State Education Trends
Academic Performance and Spending over the Past 40 Years
By Andrew J. Coulson

EXECUTIVE SUMMARY

Long-term trends in academic performance and spending are valuable tools for evaluating past education policies and informing current ones. But such data have been scarce at the state level, where the most important education policy decisions are made. State spending data exist reaching back to the 1960s, but the figures have been scattered across many different publications. State-level academic performance data are either nonexistent prior to 1990 or, as in the case of the SAT, are unrepresentative of statewide student populations. Using a time-series regression approach described in a separate publication, this paper adjusts state SAT score averages for factors such as participation rate and student demographics, which are known to affect outcomes, then validates the results against recent state-level National Assessment of Educational Progress (NAEP) test scores. This produces continuous, state-representative estimated SAT score trends reaching back to 1972. The present paper charts these trends against both inflation-adjusted per pupil spending and the raw, unadjusted SAT results, providing an unprecedented perspective on American education inputs and outcomes over the past 40 years.
The performance of 17-year-olds has been essentially stagnant across all subjects despite a near tripling of the inflation-adjusted cost of putting a child through the K–12 system.

—Horace Mann, 1837, “The Means and Objects of Common-School Education”

Parents often share the view expressed by Horace Mann, godfather of American public schooling: they want their children to have better educational options than they had. They want the best. Aware of this fact, state policymakers constantly seek to improve public school outcomes (or, for the politically jaded, they at least wish to appear to be doing so). But how well are they succeeding?

At the national level, the results do not look good. The performance of 17-year-olds has been essentially stagnant across all subjects since the federal government began collecting trend data around 1970, despite a near tripling of the inflation-adjusted cost of putting a child through the K–12 system.

And yet, nationwide patterns are not always seen as relevant to the outcomes of any particular state. Public opinion polls regularly show that Americans simultaneously think the nation’s schools are in dire straits while believing their own schools to be performing better. We can’t all be right. But who, in particular, is wrong?

Until now, there has been no way to answer that question with respect to long-term trends in state educational performance. State-level test score trends are either nonexistent prior to 1990 or, as in the case of college entrance tests like the SAT, are unrepresentative of statewide performance.

Figure 1
Trends in American Public Schooling Since 1970


Note: “Total cost” is the full amount spent on the K-12 education of a student graduating in the given year, adjusted for inflation. In 1970, the amount was $56,903; in 2010, the amount was $164,426.
student populations. The size and composition of a state’s SAT-taking population varies over time, affecting its average score.

Fortunately, it is possible to adjust state-average SAT scores to compensate for varying participation rates and student demographics, as was demonstrated in a 1993 paper for the *Economics of Education Review* by Mark Dynarski and Philip Gleason.² In a recent time-series regression study, I extended and improved on the Dynarski and Gleason model to allow adjusted SAT scores to be calculated for all 50 states between 1972 and 2012.³ These adjusted SAT scores were validated against the available state-level NAEP data with good results, suggesting that they offer a plausible estimate of overall state performance on the SAT.⁴

Of course, this is only a useful endeavor to the extent that the SAT measures things that people value, and that it measures them fairly across different student subgroups. These questions are taken up in the section titled “Is the SAT a Useful Metric?”

The results themselves are charted in the section titled “State Education Trends.” The first chart shows the percent change over time in adjusted SAT scores and in inflation-adjusted public school spending per pupil. This offers an indication of the returns states have enjoyed on their educational investments. The second chart compares the percent change over time in the adjusted SAT scores and the raw unadjusted SAT scores. The results of that comparison indicate how unwise it is to rely on unadjusted SAT scores to gauge changes in states’ educational outcomes over time.

**IS THE SAT A USEFUL METRIC?**

The first point worth making is that SAT scores are obviously not a comprehensive metric of educational outcomes. Numerous factors unmeasured by the SAT (e.g., character, grit, artistic skills, subject area knowledge) are of interest to families and are important to life quality and success. The question addressed here is only whether or not the things that the SAT does measure are also of general interest.

Though the SAT is known chiefly as a college entrance exam, it measures reading comprehension and mathematical skills that are intrinsically useful and that schools take great pains to teach. Even the SAT’s more obscure vocabulary questions are revealing, because a person’s vocabulary and their overall comprehension are directly tied to the amount of reading they’ve done and the richness of the texts they’ve read.⁵ Since developing avid readers is a universal educational goal, this is useful information.

