Educational Test Scores, Education Spending, and Productivity in Public Education: National Trends and Evidence Across States and Over Time, 1990-2015

John Garen
Rex Bray

January 2018

Institute for the Study of Free Enterprise
Working Paper 20

University of Kentucky
244 Gatton College of Business and Economics
Lexington, KY 40506-0034
http://isfe.uky.edu/

Educational Test Scores, Education Spending, and Productivity in Public Education: National Trends and Evidence Across States and Over Time, 1990 – 2015

John Garen
BB&T Professor of Economics
University of Kentucky

Rex Bray
Harvard Law School

January 2018

Abstract

We examine national trends in educational funding, test score outcomes, and productivity as well as variations in funding and test scores over time and across states to assess how changes in educational spending are (or are not) related to changes in educational test score outcomes for states. National trends show small increases in test scores, large increases in educational funding (until the last recession), and a continued fall in educational productivity. The cross-state, over time analysis indicates a statistically significant but very small association of state funding to test scores; so small that large changes in funding have little effect on scores. This is consistent with the continued decline in educational productivity. We also find similar results for black students, implying that the increased funding has not served to reduce racial inequality. We suggest that the continued decline in productivity of public schools adds further reason to question the ability of non-competitive, public organizations to improve educational performance and to look for alternatives that embrace or emulate private-sector, competitive organizations.

*For helpful comments, we thank Aaron Yelowitz, Frank Scott, James Fackler, John Merrifield and session participants at the Southern Economic Association annual conference and the Association of Private Enterprise Education annual meetings. We are also grateful to the John H. Schnatter Institute for Free Enterprise and the BB&T Program for the Study of Capitalism, both at the University of Kentucky, for support.

I. Introduction

Concerns about the performance of traditional public schools have been with us for quite a number of years. Though disputes regarding school funding seem to occur regularly, the data are clear in showing increasing resources being devoted to public schools over the past several decades. The increasing use of resources has reinforced concerns regarding public school performance. Essentially, this is a question about what is being attained with the dollars spent on K-12 education.

Thus, an important aspect of this paper is to examine the relationship of school funding to student outcomes. A closely related concept is the productivity of education spending, i.e., educational "output" (or outcome) per dollar spent. In an examination of this question over ten years ago, Hoxby (2004) finds that the productivity of public education declined substantially – by nearly 50 percent – from the 1970s to 2000. Her measure of educational outcomes is the National Assessment of Education Progress (NAEP) test scores. From 1970 to 2000, these scores hardly changed, yet inflation-adjusted, per-pupil spending almost doubled.

There have been some signs of improvement in the NAEP test scores in the 2000s, though these experiences have varied across states, as have changes in resources devoted to schools. Accordingly, this paper considers two related aspects of these issues. First, we update overall national trends on educational funding, test score outcomes, and productivity. Second, we examine variations in funding and test scores over time and across states to assess how changes in educational spending are (or are not) related to changes in educational test score outcomes for states.

Our findings regarding national trends show small increases in test scores, large increases in educational funding (until the last recession), and a continued fall in educational productivity.

The cross-state, over time analysis indicates a statistically significant but very small association of state funding to test scores. Our preferred estimates imply that the magnitude is so small that higher funding of $1,000 is associated with trivially higher NAEP test scores.[1] This is consistent with the continued decline in productivity that we verify in the cross-state data. Also, note that we find the patterns for black students are essentially the same as for all students, implying that the increased funding has not served to raise minority outcomes and reduce racial inequality.

Hoxby (2004) suggests that the decline in the productivity of K-12 education is due to the decline of competition among jurisdictions for students and provides evidence to support this. Moreover, there is a good deal of literature regarding the incentive problems of government-operated organizations that face little competition, which characterizes most public schools. For an overview and discussion of this literature as it pertains to schools, see Garen (2016). The results of our paper are consistent with Hoxby (2004). Though it is appropriate to be cautious regarding causality with respect to funding and outcomes, the continued decline in productivity of public schools that we find adds further reason to question the ability of non-competitive, public organizations to improve educational performance and to look for alternatives that embrace or emulate private-sector, competitive organizations.

The question of the effect of school resources and spending on educational outcomes has a long history. Coleman (1966) was perhaps the first to do a broad-based examination of the importance of school resources in K-12 education, finding that other factors were much more critical. Hanushek (1986) reviews the empirical work that followed Coleman (1966), indicating that differences in school quality do not seem to reflect variations in expenditure, class size, or other commonly measured attributes of schools and teachers. A good deal of empirical work on

[1] See Merrifield (2009) for a related analysis. He takes extant estimates of the effects of resources on school outcomes and simulates the effect of substantially greater resources. He finds very small effects.

this topic continued with improved data and more advanced methods. Later surveys by Hanushek (2003), Gustafsson (2003), and Glewwe (2013) find strong effects of teachers but an absence of consistent effects of school expenditures in reviews of research pertaining to many countries around the world. These lack-of-effects findings are the norm. However, there are some exceptions, and some studies find positive effects on the subsequent earnings and employment experiences of students "exposed" to higher education spending when young. See, for example, Jackson, Johnson, and Persico (2016), Fredriksson, Ockert, Oosterbeek (2013), and Card and Krueger (1992). Betts (1995), however, following a similar methodology, finds no effects of school resources, consistent with the bulk of the literature.

Our paper fits into this stream of research with focus on the recent experiences of states in the U.S. Though the small effects of expenditures on public school performance suggests an evaluation of alternative types of school organizations (e.g., charter schools and voucher programs), this is beyond the scope of this paper. The reader is referred to Garen (2016) for a discussion and critique of that literature.

The remainder of the paper is organized as follows. Section II discusses our measure of educational outcomes, as well as our data on school expenditures. We follow Hoxby (2004) and others in using the National Assessment of Educational Progress (NAEP) test scores to measure educational outcomes. Section III provides an update on and discussion of the long-term trend in productivity in education in the U.S. from 1971 to 2012. As we detail below, though test scores have improved in the recent past, spending rose even faster and productivity continues to decline. Section IV examines the data on test scores and spending by state and over time to assess the association of test score improvements with increased state spending. We find a positive and statistically significant association, though very small in magnitude, e.g., a $1,000 increment to

per pupil spending is associated with minimal changes in NAEP test scores. The findings for black students are approximately the same, indicating that the funding has not reduced racial inequality in test scores. We also find somewhat more robust effects of local funding relative to state and federal, though magnitudes remain small. The decline in educational productivity is also verified in the cross-state panel data. Lastly, section V concludes.

II. Measuring Educational Outcomes

We follow Hoxby (2004) in using the National Assessment of Educational Progress (NAEP) test scores as our measure of educational outcomes. We utilize the mathematics test scores and the reading test scores of 4th and 8th graders. The NAEP is the largest nationally representative and continuing assessment of the nation's students. Since NAEP assessments are administered uniformly using the same sets of test booklets across the nation, they provide a common measure of student achievement across the country. The assessments stay essentially the same from year to year, with only carefully documented changes to reflect changes in curriculum in the nation's schools. NAEP's long-term trend assessment is a national sample begun in the 1970s and currently is given every four years. The state assessments began in the 1990s and allows comparisons across states. It is now conducted every two years. These assessments are referred to as *The Nation's Report Card*. More detail is in National Center for Educational Statistics (2017).

The common measures across states and over time are a great advantage over other test-based data sets. Many states have conducted tests for a long time, but they are not consistent over time or across states. However, there are criticisms of test scores as measures of educational outcomes. Ultimately, the desired outcome from education is its enabling people to improve their lives. This can be through improvements in a variety of ways, such as school

completion, improved earnings, higher levels of employment, better health, and other outcomes. As noted above, there are studies that relate the subsequent labor market experiences of people to their school environment when young. However, these data are often limited in scope. Also, the NAEP tests, and others like them, measure important aspects of cognitive skills. However, it is becoming increasingly recognized that non-cognitive skills, such as persistence, motivation, and dependability, have great importance in determining success in life. See Heckman, Stixrud, and Urzua (2006). NAEP test score outcomes only indirectly measure these attributes.

