
Reviewed by:
Clive Belfield
Queens College, City University of New York

April 2024

National Education Policy Center
School of Education
University of Colorado Boulder
nepc.colorado.edu
Acknowledgements

NEPC Staff

Faith Boninger
Publications Manager

Francesca López
Academic Editor

Elaine Duggan
Production Design

Alex Molnar
NEPC Director

Kevin Welner
NEPC Director


Funding: This review was made possible in part by funding from the Great Lakes Center for Educational Research and Practice.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

This publication is provided free of cost to NEPC’s readers, who may make non-commercial use of it as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at nepc@colorado.edu.

The National Education Policy Center (NEPC), a university research center housed at the University of Colorado Boulder School of Education, sponsors research, produces policy briefs, and publishes expert third-party reviews of think tank reports. NEPC publications are written in accessible language and are intended for a broad audience that includes academic experts, policymakers, the media, and the general public. Our mission is to provide high-quality information in support of democratic deliberation about education policy. We are guided by the belief that the democratic governance of public education is strengthened when policies are based on sound evidence and support a multiracial society that is inclusive, kind, and just. Visit us at: http://nepc.colorado.edu
Summary

A growing and persuasive body of research has established the causal relationship between school funding and academic outcomes. Money matters. But a new report from the Reason Foundation uses some superficial comparisons to conclude that there “isn’t a consistent relationship between funding growth and outcomes across states.” This report uses national education data to describe trends in spending, teacher pay, and academic outcomes from 2002-03 up to the start of the pandemic. It shows trends of increasing revenue, declining teacher pay, and growing employment in non-instructional support positions. For academic outcomes, the report shows that NAEP test scores in math and reading had flatlined over the two decades before the pandemic. The report contends that, in order to move forward in the aftermath of the pandemic, it is important to understand key pre-pandemic trends, but it offers no arguments as to how those trends provide appropriate context for post-pandemic educational policy. In any case, the report is insubstantial and inconclusive and offers no useful new analyses; in fact, it reprises evidence that is already widely available. It does not account for any economic, social, or demographic trends within states or nationally. Its analysis is very weak and in some cases nonexistent, resulting in conclusions that are vacuous. Policymakers would do well to ignore it.

Reviewed by:
Clive Belfield
Queens College, City University of New York

April 2024

I. Introduction

The COVID-19 pandemic was a shock to the U.S. K-12 education system. It will most likely require significant research inquiry over many years to fully understand its extent and to quantify its impact over time. One possibility is that the pandemic was a one-time shock to the system and that its consequences will fade over time.¹ The other possibility is that the pandemic permanently altered the education system; and policymakers and parents will have to adjust to a “new normal.” Education research plays an important role in describing what the U.S. education system was like before the pandemic and what to expect as the pandemic has waned.

The Reason report, Public Education at a Crossroads: A Comprehensive Look at K-12 Resources and Outcomes for All 50 States, describes trends in the U.S. education system in the two decades before the pandemic.² The report contends that, in order to move forward in the aftermath of the pandemic, it is important to understand key trends before it started. The report looks at a range of metrics: K-12 revenues, teacher pay, staff deployment, and education outcomes. Evidence is reported at the national level and for each state in 2002-03 and 2019-20. Thus, the report attempts to offer an expansive, historical picture of the U.S. education system before the pandemic hit in March 2020.

II. Findings and Conclusions of the Report

The report presents five main findings for the U.S. education system from 2002-03 up to just prior to the pre-pandemic (2019-20). These findings are reported at the national level
and for each state.

First, inflation-adjusted revenue per student was 25% higher nationally in 2020 than in 2002. Revenue grew across almost all states and was more than 50% higher in five states. Much of this increase was allocated to payments for support services and employee benefits. Revenue per student had fallen over the two decades prior (and by less than 1%) in only one state (North Carolina).

Second, U.S. public school enrollments grew by 7% (from 47.7 million to 50.8 million students). However, this growth was very uneven across states: Enrollment grew by more than 25% in five states but fell by more than 9% in five others.

