

Point column on school spending: Money matters

C. Kirabo Jackson¹ | Claudia Persico²

¹Department of Human Development and Social Policy, Northwestern University

²Department of Public Administration and Policy, American University

Correspondence

C. Kirabo Jackson, Department of Human Development and Social Policy, Northwestern University, 2120 Campus Drive, Room 204, Annenberg Hall, Evanston, IL 60208.

Email: Kirabo.jackson@gmail.com

Handling editor: Kevin Kelly and Paul Decker

Here, we will make our best guess, based on the highest-quality evidence available, of what the effect would be of a policy that increases school spending on student outcomes. First, we establish a few key empirical facts about the historical relationship between school spending and student outcomes and also what the existing older literature found about this relationship.

TEST SCORES HAVE INCREASED AT THE SAME TIME AS SCHOOL SPENDING

A prevalent counterargument to the notion that higher school spending correlates with improved student achievement relies on the examination of national test scores over time, as measured by the National Assessment of Educational Progress (NAEP). Because school spending has increased by 130% since the 1970s, one might expect a corresponding increase in achievement, assuming all other factors remain constant. Naturally, not all factors are constant; however, thoughtful analysis of the data reveals that test scores increased dramatically over this period.

The NAEP is administered in English and math in grades 4, 8, and 12. Figure 1 displays scores on these tests from 1978 to 2012 for all students and then by ethnicity (White, Black, Hispanic). Math scores have risen for all grade levels and ethnic groups. In English, gains have been less pronounced, which aligns with the common finding that school-level interventions generally have a greater impact on math than language skills. There were modest gains for all groups in 4th and 8th grades. The average 9-year-old in 2011 had the same English and math proficiency as a 10- and 11-year-old in 1978, respectively. While 12th-grade English scores changed little on average, there were substantial gains among Black and Hispanic students (groups that disproportionately experienced increased school spending). In sum, test scores have improved for almost all grade levels and ethnic groups. This is progress.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Policy Analysis and Management* published by Wiley Periodicals LLC on behalf of Association for Public Policy and Management.

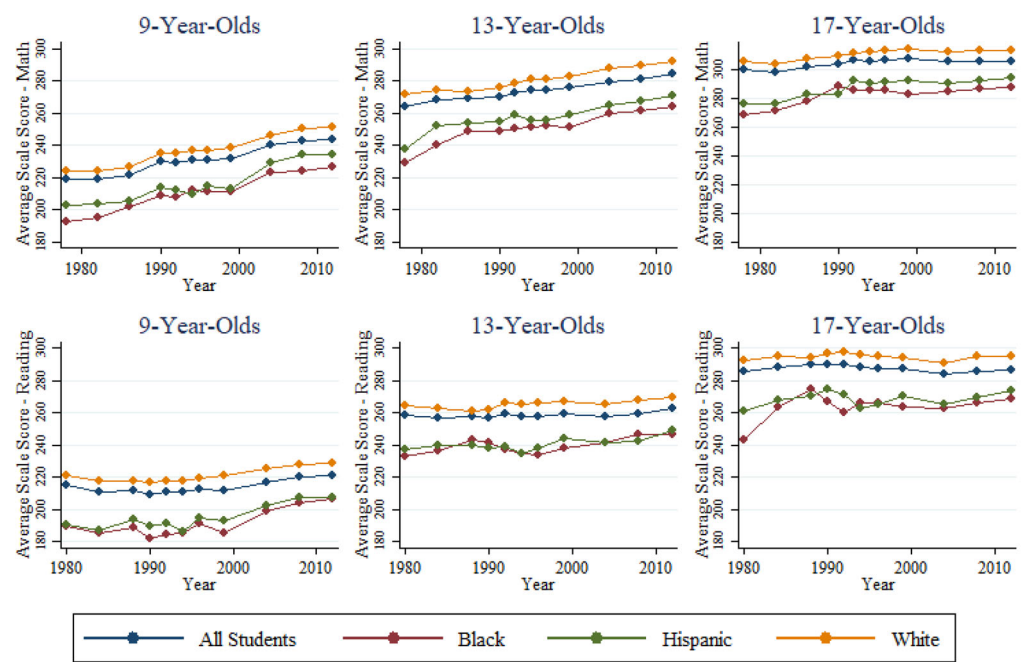


FIGURE 1 NAEP math and reading score improvement, 1978 to 2012, by race and age. [Color figure can be viewed at wileyonlinelibrary.com]

