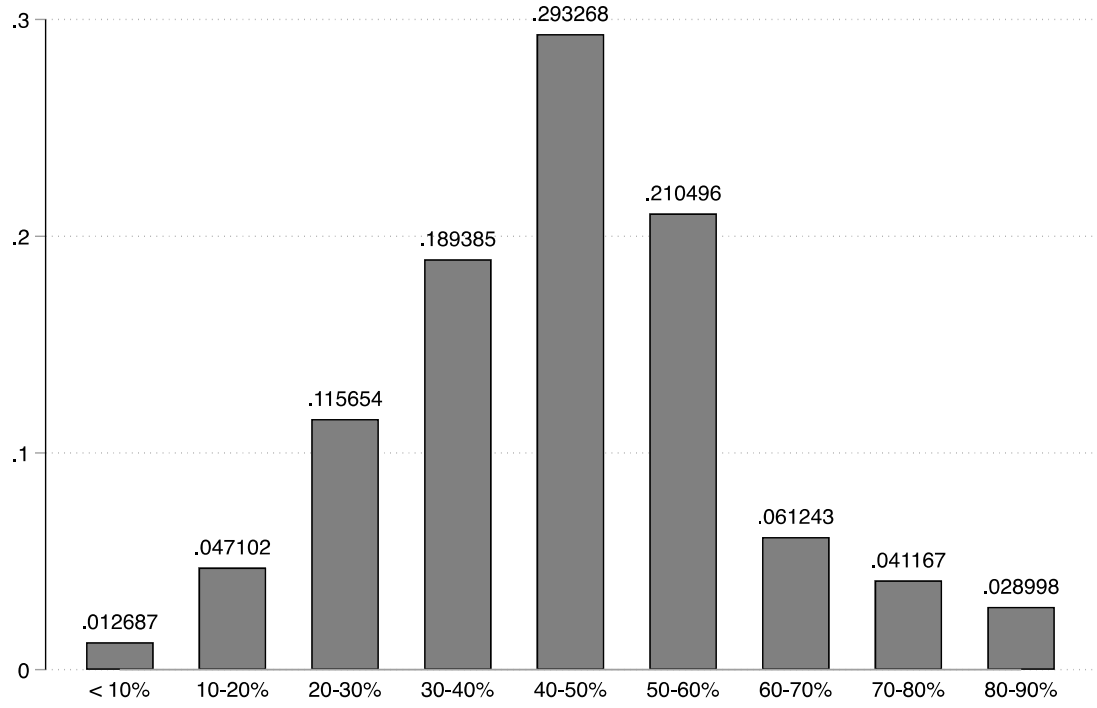
Table 11: *Percentiles of the share of Low-Income students*

	5%-ile	10%-ile	25%-ile	50%-ile	75%-ile	90%-ile	95%-ile
Share Low Income	0.1044	0.1599	0.2820	0.4085	0.5113	0.6127	0.6854

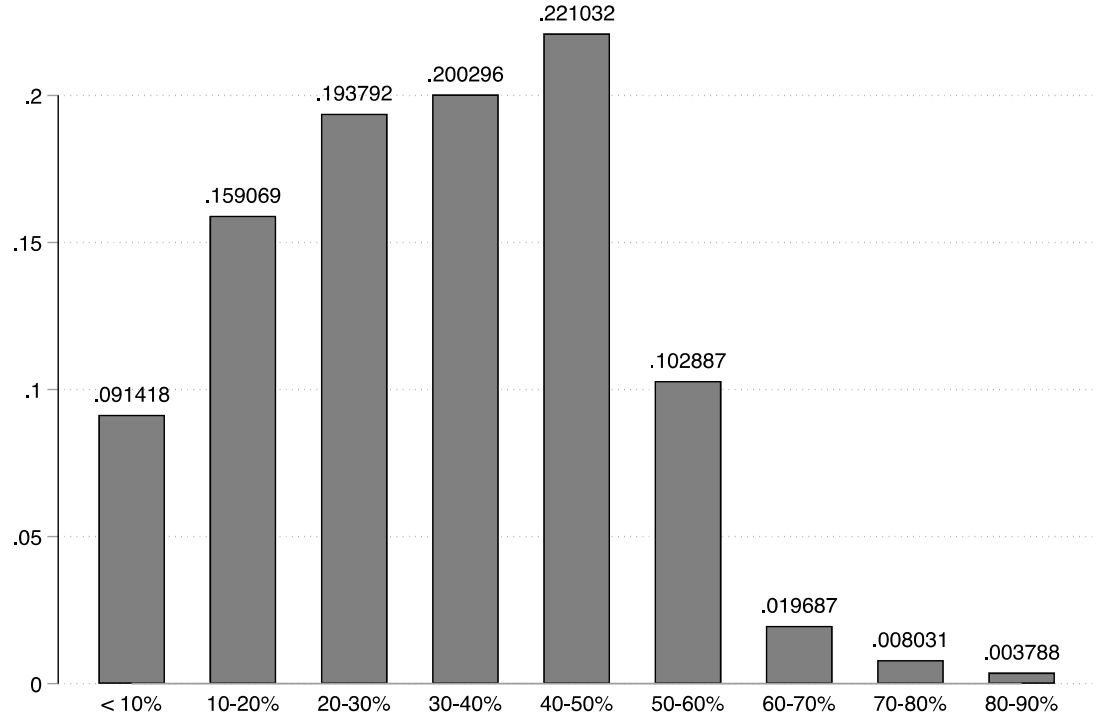
Across the state, about 30% of students are low income. If there were a perfectly even distribution of low-income students, all schools would have 30% percent low-income students. As one can see, the low-income students were not evenly distributed across all schools. About half of all schools enroll between 28.2 and 51.1 percent low-income, one-quarter of school have fewer than 28.2% low income, and one-quarter have more than 51.1% low income. Looking at the highest and lowest shares, 10 percent of schools had more than 61.6% low-income students, while 10 percent of schools had fewer than 16% low-income students.