Third, schools hired many more non-teaching staff, with an increase of 20% over the two decades. By contrast, the number of teaching staff grew by only 7%. Hence, teaching staff numbers were growing as fast as enrollment numbers; but non-teaching staff numbers were growing much faster.

Fourth, teacher salaries stagnated in real terms: In 2002, the average salary was $64,500; by 2020, it was $64,100. Again, salary growth varied by state: Average pay was more than 9% higher in five states whereas it was 12% lower in five others.

Finally, the report presents evidence on National Assessment of Educational Progress (NAEP) test scores. Focusing on fourth and eighth grade math and reading, the report finds that there was almost no improvement in these subjects nationally. For example, eighth grade reading scores are identical in 2003 and 2019 (at 263); and eighth grade math scores increased by only four points (from 278 to 282). (Looking only at students on Free or Reduced Price lunch, there was some modest growth in test scores.) For these outcomes, the variation across states was modest.

The report is mostly descriptive: It draws few conclusions from the evidence. The main conclusion is that there “isn’t a consistent relationship between funding growth and outcomes across states.” Specifically, states where revenue grew fastest between 2003 and 2019 did not have significantly higher NAEP scores. Also, the report is skeptical that any increases in funding will translate into increases in teacher pay; by inference, additional funding will likely be allocated either to non-instructional support staff or to employee benefits and pensions.

III. The Report’s Rationale for Its Findings and Conclusions

The report relies on national datasets from the Census and the National Center for Educational Statistics. These sources are valid and provide credible, harmonized data over time and across states. To generate its findings, the report simply tabulates information on revenues, pay, and academic outcomes in 2002-03 and 2019-20.

The report’s main conclusion—no consistent relationship between funding growth and academic outcomes—is based on simple ranking. The rationale is as follows: If funding growth
leads to better academic outcomes, then states that rank highest in terms of funding should also rank highest in terms of academic outcomes. From a general review of these respective ranks, the report does not identify any clear positive correlation and dismisses the argument that funding boosts outcomes.

IV. The Report’s Use of Research Literature

The report’s use of research literature is extremely limited and very weak. The report does not refer to a single peer-reviewed journal article on educational funding, teacher labor markets, efficiency, or pensions. Instead, the report relies solely on sources from a few websites (mostly postings from the Hoover Institution and pieces from the advocacy journal Education Next). The validity of these references has not been independently verified and there is no consideration as to whether these references represent a consensus or are generally accepted by the research community. Even with its exclusive reliance on material from websites, the report makes no effort to summarize the evidence. Hence, it provides a very cursory context for understanding how schools obtain revenue, how teachers are paid, or how pension plans are structured.

The report provides no context for understanding why states might vary across each of these topic areas. There are many reasons why trends might vary per state. For example, unions can play an important role in pay-setting and benefits entitlements but the strength of these unions varies significantly by state. As another example, family demographics affect how students learn; these demographics vary by state on many dimensions, including race, incomes, and immigrant status. The report does not include literature on variations such as these.

The report does not consider research on NAEP scores. It does not discuss how these scores should be interpreted (the NAEP website provides considerable information on methodology and on how to interpret the results). The report also does not mention research studies that have examined what determines NAEP scores by state. For example, the American Institutes for Research (AIR) has 16 separate research studies examining variation in NAEP.

Notably, the report ignores the vast literature on the link between education funding and outcomes. This literature stretches back decades over 30 years, with studies on school finance reforms, bond issuance, and other techniques to identify a causal relationship between money and outcomes. None of this literature is considered in the report.

Finally, the report includes zero research evidence on the impacts of the pandemic on education systems. This research evidence is now substantial. For example, there is high-quality evidence on how online learning affected students’ academic skills both nationally and across states. There is also evidence on K-12 funding and pandemic learning loss, as well as on how family incomes and teacher pay fell. The report includes no evidence on the pandemic’s impact on any of these issues.
V. Review of the Report’s Methods

The report’s method is to report data tabulations, provide rankings, and to correlate state rankings against funding outcomes. This method is inadequate.