To the extent that the SAT also helps to predict success in college, it provides additional information on educational outcomes that families value. There is, however, a common criticism that the SAT only explains a quarter or less of the variation in students’ college grade-point averages (GPAs). What this criticism fails to acknowledge is that the SAT/GPA studies typically measure that relationship within colleges. They compare students’ entering SAT scores to their first- or second-year GPAs, within a given institution. But, as Temple University mathematician John Allen Paulos observes,

Colleges usually accept students from a fairly narrow swath of the SAT spectrum. The SAT scores of students at elite schools, say, are considerably higher, on average, than those of students at community colleges, yet both sets of students probably have similar college grade distributions at their respective institutions.

If both sets of students were admitted to elite schools or both sets attended community colleges, there would be a considerably stronger correlation between SATs and college grades at these schools.

Those schools that attract students with a wide range of SAT scores generally have higher correlations between the scores and first-year grades.⁶

In other words, much of the SAT’s ability to predict college success is manifested in the different tiers of colleges to which students with different SAT scores have access. To look only at the relationship between SATs and GPAs within

---

*While SAT scores are not a comprehensive metric of educational outcomes, the SAT measures reading comprehension and mathematical skills that are intrinsically useful.*
The variation in the SAT’s ‘predictive validity’ across racial and ethnic subgroups is not large.

particular colleges misses this important variation and thus understates the strength of the relationship between SAT scores and proficiency at college-level work.

Nevertheless, even within the top 1 percent of SAT-scorers, those with the very highest scores tend to achieve more than those with relatively lower scores. A team of researchers from Vanderbilt University has documented this pattern for a variety of life outcomes including eventual income, publication in peer-reviewed journals, holding advanced degrees, and holding patents.7

While it has been suggested that the predictive power of SAT scores vanishes after controlling for socioeconomic status, grades, and subject-area test scores (such as the SAT II), that is a tautological observation. Many of the same reading, vocabulary, and mathematics skills measured by the SAT are also measured by grades and subject-area tests, so controlling for them using those other measures necessarily leaves little for the SAT to explain. It is true that controlling for socioeconomic status does reduce the SAT’s ability to predict college GPA, but the effect is small.8

It is also sometimes alleged that the SAT is biased against nonwhite students. This claim is based on the large and persistent gaps between the scores of some minority subgroups and the scores of whites. However, test bias is not the only possible cause for these subgroup test score differences—differential levels of academic preparedness across subgroups could also explain the observed results.

As it happens, the variation in the SAT’s “predictive validity” across racial and ethnic subgroups is not large. The correlation between SAT scores and within-college second-year GPAs ranges from .49 for African Americans, to .54 for Asians and Pacific Islanders, .55 for Hispanics, and .56 for whites.9 As noted above, the use of within-college SAT/GPA correlations discards information about the link between the SAT score and the tier of college to which students are able to gain admission, and so these correlation figures should be considered conservative lower bounds on the actual link between the SAT and performance on college-level material.

Interestingly, the benefits of gaining admission to a more selective college via a higher SAT score may be larger for African Americans than for other subgroups. A 2012 study comparing the eventual earnings of graduates of more- and less-selective colleges in Texas finds an overall benefit to attending a more-selective college, but notes that “historically under-represented minorities experience the highest returns in the upper tails of the earnings distribution.”10

A somewhat similar pattern was reported by Stacy Dale and Alan Krueger in the same year. Even in their most heavily controlled model, they find that low-income and minority students who attended the most selective colleges enjoyed large subsequent earnings benefits.11

STATE EDUCATION TRENDS—THE FINDINGS

The state-by-state results of this investigation are reported in the subsections that follow, but the overall picture can be summarized in a single value: 0.075. That is the correlation between the spending and academic performance changes of the past 40 years, for all 50 states. Correlations are measured on a scale from 0 to 1, where 0 represents absolutely no correlation between two data series and 1 represents a perfect correlation. Anything below 0.3 or 0.4 is considered a weak correlation. The 0.075 figure reported here suggests that there is essentially no link between state education spending (which has exploded) and the performance of students at the end of high school (which has generally stagnated or declined).
Figure 2
Alabama