Nevertheless, it is well established that scores on tests similar to the NAEP are strongly correlated with a person's labor market earnings. See Currie and Thomas (2001), for example. Moreover, Hanushek, Ruhose, and Woessmann (2017) find that U.S. state GDP growth is related to the state's average NAEP test score. The NAEP test score outcomes measure something that matters.

We collect data on education spending and school enrollment from the National Center for Educational Statistics' *Digest of Education Statistics*. Several measures of spending are collected. We collect data on the total school revenue collected by the state, as well as the total from each source; local, state, or federal. We also obtain data on total expenditures, current expenditures, and instructional expenditures (available only since 1986). Each is expressed on a per pupil basis using enrollment data. It is also adjusted for inflation and expressed in constant 2016 dollars.

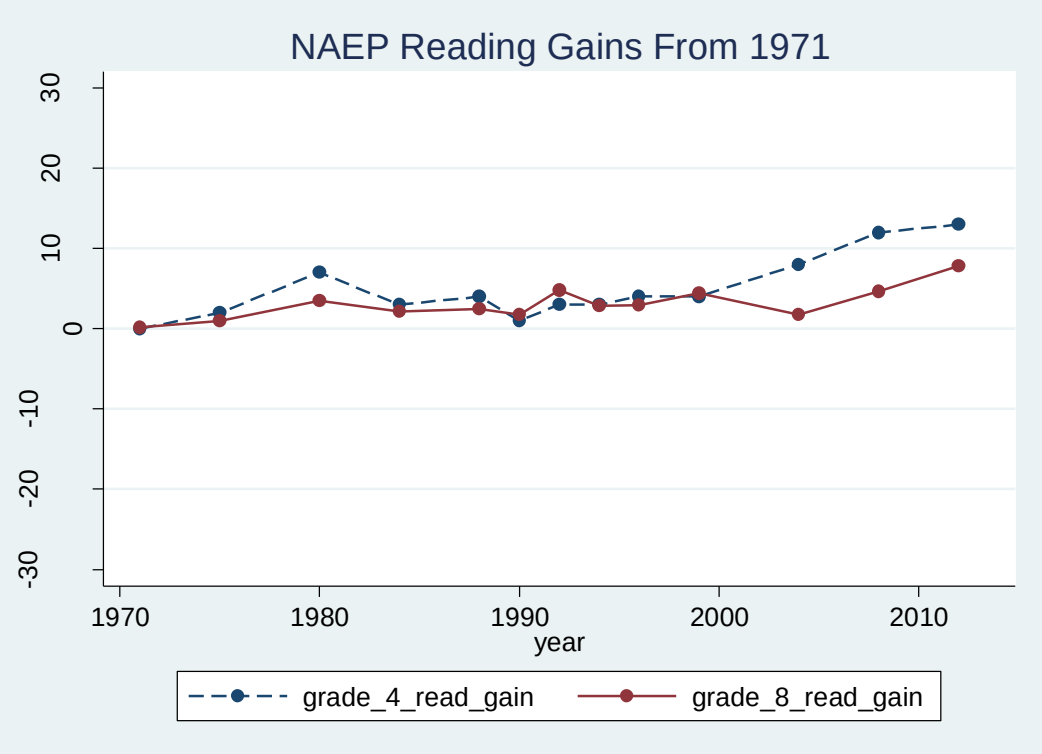
III. National Trends in Scores, Spending, and Productivity

This section presents national long-term trends in NAEP scores from NAEP's *Long-Term Trend Assessment*, along with national educational spending data from the *Digest of Education*

Statistics. Test scores are national averages. Revenue and expenditure data are per pupil national averages.

Figure 1 presents a graph of the gains in NAEP reading scores since 1971 for both grade 4 and grade 8. This is calculated by subtracting the 1971 score from the actual score in each year the test was given.[2] The latest year available is 2012. The raw data for these computations is in appendix Table A1. The dashed (blue) line is for grade 4 and the solid (red) line represents grade 8. As can be seen, the gains in test scores are small, especially through the late 1990s. Grade 4 scores have improved through the 2000s. There is an uptick in grade 8 scores beginning in 2008.

Figure 1



[2] The scale of this graph (-30 to +30) is selected because the standard deviation of individual NAEP test scores is roughly 30 each year.

Figure 2 shows similar calculations for grade 4 and grade 8 NAEP math scores. These are gains from 1978 since that is the first year the math tests were given. Again, the raw data is in appendix Table A1. Here, there is a persistent upward trend in test scores through most of this time period.

Figure 2

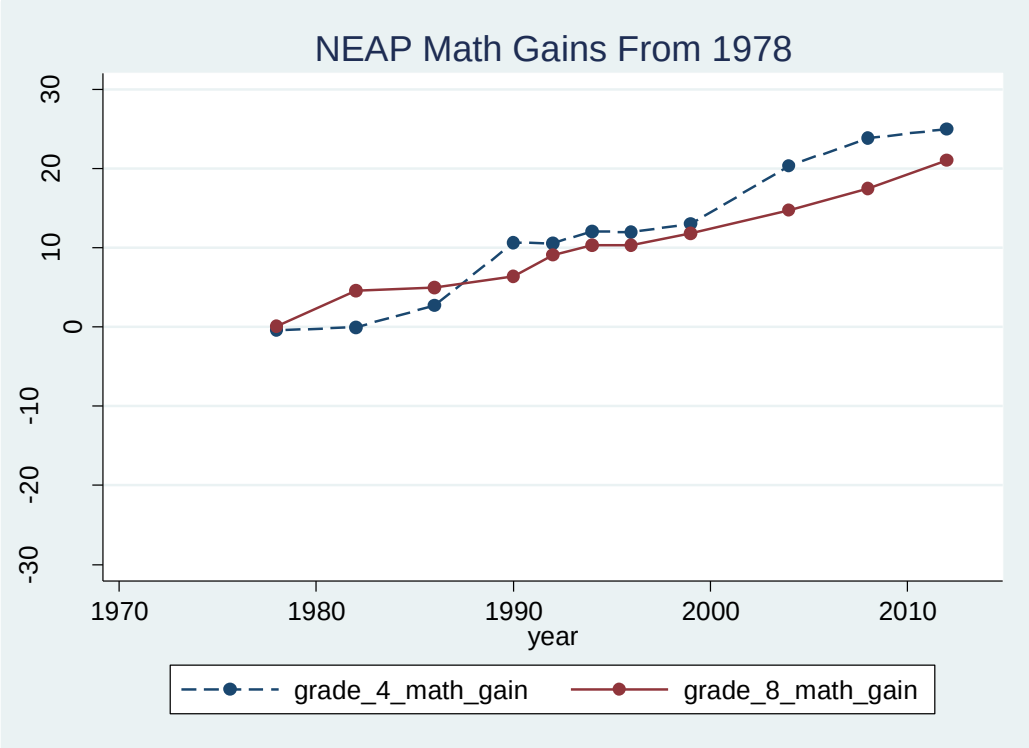


Figure 3 presents a graph of per pupil revenue and expenditure for public schools in the U.S. from 1971 to 2012. They are adjusted for inflation and expressed in 2016 dollars. The data for Figure 3, as well as Figures 4 and 5, are in appendix Table A2. In Figure 3, revenue per pupil is the dashed (blue) line and expenditure is the solid (red) line. They track one another very closely and their correlation coefficient is .9988. There is a strong and persistent upward trend. There have been only a handful of years until 2010 in which education revenue or expenditure

fell or was unchanged. The declines in years 2010 to 2012 are due to the effects of the recent recession and the consequent reductions in state and local budgets.