If the scores at all grade levels were equally reliable, one might argue that 12th-grade scores (which increased by less) instead of younger-grade scores provide a better reflection of what students ultimately know at the end of their K–12 schooling. However, the tests are not equally reliable because changes in the 12th-grade NAEP reflect both changes in knowledge and also potentially sizable compositional changes. The compositional changes that bias trends in the 12th-grade NAEP (but not the 4th- or 8th-grade NAEP) occur for two reasons. First, 4th- and 8th-grade scores represent the knowledge of students educated in the public school system, as most students have not dropped out by this grade. In contrast, 12th-grade scores only account for the performance of those who remain in school through 12th grade. As high school graduation rates have increased over time, more students who would likely have low NAEP scores take the 12th-grade test. This may cause 12th-grade scores to decrease over time, even if there were no changes in knowledge. Second, because schools must participate in the 4th- and 8th-grade NAEP (but not 12th grade) to receive Title I funding, participation rates for the 4th- and 8th-grade NAEP have been consistently high (over 90%) while those for the 12th-grade NAEP have varied between 66% and 91% (National Center for Education Statistics, 2017). For these two reasons, trends in 12th-grade NAEP scores may represent both knowledge and significant composition changes, making them unreliable. A reasonable person would trust 4th- and 8th-grade scores more, as they are not biased by compositional changes. Importantly, these more reliable tests have been rising over time along with school spending.

As evidence that changes in 8th-grade scores reflect persistent changes in skill (which changes in 12th-grade scores may not), Doty et al. (2022) found that birth cohorts with higher 8th-grade NAEP scores have improved educational attainment, earn higher incomes, and have lower teen motherhood, incarceration, and arrest rates. These authors find no such relationship for 12th-grade NAEP scores, making it indefensible to ignore the 8th-grade NAEP trends in favor of the potentially-biased 12th-grade NAEP trends.

Using the more reliable 4th- and 8th-grade scores, one can compute the association between the two in the time series as a useful benchmark. Roughly, the time series relationship suggests a 0.052σ and 0.037σ increase in elementary school test scores per \$1,000 increase in per-pupil spending for math and English, respectively (implying 0.044σ for both subjects combined). These estimates are similar to the most credible causal evidence.

THE OLDER DESCRIPTIVE AND CORRELATIONAL STUDIES SHOW A STRONG ASSOCIATION BETWEEN INCREASED SCHOOL SPENDING AND STUDENT ACHIEVEMENT

The research literature has never supported the idea that increased school spending doesn't improve student outcomes. Hanushek (2003) reviewed 163 studies published before 1995, relating school resources to student achievement. He found over 10 times as many positive and significant studies as expected by random chance if spending had no impact, and nearly 4 times as many positive and significant estimates as negative ones—strong evidence of a positive association between school spending and student achievement in older studies. However, he concluded that “overall resource policies have not led to discernible improvements in student performance” (p. F89). We explain below that this conclusion relies on flawed statistical reasoning that confuses statistical significance with the existence of an effect.

A single study can be statistically insignificant for two reasons: (1) lack of effects, and (2) lack of statistical power to detect an effect (e.g., due to small sample size or noisy data). To illustrate the problem with conflating the two, consider a scenario with a sample of 10,000 students, where a strong positive and highly significant relationship exists between increased school spending and student outcomes (with a t -statistic of 4, and p -value < 0.0001). Now, imagine dividing the data into 100 samples with 100 observations each. Each subsample would have a similar relationship but with standard errors approximately 10 times larger than the full data (yielding t -statistics around 0.4, and p -value ≈ 0.68). Each sample would likely show a positive, but insignificant, relationship. Using Hanushek's flawed logic (i.e., disregarding statistical power and merely counting *statistically significant* effects), one would mistakenly conclude that there is no systematic relationship overall.

Analyses of a set of studies similar to those in Hanushek (2003) that avoid this statistical error show a strong positive association between school spending and student outcomes (Hedges et al., 1994). To be crystal clear, a scientific reading of the old literature has never supported the view that there is no association between school spending and student achievement.

THE MORE RECENT CREDIBLY-CAUSAL STUDIES SHOW A STRONG CAUSAL RELATIONSHIP BETWEEN INCREASED SCHOOL SPENDING AND STUDENT ACHIEVEMENT

While suggestive of positive school spending effect on test scores, the evidence above does not represent credibly-causal relationships. To better understand the likely causal relationship, we must examine more recent empirical literature. This literature aims to uncover the relationship between increased school spending and student outcomes while carefully defining a comparison group and isolating changes in spending due to specific identifiable policies.