Alabama Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Alabama SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 5
Arkansas Education Trends

Arkansas SAT Trends

Figure 6
California Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972


Dollars per Pupil (Inflation Adjusted)

Percent Change Relative to 1972


California SAT Trends

- Raw
- Adjusted for Participation and Demographics

Percent Change Relative to 1972


Figure 7
Colorado

Colorado Education Trends

- - Dollars per Pupil (Inflation Adjusted)
- - SAT Score Adjusted for Participation and Demographics

Colorado SAT Trends

- - Raw
- - Adjusted for Participation and Demographics

Delaware Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Delaware SAT Trends

- Raw
- Adjusted for Participation and Demographics

The raw College Board SAT scores used in this study cover all graduating seniors who took the SAT at some point during their high school career. The share of graduating seniors for whom this is true is especially high in Delaware, since the state introduced universal SAT testing for high school juniors in 2011. The resulting higher participation rate drives down Delaware’s raw SAT scores because more lower-performing students take the test than is the case in most other states. However, the SAT score adjustment formula used in this study takes into account SAT participation rate.

Despite that fact, Delaware’s adjusted SAT scores also fell after the enactment of its universal SAT policy in 2011, albeit to a lesser extent than the raw scores. The reason that the decline persists even after adjusting for participation rate is two-fold. First, the adjustment formula looks at SAT participation as a function of the 18-year-old population; and, second, Delaware has a relatively large number of 18-year-olds who do not remain in school through to graduation. So while Delaware has a high share of graduating seniors with SAT scores, many of its 18-year-olds never graduate, so the state’s test-takers actually represent a smaller share of the eligible population than it might at first seem.

The importance of this factor can be seen in the rather different results for the state of Maine, which also introduced a universal SAT program in 2006. Like Delaware’s, Maine’s raw SAT scores tumbled after the program was introduced, but its adjusted SAT scores did not. Indeed, they have risen in the years since the program was introduced. The difference is that far more 18-year-olds in Maine remain in school through to graduation, and so Maine’s SAT participation rate (83 percent) is substantially higher than Delaware’s (73 percent).

Maine does disproportionately well on the SAT considering the high percentage of its young people who take it, whereas Delaware does more poorly than would be expected for its rate of SAT participation.
Figure 10
Florida

Florida Education Trends

- - - Dollars per Pupil (Inflation Adjusted)
- - - SAT Score Adjusted for Participation and Demographics

Florida SAT Trends

- - - Raw
- - - Adjusted for Participation and Demographics

Georgia Education Trends

Dollars per Pupil (Inflation Adjusted)
SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972


Figure 12
Hawaii

Hawaii Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Hawaii SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 13
Idaho

Idaho Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Idaho SAT Trends

- Raw
- Adjusted for Participation and Demographics

Indiana Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Indiana SAT Trends

- Raw
- Adjusted for Participation and Demographics

Kansas Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Kansas SAT Trends

- Raw
- Adjusted for Participation and Demographics

Kentucky Education Trends

- - - Dollars per Pupil (Inflation Adjusted)
- - - SAT Score Adjusted for Participation and Demographics

Kentucky SAT Trends

- - - Raw
- - - Adjusted for Participation and Demographics

Louisiana Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Louisiana SAT Trends

- Raw
- Adjusted for Participation and Demographics

Though Maine suffered a substantial drop in raw SAT scores after 2005, this can be attributed entirely to the state’s introduction in 2006 of a mandatory universal SAT program. A higher participation rate generally means that more lower-achieving students are taking the test, which drives down the average. Factoring in the state’s participation-rate controls for this source of variation in average state scores provides a better indication of underlying trends in the performance of the state’s student population as a whole. That fact is illustrated by the divergence between the raw and adjusted SAT scores for Maine starting in 2006.