Figure 3

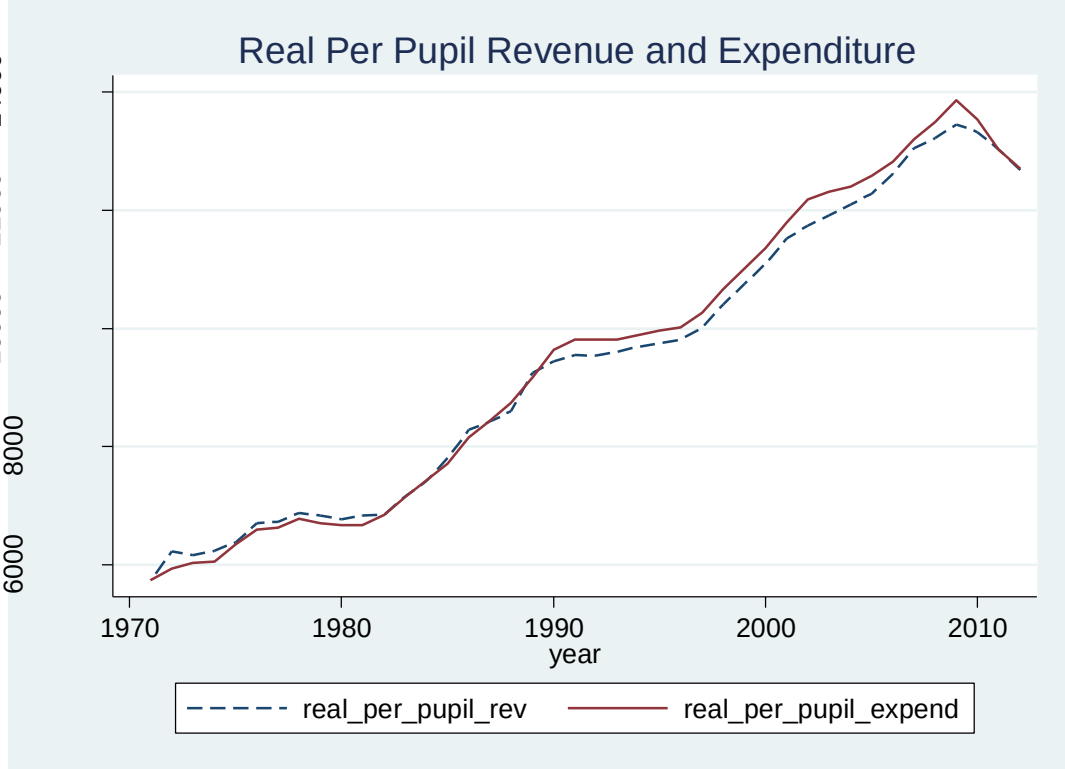
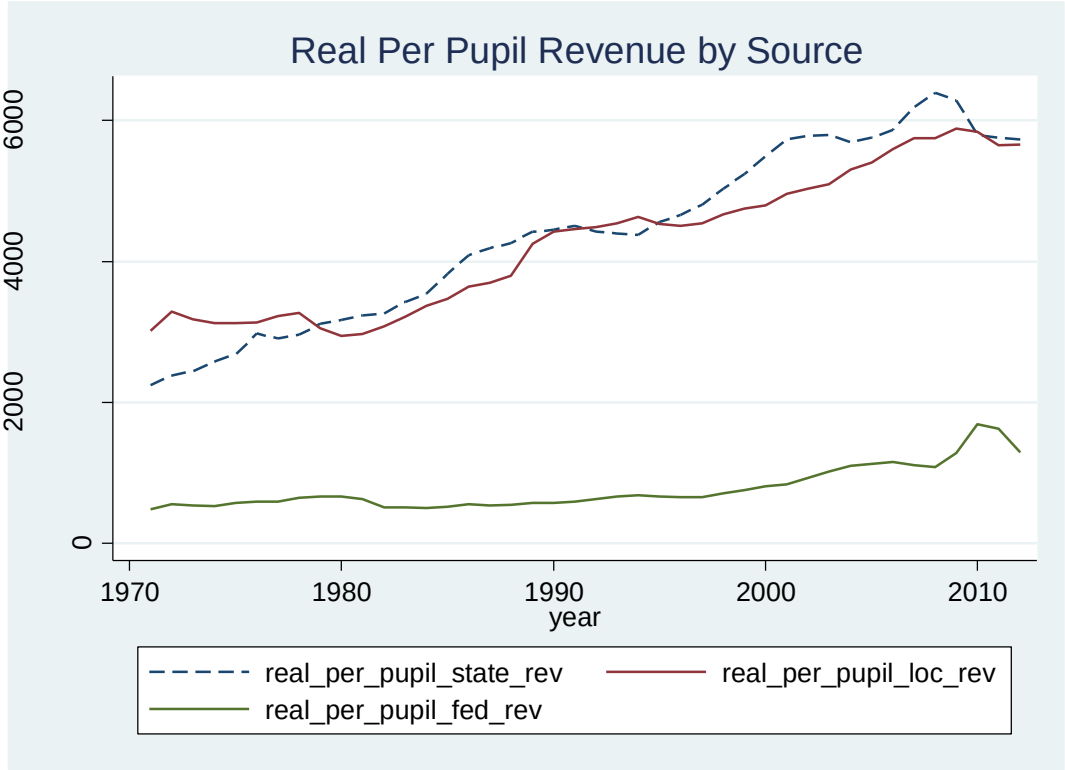


Figure 4 shows the graph of per pupil education revenue by the source of that revenue. The dashed (blue) line is state revenue, the higher solid (red) line is local revenue, and the lower solid (green) line is federal revenue. Both state and local revenue generally show a similar pattern to that of total revenue; strong upward trends until the years of the recent recession. Federal revenue shows a much more modest increase, though this picked up in the early 2000s. In the later 2000s, there was a sharp increase in the aftermath of the recession, then a decline.

Figure 5 displays some components of total expenditures. The dashed (blue) is the time path of current expenditures (all expenditures less capital expenses). The solid (red) line is instructional expenditure, available only since 1986. These both show a pattern very similar to

Figure 4



that of total expenditures; rising substantially throughout this time period except for the period immediately following the recent recession.

The next set of figures (6 and 7) examine and update the trend in the productivity of educational funding. This is computed as the NAEP test score per $1,000 of educational revenue. This can be thought as the effectiveness of education dollars in producing educational outcomes. We compute this ratio for each of the four NAEP tests we are considering. These productivity data are reported in appendix table A3. Figure 6 shows the plot over time for grades 4 and 8 reading scores.