The studies in this literature vary in policies examined, contexts studied, and estimation approaches used. However, they consistently demonstrate a positive, credibly causal relationship between school spending and students' educational outcomes. Several studies use national data to show that increased spending due to the passage of school finance reforms is associated with improved test scores, educational attainment, and income (e.g., Jackson et al., 2016; Murray et al., 1998; Rothstein & Schanzenbach, 2022). However, many studies do not rely on the introduction of a reform. Many

studies use increases in spending due to discontinuities and deterministic functions embedded in state school funding formulas to reveal improved test scores or educational attainment in states such as Massachusetts, Michigan, and New York (Gigliotti & Sorensen, 2018; Guryan, 2001; Hyman, 2017; Papke, 2005; Roy, 2011). Using the near passage of school spending referenda, research like Baron (2022) found that spending increases improve test scores and educational attainment, while Miller (2018) associated spending increases from house price appreciation with improved outcomes. When examining spending declines, Jackson et al. (2021) linked recessionary cuts in per-pupil spending to lower test scores and reduced college attendance, and Downes and Figlio (1998) associated lower spending due to spending limits with increased dropout rates and reduced test performance. Although one can always critique any individual study, collectively, they consistently support a robust positive causal relationship between increased school spending and student outcomes on average.

There are now two recent summaries of this more recent literature: one by Jackson and Mackevicius (in press) and the other by Handel and Hanushek (2023). Despite the authors of these studies differing on interpretation (we detail this below), both of these overall summary papers found that, on average, there is a strong statistically significant relationship between increased school spending and student outcomes. Jackson and Mackevicius showed that based on 31 of these recent studies (published between 2001 and 2000), on average, a \$1,000 increase in per-pupil public operational school spending (for 4 years) increases test scores by about 0.035σ and college-going by 2.8 percentage points. Handel and Hanushek examined 43 estimates from papers published between 1999 and 2022. Translating their estimates (reported in percent changes) into dollar amounts at current spending levels, increasing school spending by \$1,000 per pupil increases test scores by 0.0466σ and college going by 3.8 percentage points. In both meta-studies, almost all of the estimates are positive, and most are statistically significant. That is, in both studies, when districts increase spending, outcomes improve the vast majority of the time. Averaging across both studies, a \$1,000 increase in per-pupil spending would raise test scores by about 0.04σ and college attendance by around 3.3 percentage points.

We now provide context for these estimates. As test scores only measure part of the benefits of better schools, these estimates likely represent a lower bound on the overall benefits of spending. Using test scores, Chetty et al. (2014) found a 0.5% wage increase for each percentile increase. A 0.04σ increase equates to about 1.6 percentile points, implying a 0.8% wage increase for a roughly \$4000 spending increase. With lifetime earnings of around \$600,000, this results in a \$4,800 benefit, indicating a benefit-cost ratio of about 1.2. However, this is a probable lower bound. To approach the true benefit, we examine educational attainment effects. Zimmerman (2014) showed that a 10 percentage point increase in college attendance leads to a 5% wage increase. Assuming this holds true, a 3.3 percentage point increase in college attendance would result in a 1.34% wage increase for a roughly \$4,000 spending increase. With a present discounted value of lifetime earnings of about \$600,000, this translates to \$9,900, yielding a cost-benefit ratio of about 2.475. Note that this calculation doesn't account for benefits such as reduced criminality or other advantages to students—likely understating the full benefits of increased school spending. In summary, recent literature suggests that each dollar spent on operational school spending generates at least a \$1.20 return on average for society, with the true benefit possibly exceeding \$2.47.

WHAT EFFECTS CAN WE EXPECT AT CURRENT SPENDING LEVELS?

While one might worry that results from studies of older reforms on educational attainment may not apply today, this concern doesn't hold for test score studies, as most examine policies implemented in the 1990s or later. Jackson and Mackevicius (in press) tested whether the marginal effect of school spending is larger when baseline spending levels are low, finding no evidence of such a relationship for either test scores or educational attainment. The effect of an additional \$1,000 is consistent across baseline spending levels, suggesting the average effect of school spending will resemble the recent literature reviews. Notably, a recent study in NY State, with well above-average baseline spending,

reported a marginal effect similar to the pooled average. In summary, both formal tests and the effects of recent reforms indicate that the benefits reported in the studies summarized above are likely applicable today.

HOW RELIABLE IS THIS ESTIMATE FOR A PARTICULAR LOCATION OR PARTICULAR POLICY?

The finding that test scores will increase by an average of 0.04σ and college-going by 3.3 pp doesn't guarantee these results in every context. It implies that approximately half the time, one might observe smaller benefits, while in other instances, the effects could be larger. Considering the benefit-cost ratios, this means that the benefits of increased spending will usually outweigh the costs. However, large social returns cannot be guaranteed in all cases or settings.