The analysis above shows that some schools have very high concentrations of poverty, while others have very low concentrations. However, if the highest-poverty schools are small relative to more integrated schools, the proportion of low-income students attending schools with concentrated poverty may be small. Conversely, if the highest poverty schools are large relative to other schools, the proportion of low-income students attending schools with concentrated poverty may be very high. To shed light on the exposure of low-income students to concentrated poverty, we examine the share of all low-income students attending schools with different poverty shares in the top panel of Figure 22. For a basis for comparison, we also show analogous shares for non-low income students in the lower panel.

Figure 22: *Percent of Low-Income Students attending Schools by Share Low-Income at School*



a.



b.

The data show that many students attend somewhat mixed schools. That is, about 30-40% percent of students from both income groups attend schools with between 20.1 and 40 percent low-income. However, there are marked differences for schools with high (above 50%) and low (20% or lower) shares of low-income students. While about a quarter (25%) of non-low income students attend schools with fewer than 20 percent low income, only about 6% of low-income students do. However, a third of low-income students attend schools with more than 50% low-income compared to only 13 percent for those who are not low-income. Looking at schools with even higher levels of concentrated poverty, about 13 percent of low-income students attend a school with 60 percent or more low-income students (compared to about 3% for non-low income students).

This exposure to high levels of poverty is an important result because higher levels of concentrated poverty are associated with worse outcomes independently of a students' own poverty-status. That is, all students tend to have worse outcomes in schools that have high concentrations of poverty irrespective of their own poverty status. For example, higher concentrations of poverty are associated with higher teacher turnover (e.g., Loeb et al. 2005), lower teacher qualifications (e.g., Jackson 2009), and lower levels of school safety (e.g., Garcia and Weiss 2019). All of these factors influence the educational experience but have nothing to do with the actions of individual parents or children. To provide an equal quality of educational ***inputs*** (i.e., instructional quality, the safety of the school environment, etc.) to children in low- and high-poverty schools would require spending more in high-poverty schools than in low-poverty schools. I use a specific example to make this clear. It is well documented that teachers, on average, tend to choose to teach in schools with low levels of poverty than schools with high levels of poverty. Given this, paying teachers the same in all schools tends to result in higher teacher turnover and more substitute teaching in high poverty schools than in low poverty schools. This example makes clear that equal spending does not ensure equal instructional quality. If one wanted to provide the same quality of instruction across high- and low-poverty schools, one would need to pay teachers *more* in the higher poverty schools.

Another implication of these patterns is that policies that seek to target low-income students can do so by promoting improvements in schools that enroll high-shares

of low-income students. That is, interventions targeted to schools with more than 50% low income would potentially impact a third of low-income students but only about 13 percent of non-low income students. *It also bears mentioning that there is a threshold that determines categorization as low-income. As such, the non-low-income students at schools with high shares of low-income students are likely to be relatively low-income families themselves.* To reach a greater share of low-income students would require targeting schools that also enroll nontrivial shares of non-low income students. Indeed, targeting intervention at schools that enroll more than 20% low-income students would affect over 90% of all low-income students and about 75% of non-low income students.