Figure 5

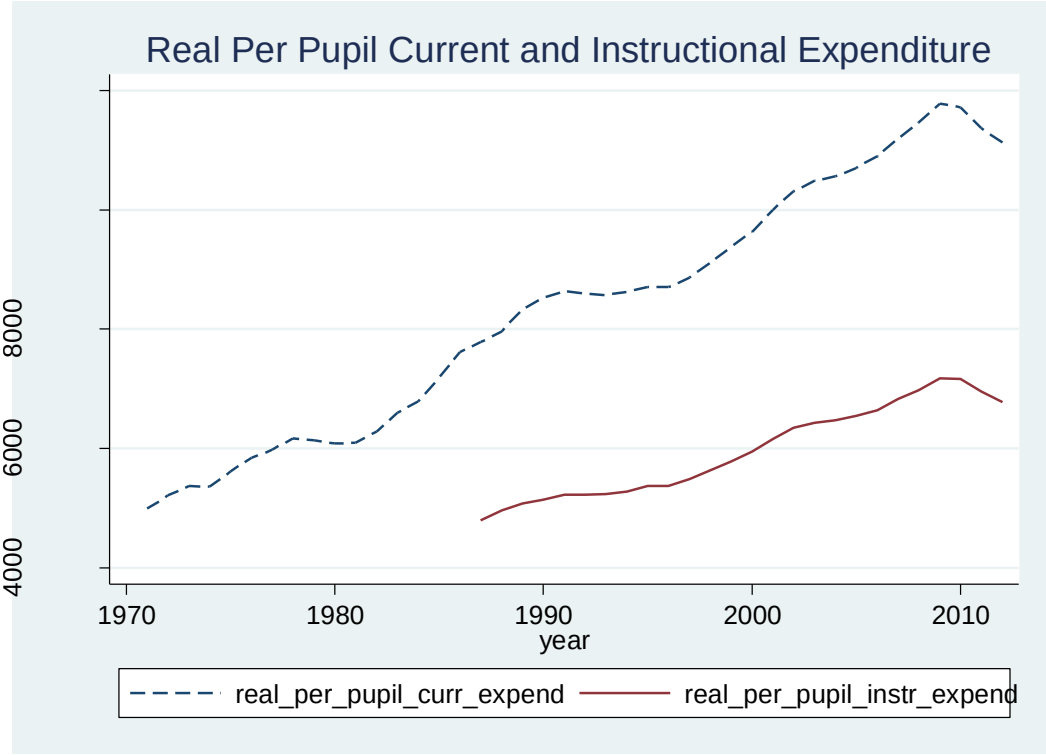
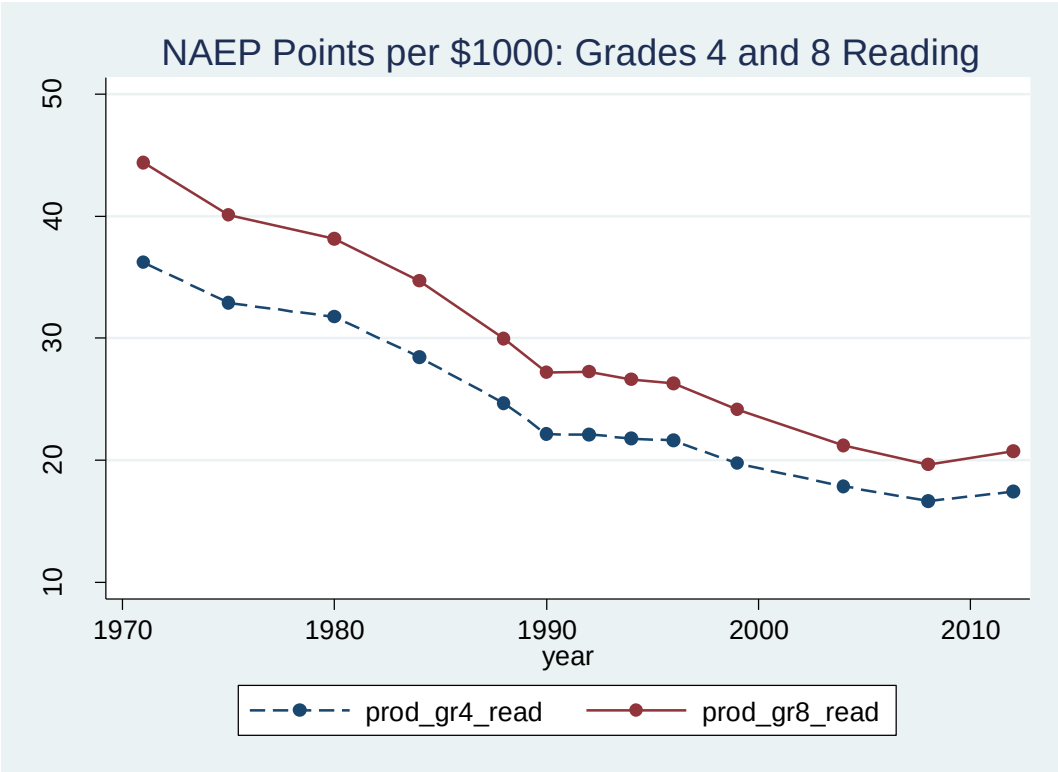


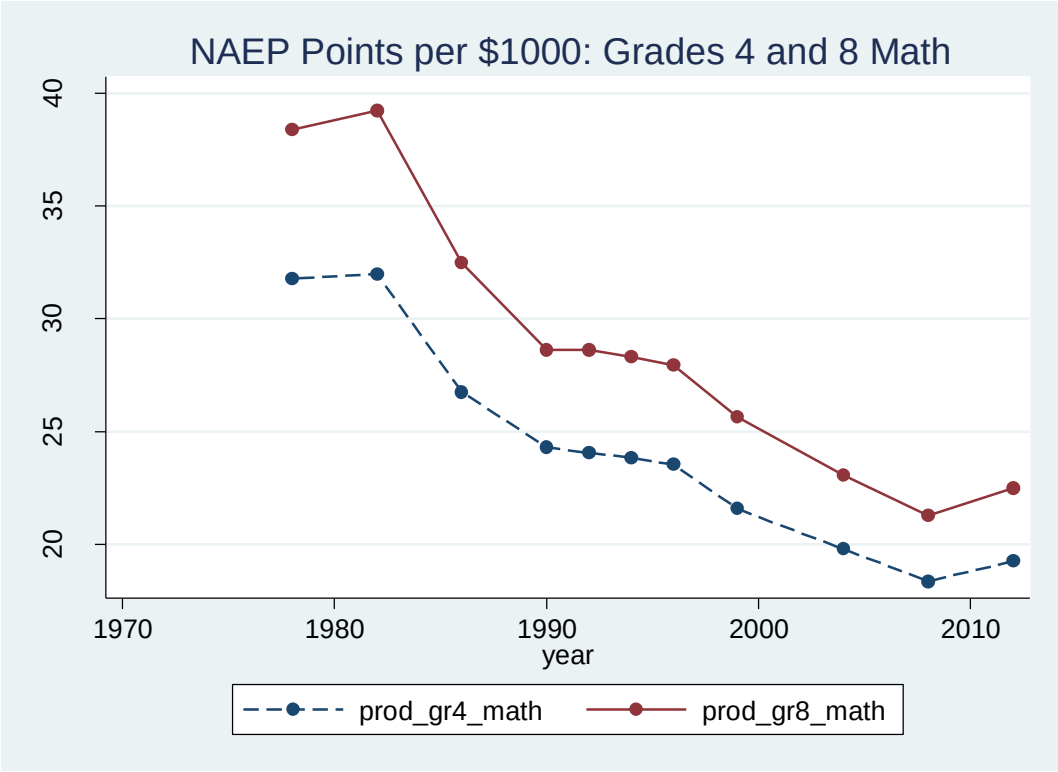
Figure 6



The dashed (blue) line in Figure 6 is for grade 4 and the solid (red) line is for grade 8. Both display an essentially continuous downward trend for the entire time period. The exception is the last two years of tests that show an uptick. This is largely due to the post-recession decline in education spending. From 2008 to 2012, education revenue per pupil fell by 4%. However, grade 4 reading scores increased by ½ of 1% and grade 8 reading scores rose by 1%. Thus, most of the increased productivity in this time frame was due to declining funding.

Figure 7 shows the plot of productivity for grade 4 and grade 8 math. There is a general similarity to the reading score plot; an overall and mostly continuous downward trend. Here, there is an exception between the 1978 and 1982 tests, as well as for the final 2008 to 2012 period.

Figure 7



Visually, the rates of decline seem sharpest prior to 1990, followed by smaller declines in the 1990s, then further declines in the 2000s that are not as large as in the earliest period. This is verified in Table 1 below. We show the annual average decline in productivity for each test for three time periods: 1990 and before, 1990 to 1999, and 1999 and after. For the last period, we consider both the 1999 – 2008 period and the 1999 – 2012 period, where the former excludes the effect of the last recession.