First, the tabulations are unadjusted for differences over time both nationally and across states. To compare two time periods, it is necessary to adjust for external factors that may have changed. For example, if the average teacher had less experience in 2003 than in 2020, we would expect average teacher pay to have increased over time. Or, if the prevalence of mental health issues among students was rising, we would expect non-instructional staffing to increase. The report ignores all external factors and presents unadjusted tabulations for 2003 and 2020. It also applies a single uniform price index; this is invalid because inflation of educational services is not constant across all states. In effect, the report is assuming that—across the U.S. and within each state—the economy, family structures, and educational needs were unchanged over two decades.

Second, the report does not present trends in sufficient detail. Instead, it offers two snapshots: 2002-03 and 2019-20. Years in between these two snapshots are not examined. This approach is misleading. For example, the report documents a 6.6% increase in student enrollments in public school over the period. In actuality, enrollment not only did not steadily increase each year (enrollment was almost identical in 2010 and 2011, for example) but has declined since 2016.

Finally, the report’s method is puzzling in light of its stated logic, which is to examine pre-pandemic trends in order to “provide policymakers with a critical anchor for navigating post-pandemic decisions.” However, the report mostly ignores the educational consequences of the pandemic itself. Readers may learn about the U.S. education system before the pandemic but will not learn what has happened since. The report does not discuss in any detail (and it provides no evidence on) how the pandemic has affected K-12 education since 2020. For example, it shows that teacher pay flatlined from 2003 up to 2020 but it provides no evidence on whether teacher pay increased (or decreased) during the pandemic. Perhaps pre-pandemic trends were offset by the pandemic; perhaps they were exacerbated. The report provides very little information either way. Historical analysis is of limited use when—as in this report—it pays no attention to the most eventful years.

VI. Review of the Validity of the Findings and Conclusions

Most of the findings of the report are based on tabulations from other sources. In that respect, these tabulations are reasonable (although it is unclear why readers should consult the report rather than access the tabulations directly from their sources).

However, the comparison over time and between states is not well-founded. As the report does not adjust for any external factors it is difficult to determine if any trends are a result of new educational policy decisions or demographic changes. This problem is exacerbated by the spurious ranking method used. Leaving aside measurement issues, ranking by revenue
can be misleading. For example, California is ranked ninth in funding growth as a percentage of baseline revenues; New Jersey is ranked 21st. But the absolute dollar increase in funding was 13% higher in New Jersey than California.\textsuperscript{14} Similarly, ranking by NAEP scores can be misleading; most states have very similar scores that are statistically indistinguishable when sampling variance is accounted for.

The report’s conclusion on funding and outcomes has no value. There are many studies on funding and outcomes. At a high level of generality, it is accepted that increases in funding are likely to improve educational outcomes for students. But there are caveats: Specific results and findings vary depending on how funding is counted, what outcomes are measured, and which statistical tests are applied to establish causality.\textsuperscript{15} Nevertheless, significant methodological progress has been made: Rigorous standards have been established as to what constitutes a valid test of the relationship between funding and outcomes.

The report’s analysis of funding and outcomes does not meet even minimal research standards. In fact, the report does not perform any formal research on the relationship between funding and outcomes. Rather, the report ranks each state in terms of changes in per-student revenue and NAEP scores. It then shows that the most highly funded states are not the same as the highest scoring states and so concludes that there is no consistent relationship between these two entities.

This analysis is deficient in many respects but three deficiencies are striking. First, it includes no formal tests (not even of rank correlation coefficients or pairwise correlations). By extension, no sensitivity testing is employed either. Second, it ignores sampling error. After accounting for sampling variation, NAEP scores in most states are statistically equivalent and so states cannot be individually ranked.\textsuperscript{16} Finally, and most egregiously, the report does not account for any possible confounding factors. Most obviously, states with high proportions of at-risk students may need to spend more in order to compensate for the extra learning needs of these students. These three deficiencies are such that they render the report’s conclusion vacuous.