To understand how reliable increased school spending is as an investment (i.e., the likelihood a policy will improve student outcomes), one must understand the spread of actual school spending impacts. Both Jackson and Mackevicius (in press) and Handel and Hanushek (2023) tried to address this issue. However, Handel and Hanushek confused the distinction between an *estimate* and an *effect* in their analysis, resulting in a drastic overstatement of the uncertainty surrounding the potential for policies to improve outcomes in specific settings. Below, we explain why this matters.

In any given setting i , we have the true effect of a \$1,000 per-pupil increase in school spending (θ_i). We draw a random sample from the population and try to estimate this effect. However, due to sampling variability, measurement errors, and other sources of noise, we do not observe θ_i itself, but a noisy version of it $Y_i = \theta_i + \varepsilon_i$. Note that the noise (ε_i) is itself a source of variability. For example, Roy (2011) and Chaudhary (2009) both examined the effects of the same policy, but one has an estimate of 0.38, while the other has an estimate of 0.0179. A failure to properly account for the large chance differences in estimates one can observe of the same effect can lead one to overinterpret noise for signal and misinterpret the variability in estimates as variability in effects. Sadly, Handel and Hanushek (2023) did exactly this, while saying things like “the point estimates are widely different across studies” and “the outliers are dramatically different” (p. 31). As we highlighted above, estimates can be drastically different even when the underlying effect is the same—there is nothing startling or dramatic about chance variability. To understand how heterogeneous effects are likely to be, one must estimate it directly.

The basic idea behind estimating true heterogeneity is that one treats the reported standard error from each study as a measure of the noise of each study (or within-study error), and then one estimates how much variability exists that cannot be explained by noise alone. This heterogeneity parameter (τ), measures how much any two random effects are likely to differ from each other and can be used to predict how likely one is to observe the effects of any given size. Jackson and Mackevicius (in press) reported the standard deviation of true heterogeneity in test score impacts is about two-thirds the size of the mean. Accordingly, true test score effects are almost always positive—ranging between -0.004σ and 0.067σ ninety percent of the time. For educational attainment, the estimated standard deviation of heterogeneity is less than half the size of the mean so that true effects on college-going are virtually always positive—ranging between 0.5 and 5.1 pp 90% of the time.

Using the previous numbers, a college-going effect of 1.35 percentage points would pass a cost-benefit test. The pooled average and heterogeneity reported indicate this would occur at least 85 percent of the time. That is, each dollar spent on public schools would yield more than one dollar in social return at least 85% of the time. This does not account for benefits through taxes, reduced crime, or other channels, so the true likelihood is higher than this.

CONCLUSIONS

All of the evidence points to a strong causal relationship between increased school spending and student performance. The best evidence suggests that increasing school spending by \$1,000 per pupil would increase student test scores by about 0.04σ and college attendance by roughly 3.3 percentage points. The most careful examination of heterogeneity shows that while positive social returns is not guaranteed, it is exceedingly likely.

Note that the range of estimates relates to the broad range of things schools tend to do with increased money. As such, the range of estimates is indicative of what one would see in a typical state or district when they are given more money to spend. Accordingly, the range of plausible effects found indicates that by simply letting districts do what they typically do with an additional \$1,000 per pupil, we can expect positive test score effects more than 90% of the time, positive educational attainment impacts more than 99% of the time, and that the money will yield more in return than the cost at least 85% of the time.

There is a growing literature on how various school inputs, such as class size, teacher quality, and high-quality professional development, improve educational outcomes. All of these cost money. There is also evidence that incentives and good governance matter. As such, there are reasons why some areas are more efficient with their additional funds than others. However, the fact that certain kinds of spending may be more effective at improving outcomes than others is largely irrelevant to the question of whether increasing school spending is worthwhile. Whatever the budget may be, it should be spent as effectively as possible. The debate around school spending often seems to be framed as one of providing more resources versus spending resources more effectively. This dichotomy is largely rhetorical and has little economic or policy content. So long as the additional funds pass a cost-benefit test (which research suggests it typically does), one should do both. While the evidence shows that providing largely unrestricted additional money to schools does improve outcomes in general, this does not mean that increasing school spending (without regard for what it may be spent on) is the best way to improve outcomes. However, the results clearly indicate that increasing school spending in ways that have been done in the past at current spending levels will likely add social value. As such, it should be part of the policymaker's toolkit.