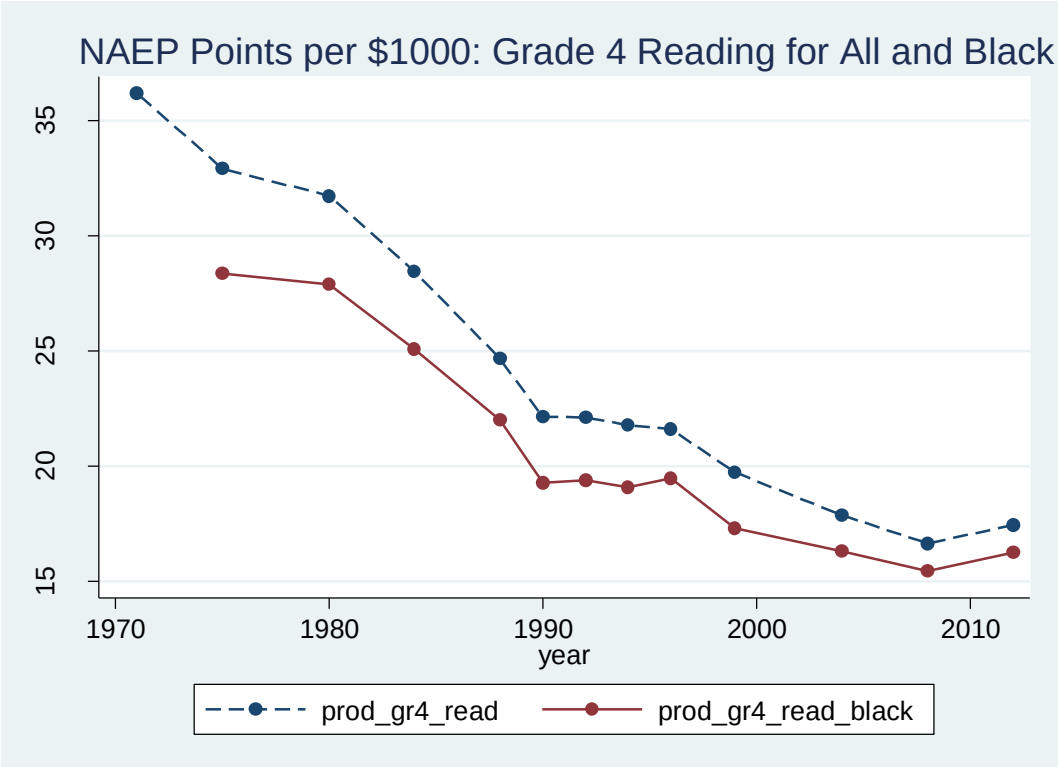
Table 1 verifies the visual analysis. Up through 1990, the annual average absolute decline in productivity for each test ranges from 0.62 points per year to 0.90. The percent decline is about 2% per year for each test. During the 1990s, the decline in productivity continued, but at a slower rate. For each test, the absolute decline was roughly 0.30 points per year, or a little over 1% reduced productivity per year. From 1999 to 2008, productivity fell faster than in the 1990s, but slower than the earlier period. Declines in lost points per year ranged from 0.34 to 0.5, translating into percentage reductions between 1.6% and 2.0%. With the time period 1999 to 2012, absolute and percentage reductions are the smallest due to the 2008 to 2012 uptick. Note, though, in none of these time periods did productivity rise. The differences are how quickly it declines.

Figures 8 and 9 plot the productivity for all students (as in previous figures), along with the productivity for black student scores, i.e., the average NAEP score for black students per $1,000 of educational revenue. This addresses the issue of whether the additional funding for schools over this time period was targeted toward minorities in order to raise their scores and educational outcomes, and thereby reduce racial inequality. If this were the case, we expect to see higher black student test scores per $1,000 of spending relative to all students.

Table 1

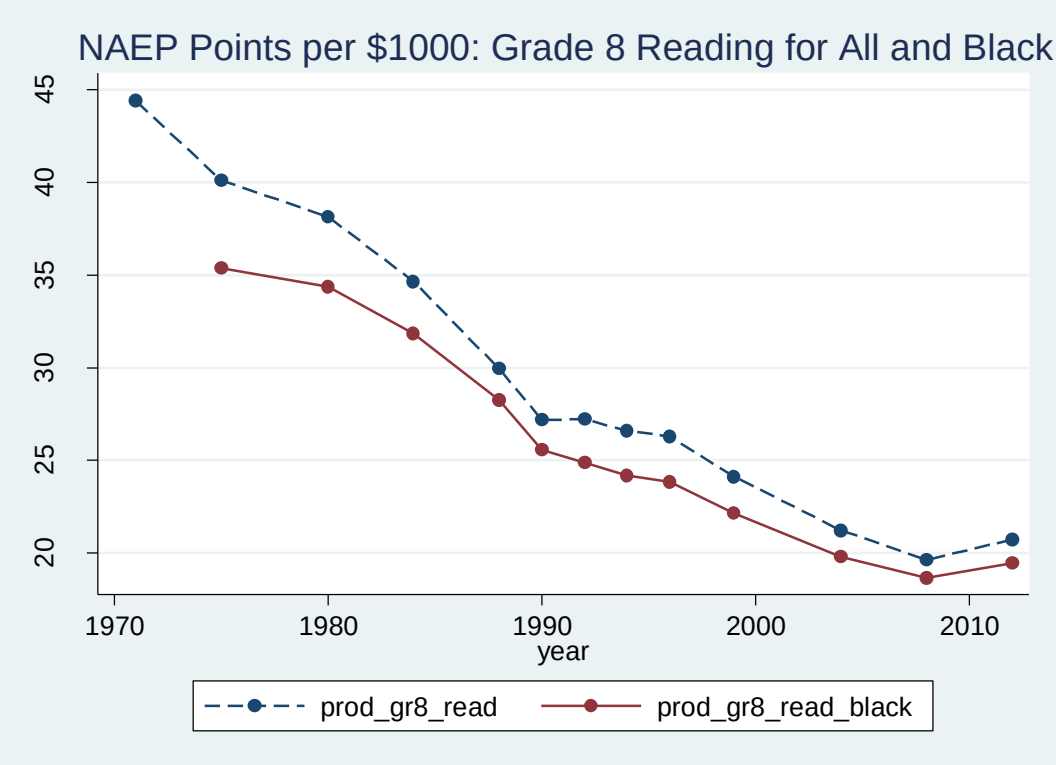
Annual Average Absolute and Percent Decline in Productivity (NAEP Score per $1000): Grades 4 and 8 Reading and Math								
	Grade 4 Reading		Grade 8 Reading		Grade 4 Math		Grade 8 Math	
	Absolute Annual Prod. Decline	Percent Annual Prod. Decline	Absolute Annual Prod. Decline	Percent Annual Prod. Decline	Absolute Annual Prod. Decline	Percent Annual Prod. Decline	Absolute Annual Prod. Decline	Percent Annual Prod. Decline
1990 and before	0.74	2.1%	0.90	2.0%	0.62	2.0%	0.82	2.1%
1990 to 1999	0.27	1.2%	0.34	1.3%	0.30	1.2%	0.32	1.1%
1999 to 2008	0.34	1.7%	0.5	2.0%	0.35	1.6%	0.49	1.9%
1999 to 2012	0.18	0.9%	0.26	1.1%	0.18	0.95%	0.24	1.0%

Figure 8



Figures 8 and 9 show that this did not occur. Figure 8 is for grade 4 reading and Figure 9 is for grade 8 reading. In both figures, the dashed (blue) line represents all students and the solid (red) line for black students. The trend in productivity for the two groups track one another almost exactly. Black student productivity outcomes show the same pattern of decline as all students. Repeating this exercise for math scores yields the same outcome.