Poverty and Disability Status

As one considers the relationship between spending and poverty status, it is important to consider the relationship between the low-income share and the share of students who are classified as disabled. Under the Individuals with Disabilities Act (IDEA) of 2004, districts and charters have an obligation to ensure that all students with disabilities have available to them a free and appropriate public education (FAPE) in the least restrictive environment (LRE) that emphasizes special education and related services designed to meet their unique needs, and prepares them for further education, employment, and independent living. Districts and charters must ensure that the rights of children with disabilities and their families are protected (link to source).

Given that additional supports are required for schools that enroll students with disabilities, one must account for these additional funds when making comparisons between spending levels in high- and low-poverty schools. To assess the extent to which this may be something to consider, we estimated the relationship between the share of students with a disability and the share of low-income students. Figure 3 shows a scatterplot of the percent low income against the percent of students with a disability at the school (among schools with less than 30% with a disability – over 95 percent of schools). This figure shows a clear positive relationship. Schools with higher proportions of low-income students are also schools with higher shares of students with disabilities. Remarkably, a linear regression reveals that for each one-percentage-point increase in the share of students with disabilities, the share low-income increases by 1.2 percentage points.

This pattern has important implications for how we compare spending levels across high and low-poverty schools. Given that the state spends additional resources at schools with larger shares of students with disabilities, the spending levels at high-poverty schools (which will include more of these additional funds for students with disabilities) do not accurately reflect the level of state support for non-special education low-income students. Indeed, low poverty schools could spend more on a per-pupil basis for non-special education services, but have lower levels of overall spending that high-poverty schools (since a much greater share of spending at high poverty schools may go to special education services). This pattern highlights the fact that comparisons between high and low-poverty schools will *understate* the extent to which high-poverty schools may spend less than low-poverty schools.

Note on 2019 spending data

As discussed earlier in the report, we do not use the 2019 data for our analysis because the 2019 spending data are not complete. That is, while the ESSA sheets provided for 2019 had reasonably complete looking data on district's finances, the school-level financial data were incomplete. Indeed, in one instance, District 17 (sheet: SD_0115662) is missing many, though not all, entries for Direct Building Expenditures. In lieu of using the 2019 data, about which we have concerns, we focus our analysis on the 2018 data, which appear to be reliable.

Papers Cited:

Baron, E. Jason. "School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin." *Available at SSRN 3430766* (2019).

Borenstein, Michael, Larry V. Hedges, Julian PT Higgins, and Hannah R. Rothstein. *Introduction to meta-analysis*. John Wiley & Sons, 2011.

Carlson, Deven, and Stéphane Lavertu. "School improvement grants in Ohio: Effects on student achievement and school administration." *Educational Evaluation and Policy Analysis* 40, no. 3 (2018): 287-315.

Cascio, Elizabeth U., Nora Gordon, and Sarah Reber. "Local responses to federal grants: Evidence from the introduction of Title I in the South." *American Economic Journal: Economic Policy* 5, no. 3 (2013): 126-59.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. "The value of school facility investments: Evidence from a dynamic regression discontinuity design." *The Quarterly Journal of Economics* 125, no. 1 (2010): 215-261.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American economic review* 104, no. 9 (2014): 2633-79.

Clark, Melissa A. "Education reform, redistribution, and student achievement: Evidence from the Kentucky Education Reform Act." *PhD diss. Princeton University* (2003).

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. "Teacher-student matching and the assessment of teacher effectiveness." *Journal of human Resources* 41, no. 4 (2006): 778-820.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. "Teacher credentials and student achievement: Longitudinal analysis with student fixed effects." *Economics of Education Review* 26, no. 6 (2007): 673-682.

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

Downes, Thomas A., Richard F. Dye, and Therese J. McGuire. "Do limits matter? Evidence on the effects of tax limitations on student performance." *Journal of Urban Economics* 43, no. 3 (1998): 401-417.

Dynarski, Susan, Joshua Hyman, and Diane Whitmore 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.

Figlio, David N. "Did the "tax revolt" reduce school performance?." *Journal of Public Economics* 65, no. 3 (1997): 245-269.

Frisvold, David E. "Nutrition and cognitive achievement: An evaluation of the School Breakfast Program." *Journal of public economics* 124 (2015): 91-104.