VII. Usefulness of the Report for Guidance of Policy and Practice

The report includes some basic evidence that may be useful for policymakers who must speculate about a post-pandemic future. This evidence relates to resources for personnel. One striking finding is that teacher pay effectively flatlined between 2003 and 2020. Thus, policymakers might temper their expectations regarding improvements in educational outcomes: It seems unreasonable to expect significant productivity gains when frontline workers (teachers) have not had a real pay increase in two decades. Also, payments for pensions are increasing. Policymakers might be concerned that future budgets will be constrained by promises to pay pension obligations. Finally, many more non-instructional staff are being hired by schools. Here, policymakers might consider why these staff are needed and why social and behavioral services are being prioritized over resources for instruction. Unfor-
fortunately, the report provides only a very cursory discussion on these speculations (or any others that might reasonably be inferred from its tabulations).

Some of the tabulations in the report are unlikely to offer policymakers much help. The report notes a trend in spending more on support services staff (non-instructional). However, it is unknown whether this is a beneficial or detrimental trend. If states had been spending too little (much) on support services staff, then the extra spending is efficient (inefficient). Without knowing what the optimal level of spending should be, it is not much help to know what the actual level of spending is. (More bluntly, as Professor Loeb remarked in a column for Brookings: “half the people working in schools aren’t classroom teachers: So what?”)

Most unfortunately, the report includes analysis that has no value for policymakers. One of its key findings—that funding growth does not yield gains in academic outcomes—is not based on a credible research method and uses data that have not been adequately transformed. As such, policymakers would be better served by disregarding this conclusion and by relying on more rigorous, peer-reviewed research.

Moreover, the state-level analysis provides limited information for policy or practice. State policymakers may discover where they are ranked in each of the topic areas but many of the ranks (especially for academic outcomes) do not reflect genuine differences between states. Also, there is no explanation of why a state is highly-ranked or what policymakers should do to move up the rankings.

Finally, the report offers no guidance on the pandemic and its consequences for decision-making. This omission is puzzling given that this was ostensibly one of the motivations for the report. The report offers an anecdotal depiction of the impact of the pandemic and an insubstantial treatment of the education system post-pandemic. Thus, it is difficult for readers of this report to use pre-pandemic trends as a guide for decisions to improve education systems in a post-pandemic future.

Given the report’s poor methodology, very weak (in some cases nonexistent) analysis, and vacuous conclusions, policymakers would do well to ignore it.
Notes and References

1 Even as current cohorts of students may be unaffected, the pandemic may take a long time to wane for students for whom instruction was directly affected in 2020 and 2021. For example, these students may suffer long-term shortfalls in earnings. See Hanushek, E.A. (2023, January 4). *The economic cost of the pandemic: State by state.* Hoover Institution. Retrieved January 15, 2023, from https://www.hoover.org/sites/default/files/research/docs/HanushekEconomicCostweb.pdf


4 Unionization rates are available from Schools and Staffing Survey, National Center for Education Statistics. Retrieved March 11, 2024, from https://nces.ed.gov/surveys/sass/tables/sass0708_043_t1s.asp


7 For an early study, see:


For an example of more recent literature, see:


http://nepc.colorado.edu/review/crossroads


There are many possible effects of the pandemic. Teacher pay may have risen because of government investment via the CARES Act, for example. But teaching as a profession may have become less desirable (e.g. because of social distancing requirements or online learning).


Two recent studies are listed below; a full catalog would be voluminous:


The standard error for NAEP scores is typically 1-2 points. Based on the growth in 8th grade NAEP reading scores, Mississippi is ranked 15th and Arkansas is ranked 21st, despite there being no statistically difference in their scores. Indeed, few states have growth rates outside of the reported standard errors. It is because scores are measured with error that the NAEP website groups states into three categories relative to the national average of “performed higher,” “no difference,” and “performed lower.”