Figure 9



Overall, the findings for nationwide trends are not that different from Hoxby's (2004), even with consideration of the last decade of data. Test scores do show some improvement over the 2000s, but education spending increased rapidly until the last recession. The productivity of educational funds continues to decline, albeit not as rapidly as in the 1970s and 1980s. The pattern for productivity is the same for black students as all students.

IV. Spending and Test Scores Across States and Over Time

This section utilizes the state-level NAEP test score data and state-level educational funding to examine the relationship between school spending and test scores, as well as the productivity of school spending. An advantage of these data is that states have varying experiences regarding changes in funding and test score outcomes over time, and states naturally vary in these respects at any point in time. There is much greater variation in the relevant data than in the national data, and so ought to be more informative in assessing the association of spending and test scores. We also consider the effects of spending on black student scores, examine if there are differences based on whether the source of funds is federal, state, or local, and reexamine the productivity of education funding with the cross-state panel.

A. Summary Statistics and Econometric Specification

Appendix Table A4 shows the years the state-based NAEP tests were given. The first began in 1990 and they were given at various times to most, though not all, states. Since 2003, the assessments that we use have been given every two years to all states.

The basic methodology used is regression analysis of the form in equation (1).

(1) $$T_{jt} = b_0 + b_1 S_{jt} + b_2 X_{jt} + \theta_j + \delta_t + \varepsilon_{jt}$$

where j indexes states, t indexes time and:

T_{jt} = state j's average test score at time t (where we estimate separate equations for each test),

S_{jt} = state j's real per pupil funding at time t,

X_{jt} = a vector of demographics for state j at time t,

θ_j = a vector of state effect dummy variables,

δ_t = a vector of year effect dummy variables, and

ε_{jt} = white noise.

We adopt and report a number of specifications that utilize different subsets of the above variables. The full specification allows for the influence of state spending on test scores, with controls for differences and changes in state demographics, state-specific influences, and time-dependent effects.

Table 2 presents means and standard deviations, as well as variable definitions, for the test score, funding, and demographic variables. The columns show these summary statistics for each test, with the data pooled across states and over time. Prior to 2003, different tests were given at different intervals, so sample sizes are not the same for each test. Also, the primary variable we use for school funding is real revenue per pupil, lagged one year. The mean for this variable in each column is approximately $12,000 per pupil in 2016 dollars.[3] We use funding with a one-year lag largely because the last year of test score data is 2015 and the last year of funding data is 2014. Thus, using lagged funding enables us to use the 2015 test score data. However, dropping 2015 or using different lag lengths do not substantially change our findings.

The demographics we use include the racial composition of the state's public school students, the percent of the state's students eligible for the federal reduced price lunch program, the percent of students identified as English language learners, and the percent of students with disability. States have some latitude in whether English language learners and students with disability are tested, so we include the percent of each that were assessed. After 1998, states were enabled to test these two groups with accommodation, and we include the percent who were assessed with accommodation.

[3] We also experimented using expenditures instead of revenue but with essentially identical findings.

Table 2: Means (standard deviations), State Level NAEP Scores, Educational Revenue, and Demographics

Variable[a]	*Grade 4 Reading*	*Grade 8 Reading*	*Grade 4 Math*	*Grade 8 Math*
NAEP State Score	218.29 (8.03)	263.23 (6.76)	234.25 (10.56)	276.87 (11.07)
Real revenue per pupil (lagged one year)	12023.81 (3422.29)	12530.18 (3389.74)	12163.85 (3432.21)	11956.71 (3439.64)
Percent White students	65.05 (20.22)	64.74 (20.23)	63.69 (19.87)	66.61 (20.64)
Percent Black students	15.73 (16.16)	15.43 (16.05)	15.32 (15.77)	15.23 (16.33)
Percent Hispanic students	11.64 (13.54)	11.57 (12.79)	12.99 (12.97)	10.69 (12.55)
Percent Reduced Price School Lunch Eligible	41.89 (13.78)	42.95 (11.47)	43.22 (13.55)	39.41 (13.42)
Pct. Students English Language Learners	7.66 (7.38)	5.74 (5.59)	6.27 (6.02)	3.70 (3.64)
Pct. of English Language Learners assessed	5.99 (6.59)	4.60 (5.07)	5.52 (5.68)	3.12 (3.39)
Pct. of English Language Learners tested with accommodation	2.69 (4.08)	2.26 (3.73)	2.56 (2.74)	1.28 (1.45)
Pct. of Students with Disability	13.59 (2.89)	13.49 (2.50)	13.58 (2.89)	12.55 (2.90)
Pct. of Students with Disability assessed	9.50 (3.56)	10.09 (2.89)	10.91 (3.63)	9.50 (3.63)
Pct. of Students with Disability tested with accommodation	5.64 (4.09)	6.92 (3.62)	7.16 (4.34)	6.20 (4.50)
N (sample size)	523	436	484	517

[a]All variables are statewide averages for each sample year.

B. Basic Findings

Naturally, we are most interested in the coefficient b_2, showing the relationship of state funding to test scores. Table 3 shows the findings of four different specifications for the grade 4

reading exam. The coefficient on funding is transformed to indicate the influence of $1,000 of per pupil funding. The full set of coefficients is in appendix Table A5. Column (1) shows the

Table 3: Coefficient Estimates, Grade 4 Reading Equation

Dependent Variable: NAEP Grade 4 Reading Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil, lagged one year ($1,000)	.6119 (6.16)	.2567 (2.94)	.6992 (6.51)	.2549 (2.18)
Demographics	no	yes	yes	yes
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.0680	0.6432	0.8853	0.9108
N	523	523	523	523

results of a univariate regression of the test score on funding, with no control variables. The coefficient of .6119 indicates that a $1,000 increase in per pupil funding is associated with a little over a 0.6 point rise in the grade 4 NAEP reading score. Column (2) adds controls for the demographic variables since a state's average score may be influenced by its demographic composition. With this specification, the coefficient on funding falls to .2567, implying that additional funding of $1,000 per student is associated with just over a ¼ point higher test score.

Because there are likely to be many difference across states not fully captured by the demographic variation, column (3) adds state fixed effects to account for them. Doing so allows for different "baseline" levels for each state and so our results reflect the effects of within-state changes over time. In this specification, the coefficient on funding is .6992, indicating that a $1,000 increase in funding by a state is associated with an almost 0.7 point higher test score. Lastly, column (4) adds time effect dummy variables. This allows for nationwide trends from other, unmeasured sources to affect test scores and is the standard difference-in-differences

specification (with covariates), and is our preferred specification. The results here reflect the effects of within-state changes in funding relative to other states as a whole. Here, the coefficient on funding is .2549, implying that a $1,000 increase in funding is associated with a little over a ¼ point increase in the test score.[4]

Care must be taken in interpreting these coefficients as causal. Clearly, funding is endogenous to policy makers, and if unobservables drive both funding and test scores, the associations we find may not be causal. However, the state and time effects may account for many of these unobservable factors.

Note that all of the estimated effects of funding have a high degree of statistical significance. However, the magnitudes of the effects are very small. Focus on the estimates in columns (3) and (4). These, we believe, are the most meaningful. Recall that the column (4) specification accounts for demographics, state effects, and time effects, and shows the effect of changes in state funding relative to the nation. Column (3) does not include the time effects, so the influence of unobserved, national trends may be attributed to the effect of funding. Thus, the column (3) result is probably an upper-bound estimate of the influence of funding.

Though the estimate of column (3) is over twice as large at that in column (4) – .6992 compared to .2549 – both are small in magnitude. A $1,000 increase in per pupil funding is associated with 0.7 test points for the former and ¼ point for the latter. The yearly average for this test score is about 218, and the yearly standard deviation varies over time but ranges from about 6 to almost 10. Thus, $1,000 "buys" only a very minimal change in the test score with either estimate. Note that $1,000 is a substantial share of per pupil spending. From Table 2, the

[4] Though column (4) is our preferred specification, we report other specifications so the reader may gain a fuller picture of the findings.

average of per pupil spending is about $12,000 per year. An increment of $1,000 is an 8.3% change.