REFERENCES

- Baron, E. J. (2022, February). School spending and student outcomes: Evidence from revenue limit elections in Wisconsin. *American Economic Journal: Economic Policy*, 14(1), 1–39. <https://doi.org/10.1257/pol.20200226>
- Chaudhary, L. (2009). Education inputs, student performance and school finance reform in Michigan. *Economics of Education Review*, 28(1), 90–98. <https://doi.org/10.1016/j.econedurev.2007.11.004>
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014, September). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9), 2593–2632. <https://doi.org/10.1257/aer.104.9.2593>
- Doty, E., Kane, T. J., Patterson, T., & Staiger, D. O. (2022). *What do changes in state test scores imply for later life outcomes?* [Working paper no. 30701]. National Bureau of Economic Research. <https://doi.org/10.3386/w30701>
- Downes, T., & Figlio, D. (1998). *School finance reforms, tax limits, and student performance: Do reforms level-up or dumb down?* Department of Economics, Tufts University.
- Gigliotti, P., & Sorensen, L. C. (2018). Educational resources and student achievement: Evidence from the Save Harmless Provision in New York State. *Economics of Education Review*, 66, 167–182. <https://doi.org/10.1016/j.econedurev.2018.08.004>
- Guryan, J. (2001). *Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts.* [Working paper no. 8269]. National Bureau of Economic Research. <https://doi.org/10.3386/w8269>
- Handel, D. V., & Hanushek, E. A. (2023). *US school finance: Resources and outcomes* [Working paper no. 30769] (Rev. ed.). National Bureau of Economic Research. Originally published 2022. <https://doi.org/10.3386/w30769>
- Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113(485), F64–98. <https://doi.org/10.1111/1468-0297.00099>
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994, May). Money does matter somewhere: A reply to Hanushek. *Educational Researcher*, 23(4), 9–10. <https://doi.org/10.3102/0013189x023004009>
- Hyman, J. (2017). Does money matter in the long run? Effects of school spending on educational attainment. *American Economic Journal: Economic Policy*, 9(4), 256–280. <https://doi.org/10.1257/pol.20150249>

- Jackson, C. K., Johnson, R. C., & Persico, C. (2016, February 1). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, 131(1), 157–218. <https://doi.org/10.1093/qje/qjv036>
- Jackson, C. K., & Mackevicius, C. L. (in press). What impacts can we expect from school spending policy? Evidence from evaluations in the U.S. *American Economic Journal: Applied Economics*. <https://doi.org/10.1257/app.20220279>
- Jackson, C. K., Wigger, C., & Xiong, H. (2021). Do school spending cuts matter? Evidence from the Great Recession. *American Economic Journal: Economic Policy*, 13(2), 304–335. <https://doi.org/10.1257/pol.20180674>
- Miller, C. L. (2018). The Effect of education spending on student achievement: Evidence from property values and school finance rules. *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, 111, 1–121. <https://www.jstor.org/stable/26939356>
- Murray, S. E., Evans, W. N., & Schwab, R. M. (1998). Education-finance reform and the distribution of education resources. *The American Economic Review*, 88(4), 789–812. <https://www.jstor.org/stable/117006>
- National Center for Education Statistics. (2017, August). *Focus on NAEP: Grade 12 participation and engagement in NAEP*. U.S. Department of Education, Institute of Education Sciences. https://www.nationsreportcard.gov/focus_on_naep/files/g12_companion.pdf
- Papke, L. E. (2005). The effects of spending on test pass rates: Evidence from Michigan. *Journal of Public Economics*, 89(5–6), 821–839. <https://doi.org/10.1016/j.jpubeco.2004.05.008>
- Rothstein, J., & Schanzenbach, D. W. (2022). Does money still matter? Attainment and earnings effects of post-1990 school finance reforms. *Journal of Labor Economics*, 40(S1), S141–178. <https://doi.org/10.1086/717934>
- Roy, J. (2011). Impact of school finance reform on resource equalization and academic performance: Evidence from Michigan. *Education Finance and Policy*, 6(2), 137–167. https://doi.org/10.1162/EDFP_a_00030
- Zimmerman, S. D. (2014, October). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4), 711–754. <https://doi.org/10.1086/676661>

How to cite this article: Jackson, C. K. & Persico, C. (2023). Point column on school spending: Money matters. *Journal of Policy Analysis and Management*, 42, 1118–1124. <https://doi.org/10.1002/pam.22520>

AUTHOR BIOGRAPHIES

C. Kirabo Jackson is Abraham Harris Professor of Education and Social Policy in the Department of Human Development and Social Policy at Northwestern University, 2120 Campus Drive, Room 204, Annenberg Hall, Evanston, IL 60208 (email: Kirabo.jackson@gmail.com).

Claudia Persico is an Associate Professor in the Department of Public Administration and Policy at American University, School of Public Affairs, 4400 Massachusetts Avenue NW, Washington, DC 20016–8070 (email: cpersico@american.edu).