Maine is not the only state to introduce universal SAT taking, and the experience described earlier in the section on Delaware provides a revealing contrast.
Figure 21
Maryland Education Trends

Maryland SAT Trends

Figure 22
Massachusetts

Massachusetts Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Massachusetts SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 23
Michigan

Michigan Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Michigan SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 26
Missouri

Missouri Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Missouri SAT Trends

- Raw
- Adjusted for Participation and Demographics

Montana Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972


Montana SAT Trends

Raw Adjusted for Participation and Demographics

Percent Change Relative to 1972


Figure 28
Nebraska

Nebraska Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Nebraska SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 31
New Jersey

New Jersey Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

New Jersey SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 32
New Mexico

New Mexico Education Trends

- - - Dollars per Pupil (Inflation Adjusted)
- - SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972

New Mexico SAT Trends

- - Raw
- Adjusted for Participation and Demographics

Percent Change Relative to 1972

New York Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Figure 34
North Carolina

North Carolina Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972

North Carolina SAT Trends

- Raw
- Adjusted for Participation and Demographics

Percent Change Relative to 1972

Oregon Education Trends

Oregon SAT Trends

Figure 39
Pennsylvania

Pennsylvania Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Pennsylvania SAT Trends

- Raw
- Adjusted for Participation and Demographics

Rhode Island Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Rhode Island SAT Trends

- Raw
- Adjusted for Participation and Demographics

Tennessee Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Tennessee SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 45
Utah

Utah Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Utah SAT Trends

- Raw
- Adjusted for Participation and Demographics

Figure 48
Washington

Washington Education Trends

Dollars per Pupil (Inflation Adjusted)
SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972

Dollars per Pupil (Inflation Adjusted)
-10 -8 -6 -4 -2 0 2

Washington SAT Trends

Raw
Adjusted for Participation and Demographics

Percent Change Relative to 1972

Raw
Adjusted for Participation and Demographics
-10 -8 -6 -4 -2 0 2

Figure 49
West Virginia Education Trends

West Virginia SAT Trends

Figure 50
Wisconsin

**Wisconsin Education Trends**

- **Dollars per Pupil (Inflation Adjusted)**
- **SAT Score Adjusted for Participation and Demographics**

**Wisconsin SAT Trends**

- **Raw**
- **Adjusted for Participation and Demographics**

Figure 51
Wyoming Education Trends

- Dollars per Pupil (Inflation Adjusted)
- SAT Score Adjusted for Participation and Demographics

Percent Change Relative to 1972

Wyoming SAT Trends

- Raw
- Adjusted for Participation and Demographics

CONCLUSION

Academic performance and preparation for college success are widely shared goals, and so it is useful for the public and policymakers to know how they have varied over time at the state level. The present paper estimates these trends by adjusting state average SAT scores for variation in student participation rates and demographic factors known to be associated with those scores.

In general, the findings are not encouraging. Adjusted state SAT scores have declined by an average of 3 percent. This echoes the picture of stagnating achievement among American 17-year-olds painted by the Long Term Trends portion of the National Assessment of Educational Progress, a series of tests administered to a nationally representative sample of students since 1970. That disappointing record comes despite a more-than-doubling in inflation-adjusted per pupil public-school spending over the same period (the average state spending increase was 120 percent). Consistent with those patterns, there has been essentially no correlation between what states have spent on education and their measured academic outcomes. In other words, America’s educational productivity appears to have collapsed, at least as measured by the NAEP and the SAT.

That is remarkably unusual. In virtually every other field, productivity has risen over this period thanks to the adoption of countless technological advances—advances that, in many cases, would seem ideally suited to facilitating learning. And yet, surrounded by this torrent of progress, education has remained anchored to the riverbed, watching the rest of the world rush past it.

Not only have dramatic spending increases been unaccompanied by improvements in performance, the same is true of the occasional spending declines experienced by some states. At one time or another over the past four decades, Alaska, California, Florida, and New York all experienced multi-year periods over which real spending fell substantially (20 percent or more of their 1972 expenditure levels). And yet, none of these states experienced noticeable declines in adjusted SAT scores—either contemporaneous-ly or lagged by a few years. Indeed, their score trends seem entirely disconnected from their rising and falling levels of spending.

Two generations seems a long time for a field to stand outside of history, particularly when those generations have witnessed so many reforms aimed at improving education. Perhaps it’s time to ask if there are inherent features in our approach to schooling that prevent it from enjoying the progress typical in other fields.

NOTES


4. The raw SAT data used in this analysis were provided by The College Board, and are copyright © 1972–2012 by The College Board. http://www.collegeBoard.com.