Garcia, Emma, and Elaine Weiss. "Challenging Working Environments (" School Climates"), Especially in High-Poverty Schools, Play a Role in the Teacher Shortage. The Fourth Report in" The Perfect Storm in the Teacher Labor Market" Series." *Economic Policy Institute* (2019).

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

Guryan, Jonathan. *Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts*. No. w8269. National Bureau of Economic Research, 2001.

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

Hong, Kai, and Ron Zimmer. "Does investing in school capital infrastructure improve student achievement?." *Economics of Education Review* 53 (2016): 143-158.

Hoxby, Caroline M. "All school finance equalizations are not created equal." *The Quarterly Journal of Economics* 116, no. 4 (2001): 1189-1231.

Husted, Thomas A., and Lawrence W. Kenny. "Evidence on the impact of state government on primary and secondary education and the equity-efficiency trade-off." *The Journal of Law and Economics* 43, no. 1 (2000): 285-308.

Hyman, Joshua. "Does money matter in the long run? Effects of school spending on educational attainment." *American Economic Journal: Economic Policy* 9, no. 4 (2017): 256-80.

Jackson, C. Kirabo. *Does school spending matter? The new literature on an old question*. No. w25368. National Bureau of Economic Research, 2018.

Jackson, C. Kirabo. "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation." *Journal of Labor Economics* 27, no. 2 (2009): 213-56. Accessed March 7, 2020. doi:10.1086/599334.

Jackson, C. Kirabo, and Elias Bruegmann. "Teaching students and teaching each other: The importance of peer learning for teachers." *American Economic Journal: Applied Economics* 1, no. 4 (2009): 85-108.

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. "The effects of school spending on educational and economic outcomes: Evidence from school finance reforms." *The Quarterly Journal of Economics* 131, no. 1 (2016): 157-218.

Jackson, C. Kirabo, Jonah E. Rockoff, and Douglas O. Staiger. "Teacher effects and teacher-related policies." *Annu. Rev. Econ.* 6, no. 1 (2014): 801-825.

Jackson, C. Kirabo, Cora Wigger, and Heyu Xiong. *Do school spending cuts matter? Evidence from the Great Recession*. No. w24203. National Bureau of Economic Research, 2018.

Johnson, Rucker C., and C. Kirabo Jackson. "Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending." *American Economic Journal: Economic Policy* 11, no. 4 (2019): 310-49.

Kogan, Vladimir, Stéphane Lavertu, and Zachary Peskowitz. "Direct democracy and administrative disruption." *Journal of Public Administration Research and Theory* 27, no. 3 (2017): 381-399.

Lafortune, Julien, and David Schonholzer. *Do school facilities matter? measuring the effects of capital expenditures on student and neighborhood outcomes*. Working Paper, 2017.

Lee, Kyung-Gon, and Solomon W. Polachek. "Do school budgets matter? The effect of budget referenda on student dropout rates." *Education Economics* 26, no. 2 (2018): 129-144.

Loeb, Susanna, Linda Darling-Hammond, and John Luczak. "How Teaching Conditions Predict Teacher Turnover in California Schools." *Peabody Journal of Education* 80, no. 3 (2005): 44-70. Accessed March 7, 2020. www.jstor.org/stable/3497042.

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

Matsudaira, Jordan D., Adrienne Hosek, and Elias Walsh. "An integrated assessment of the effects of Title I on school behavior, resources, and student achievement." *Economics of Education Review* 31, no. 3 (2012): 1-14.

Miller, Corbin L. "The Effect of Education Spending on Student Achievement: Evidence from Property Values and School Finance Rules." Working Paper, 2018.

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

Papke, Leslie E. "The effects of changes in Michigan's school finance system." *Public Finance Review* 36, no. 4 (2008): 456-474.

Roy, Joydeep. "Impact of school finance reform on resource equalization and academic performance: Evidence from Michigan." *Education Finance and Policy* 6, no. 2 (2011): 137-167.

Schanzenbach, Diane Whitmore. "Does class size matter?." (2014).

Van der Klaauw, Wilbert. "Breaking the link between poverty and low student achievement: An evaluation of Title I." *Journal of Econometrics* 142, no. 2 (2008): 731-756.

Weinstein, Meryle G., Leanna Stiefel, Amy Ellen Schwartz, and Luis Chalico. "Does Title I Increase Spending and Improve Performance? Evidence from New York City. Working Paper# 09-09." *Institute for Education and Social Policy* (2009).

Wiswall, Matthew. "The dynamics of teacher quality." *Journal of Public Economics* 100 (2013): 61-78.

/s/ C. Kirabo Jackson, Ph.D.

March 12,2020