The theme from Table 3 recurs with the other tests: effects that are statistically significant but very small in magnitude. This is evident from Table 4, which shows the findings for the grade 8 reading test. The full set of coefficients is in appendix Table A6. Column (1) estimates the effect of $1,000 more in funding at .30 point, column (2) indicates .11 points, column (3) about .25 points, and column (4) .20 points. Each is statistically significant except for the coefficient in column (2). Regarding our preferred specifications of column (4) and the upper-bound estimate of column (3), their magnitudes are not that different and both are very small. Yearly standard deviations for this test range between 6 and 7. Thus, a .20 or .25 change in the test score is quite minimal.

Table 4: Coefficient Estimates, Grade 8 Reading Equation

Dependent Variable: NAEP Grade 8 Reading Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil, lagged one year ($1,000)	.3064 (3.24)	.1131 (1.57)	.2478 (3.09)	.2033 (2.35)
Demographics	no	yes	yes	yes
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.0236	0.7636	0.9366	0.9499
N	436	436	436	436

Table 5 presents the findings for grade 4 math NAEP test scores, with full results in appendix Table A7. Focus again on columns (3) and (4). The column (3) estimate indicates a gain of 1.22 points per $1,000 of funding, while that in column (4) shows 0.29 points. The average score for this test is around 234, and yearly standard deviation ranges from 5 to 8.5.

Though the column (3) estimate is larger than that for the reading score, its magnitude is still not that large. The magnitude of the column (4) estimate is small as previously.

Table 5: Coefficient Estimates, Grade 4 Math Equation*

Dependent Variable: NAEP Grade 4 Math Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil, lagged one year ($1,000)	1.2101 (9.39)	.2094 (1.97)	1.2159 (8.14)	.2902 (2.55)
Demographics	no	yes	yes	yes
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.1546	0.6981	0.8916	0.9538
N	484	484	484	484

Lastly, Table 6 presents the findings for grade 8 math scores, with full results in appendix Table A8. Regarding the coefficients in columns (3) and (4), the former indicates a gain of 1.39 points per $1,000 of funding, while that in column (4) shows 0.52 points. Given an average test score of around 234 and a yearly standard deviation that ranges from 7 to 11, the magnitude of either of these estimates is not especially large.

Table 6: Coefficient Estimates, Grade 8 Math Equation

Dependent Variable: NAEP Grade 8 Math Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil, lagged one year ($1,000)	1.1562 (8.74)	.437 (3.41)	1.3896 (10.78)	.5193 (4.00)
Demographics	no	yes	yes	yes
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.1291	0.6638	0.9171	0.9443
N	517	517	517	517

Table 7 summarizes the association of per pupil spending with test scores for each of the NAEP tests. The first row of this table recaps the findings presented in Tables 3 through 6. In particular, we show the effect of a $1,000 increase in per pupil funding based on the column (3) and the column (4) estimates. The estimates without year effects are larger, though, as noted above, these are upper-bound estimates. Also as noted, these effects are small in magnitude.

The second row of Table 7 presents another way to illustrate the magnitude of the effects. Based on the coefficient estimates, we calculate the additional funding needed to move each test score by one-half of a standard deviation. With such a change, a state at the 33rd percentile of scores would move to the median, or a state at the median would move to the 66th percentile. Because the standard deviation of test scores varies from year to year, we use its average over time for each test.

As can be seen, these dollar values are quite large. For grade 4 reading and the column (1) value, the additional required funds is $5,422. Using the column (2) value, it is even larger at $14,863. Note that the mean state per pupil funding (from Table 2) is about $12,000 and for 2014 it is $13,362. Thus, either of our computations imply a massive funding increase to attain such a score change: a 40% rise with the column (1) computation – based on 2015 mean spending – and an over 100% increase with column (2). Of course, these large values stem entirely from the small values of the coefficient estimates.

Regarding grade 8 reading, both computations indicate over a 100% rise in funding to attain the one-half standard deviation score increase. For the grade 4 and grade 8 math scores, a somewhat different picture emerges. Referring to columns (5) and (7), the necessary funding increase is $2,788 for grade 4 and $3,209 for grade 8. These translate into 21% and 24% funding increases, respectively. The computations from columns (6) and (8), though, show much larger

values. For the grade 4 math computation in column (6), the funding increase needed is 87% and for grade 8 math in column (8) it is 63%.

Table 7: Summary Estimates: The Association of Per Pupil Funding to Test Scores

	Grade 4 Reading		Grade 8 Reading		Grade 4 Math		Grade 8 Math	
	(1)[a]	(2)[b]	(3)[a]	(4)[b]	(5)[a]	(6)[b]	(7)[a]	(8)[b]
Points from $1K of funds	.699	.255	.248	.203	1.216	.290	1.390	.529
Funding for a ½σ score change[c]	$5,422	$14,863	$13,327	$16,281	$2,788	$11,690	$3,209	$8,431

[a]Based on column (3) estimates in the above four tables with state fixed effects.
[b]Based on column (4) estimates in the above four tables with state and year fixed effects.
[c]Yearly standard deviations of test scores vary. We use the average yearly standard deviation for each test. One-half of these are 3.79 for grade 4 reading, 3.305 for grade 8 reading, 3.39 for grade 4 math, and 4.46 for grade 8 math.

C. Findings for Black Student Scores

We consider if the influence of funding has a different effect on average black student scores than on the average of all students. If the effect on black student scores is substantially larger, then the funding tends to reduce racial inequality in scores.

Table 8 shows the findings for both grade 4 and 8 reading scores, where the dependent variable is the state's average NAEP score for black students. We report on the specification with just state effects and also that with both state and year effects. Overall the findings are similar to above in that the estimates with state effects only are larger, but the magnitude of the effects are quite small. The effect of funding on black student scores is somewhat larger for some of the specification, though not for all.[5]

[5] The sample sizes for the black student results are smaller because not all states report black student scores for all years. When we estimate using the same sample of states, comparable findings appear.

Table 8: Coefficient Estimates, Grade 4 and 8 Reading, Black Students

Dependent Variable: NAEP Reading Score, Black Students
(absolute value of t-statistics in parentheses)

	Grade 4		Grade 8	
Variable	(1)	(2)	(3)	(4)
Real revenue per pupil, lagged one year ($1,000)	.9114 (4.73)	.0645 (0.32)	.5788 (3.84)	.3445 (2.00)
Demographics	yes	yes	yes	yes
State effects	yes	yes	yes	yes
Year effects	no	yes	no	yes
R^2	.7405	.8172	.7712	.7982
N	447	447	359	359

Table 9 presents the results for grade 4 and grade 8 math scores for black students, organized as in Table 8. Very similar comments apply as with the reading scores. Even though a number of the coefficients are larger for black student scores, the small magnitudes imply that

Table 9: Coefficient Estimates, Grade 4 and 8 Math, Black Students

Dependent Variable: NAEP Math Score, Black Students
(absolute value of t-statistics in parentheses)

	Grade 4		Grade 8	
Variable	(1)	(2)	(3)	(4)
Real revenue per pupil, lagged one year ($1,000)	1.3818 (6.44)	.0927 (0.53)	1.8914 (9.31)	.7355 (3.40)
Demographics	yes	yes	yes	yes
State effects	yes	yes	yes	yes
Year effects	no	yes	no	yes
R^2	.8633	.9326	.8536	.8920
N	411	411	416	416

there is very little effect on score inequality. For example, if the effect on black students is 0.50 higher than on all students (the largest of the different effects that we find), a $1,000 increase in funding raises black student scores by a half point more that all students. The differential

between the average score of black students and all of students varies across the tests, but is at least 16 points (the black-white differential is larger). Thus, such a change implies a very minimal effect on score inequality

D. Effects by Source of Funds

Funds for public schools may come from three sources: the federal government, state government, or local government.[6] Because funds from different levels of government may be earmarked for certain uses and have various restrictions and rules on their use, the effectiveness may differ by fund source. Thus, we re-estimate our test score equations by replacing the total per pupil funding with three variables: federal, state, and local per pupil funding. Tables 10 and 11 present the findings. We show only two specifications; that with state effects and that with state and year effects. The results for grades 4 and 8 reading are in Table 10, and those for the math scores are in Table 11.