7. Kimberley Ferriman Robertson, Stijn Smeets,


11. Stacy Dale and Alan Krueger, “Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data,” October 19, 2012, http://www.utexas.edu/law/journals/tri/sources/volume%2092/issue%204/Willig/DSB/Willig.fno5401.SB.pdf. It is worth noting economist Caroline Hoxby’s concern with the earlier version of the Dale and Krueger paper, which is that the predictive power of their model relies on data for students who are accepted to highly selective colleges but choose to attend relatively less selective ones. These students, Hoxby notes, are outliers, since most students attend the most selective college to which they are accepted. These outlier students may differ in unmeasured ways from the majority, thus calling into question the generalizability of the results. See Caroline Hoxby, “The Changing Selectivity of American Colleges,” *Journal of Economic Perspectives* 23, no. 4 (Fall 2009): 95–118.
RELATED PUBLICATIONS

Drawing Meaningful Trends from the SAT by Andrew J. Coulson, Cato Institute Working Paper no. 16 (March 10, 2014)


The Impact of Federal Involvement in America’s Classrooms by Andrew J. Coulson, Testimony before the Committee on Education and the Workforce, United States House of Representatives (February 10, 2011)


Private School Chains in Chile: Do Better Schools Scale Up? by Gregory Elacqua, Dante Contreras, Felipe Salazar, and Humberto Santos; Cato Institute Policy Analysis no. 682 (August 16, 2011)


Behind the Curtain: Assessing the Case for National Curriculum Standards by Neal McCluskey, Cato Institute Policy Analysis no. 661 (February 17, 2010)

The Effects of Teachers Unions on American Education by Andrew J. Coulson, *Cato Journal* 30, no. 1 (Winter 2010)

They Spend WHAT? The Real Cost of Public Schools by Adam B. Schaeffer, Cato Institute Policy Analysis no. 662 (March 10, 2010)

Dismal Science: The Shortcomings of U.S. School Choice Research and How to Address Them by John Merrifield, Cato Institute Policy Analysis no. 616 (April 16, 2008)

How Much Ivory Does This Tower Need? What We Spend on, and Get from, Higher Education by Neal McCluskey, Cato Institute Policy Analysis no. 686 (October 27, 2011)

The Poverty of Preschool Promises: Saving Children and Money with the Early Education Tax Credit by Adam B. Schaeffer, Cato Institute Policy Analysis no. 641 (August 3, 2009)
The Other Lottery: Are Philanthropists Backing the Best Charter Schools? by Andrew J. Coulson, Cato Institute Policy Analysis no. 677 (June 6, 2011)

Giving Kids the Chaff: How to Find and Keep the Teachers We Need by Marie Gryphon, Cato Institute Policy Analysis no. 579 (September 25, 2006)

RECENT STUDIES IN THE CATO INSTITUTE POLICY ANALYSIS SERIES

745. Obamacare: What We Know Now by Michael Tanner (January 27, 2014)

744. How States Talk Back to Washington and Strengthen American Federalism by John Dinan (December 3, 2013)


742. Privatizing the Transportation Security Administration by Chris Edwards (November 19, 2013)

741. Solving Egypt’s Subsidy Problem by Dalibor Rohac (November 6, 2013)


739. Antitrust Enforcement in the Obama Administration’s First Term: A Regulatory Approach by William F. Shughart II and Diana W. Thomas (October 22, 2013)

738. SNAP Failure: The Food Stamp Program Needs Reform by Michael Tanner (October 16, 2013)

737. Why Growth Is Getting Harder by Brink Lindsey (October 8, 2013)

736. The Terrorism Risk Insurance Act: Time to End the Corporate Welfare by Robert J. Rhee (September 10, 2013)


Published by the Cato Institute, Policy Analysis is a regular series evaluating government policies and offering proposals for reform. Nothing in Policy Analysis should be construed as necessarily reflecting the views of the Cato Institute or as an attempt to aid or hinder the passage of any bill before Congress. Contact the Cato Institute for reprint permission. All policy studies can be viewed online at www.cato.org. Additional printed copies of Cato Institute Policy Analysis are $6.00 each ($3.00 each for five or more). To order, call toll free (800) 767-1241 or email catostore@cato.org.