As with the previous results, the coefficient for the state effect only specifications are almost always larger for reasons noted above. Generally, the effects of local funding are the most robust in the sense of consistently attaining statistical significance and often being the largest in magnitude. However, as above, their magnitudes are still such that the association with test scores is very, very small.

State per pupil funding tends to be the smallest in magnitude and often does not attain statistical significance, with the exception of the grade 8 math scores. Federal funding has strong effects in the state effects only specifications, but generally has very weak effects when year effects are included.

[6] There are also small amounts of funds from private sources through fees and donations that we count as local.

Table 10: Coefficient Estimates, Grade 4 and 8 Reading, Effects of Funding Source

Dependent Variable: NAEP Reading Score
(absolute value of t-statistics in parentheses)

	Grade 4		Grade 8	
Variable	(1)	(2)	(3)	(4)
Real federal revenue per pupil, lagged one year ($1,000)	.6529 (1.21)	.482 (0.57)	.6409 (1.76)	.7207 (1.19)
Real state revenue per pupil, lagged one year ($1,000)	.362 (2.31)	-.1002 (0.66)	.100 (0.84)	.0434 (0.38)
Real local revenue per pupil, lagged one year ($1,000)	1.0299 (6.39)	.5758 (3.80)	.3052 (2.55)	.3286 (2.90)
Demographics	yes	yes	yes	yes
State effects	yes	yes	yes	yes
Year effects	no	yes	no	yes
R^2	.8877	.9134	.9371	.9505
N	523	523	436	436

Table 11: Coefficient Estimates, Grade 4 and 8 Math, Effects of Funding Source

Dependent Variable: NAEP Math Score
(absolute value of t-statistics in parentheses)

	Grade 4		Grade 8	
Variable	(1)	(2)	(3)	(4)
Real federal revenue per pupil, lagged one year ($1,000)	3.6484 (4.98)	-.3461 (0.42)	3.4682 (5.50)	-.8934 (0.89)
Real state revenue per pupil, lagged one year ($1,000)	.8357 (4.19)	.0858 (0.60)	1.121 (6.42)	.3458 (2.15)
Real local revenue per pupil, lagged one year ($1,000)	1.1897 (6.01)	.5033 (3.64)	1.2634 (7.31)	.7316 (4.75)
Demographics	yes	yes	yes	yes
State effects	yes	yes	yes	yes
Year effects	no	yes	no	yes
R^2	.8952	.9546	.9193	.9451
N	484	484	517	

E. Revisiting the Productivity of Educational Funding

This subsection revisits the productivity of educational funding over time with use of the cross-state panel of observations rather than with the national-level data, as in the previous section. These data provide over time as well as across state variation in funding and test scores to examine productivity. For each state and year, we compute productivity as above: NAEP test score points per $1,000 of funding. We then estimate models as in equation (2).

(2) $$P_{jt} = a_0 + a_1 X_{jt} + \theta_j + \delta_t + u_{jt}$$

where j indexes states, t indexes time. The term P_{jt} is state j's productivity at time t (where we estimate separate equations for each test), u_{jt} is white noise, and X_{jt}, θ_j, and δ_t are as above: a vector of demographics for state j at time t, a vector of state effect dummy variables, and a vector of year effect dummy variables, respectively.

We report two specifications for each test score. The first is with only time dummies and the second also includes the demographics and the state dummies. Thus, we consider the time path of productivity, both with and without adjustment for demographics and state effects. Table 12 reports the findings for grades 4 and 8 reading scores. The "baseline" is mean productivity in the first year the test was given. Coefficients show the change in productivity from the baseline year, both with no adjustment (columns (1) and (3)) and with adjustment (columns (2) and (4)).

The time effect coefficients in Table 12 are all negative with most statistically significant. Generally, the coefficients become larger with time, until 2011 when they decline. This is consistent with the national level data on productivity, and corresponds to the decline in funding after the recession. Moreover, the magnitudes of the coefficients of columns (1) and (3) in Table 12 are quite consistent with national productivity declines found in the national data.

Table 12: Productivity of Funding on Reading Scores, by Year, Grades 4 and 8

Dependent Variable: NAEP Reading Score per $1,000 of School Funding
(absolute value of t-statistics in parentheses)

	Grade 4		Grade 8	
Year	(1)	(2)	(3)	(4)
1992	24.4235[a]	24.4235[a]		
1994	-.8695 (0.94)	-.8686 (2.74)		
1998	-1.7158 (1.85)	-2.4483 (5.32)	27.9202[a]	27.9202[a]
2002	-4.3905 (4.86)	-5.7622 (11.26)	-3.6477 (3.40)	-3.6646 (11.14)
2003	-5.2197 (5.98)	-6.4491 (12.08)	-4.7267 (4.60)	-4.1996 (12.05)
2005	-5.7519 (6.59)	-6.8210 (13.90)	-5.5221 (5.37)	-4.6342 (12.07)
2007	-6.4392 (7.38)	-8.1799 (14.21)	-6.5185 (6.34)	-5.8904 (13.34)
2009	-7.2994 (8.37)	-9.5670 (15.36)	-7.4351 (7.23)	-7.3969 (14.71)
2011	-7.6469 (8.76)	-10.3871 (14.64)	-7.7497 (7.54)	-7.7747 (15.37)
2013	-6.6353 (7.61)	-8.9815 (10.70)	-6.4754 (6.30)	-4.5775 (5.73)
2015	-6.4536 (7.40)	-9.8629 (12.10)	-6.4809 (6.30)	-6.7137 (10.52)
Demographics	no	yes	no	yes
State effects	no	yes	no	yes
R^2	.2621	.9368	.1633	.9508
N	523	523	436	436

[a]Baseline year for each test.

Table 13 presents the analogous results for grades 4 and 8 math scores. All year effect coefficients are negative, are mostly significant, and their magnitudes in columns (1) and (3) line up quite closely with the national data. Regarding columns (2) and (4) in both tables, these

coefficients show that same pattern, but the magnitudes tend to be larger. This implies that if demographics and other state-level conditions remained the same as they were in the first year of the test, productivity would have dropped even more

Table 13: Productivity of Funding on Math Scores, by Year, Grades 4 and 8

Dependent Variable: NAEP Math Score per $1,000 of School Funding
(absolute value of t-statistics in parentheses)

	Grade 4		Grade 8	
Year	(1)	(2)	(3)	(4)
1990			30.0736[a]	30.0736[a]
1992	24.7569[a]	24.7569[a]	-.17887 (0.15)	-.58662 (1.46)
1996	-1.3355 (1.39)	-2.1153 (4.75)	-1.1689 (0.98)	-2.2341 (3.79)
2000	-2.6333 (2.70)	-4.105 (8.47)	-3.2697 (2.72)	-5.2611 (7.98)
2003	-4.0896 (4.42)	-5.1371 (8.42)	-5.6190 (4.94)	-7.7371 (10.10)
2005	-4.4560 (4.81)	-5.7665 (8.89)	-6.2840 (5.53)	-8.6970 (12.26)
2007	-5.2316 (5.65)	-6.7580 (10.52)	-7.1534 (6.29)	-10.2214 (12.78)
2009	-6.1206 (6.61)	-7.9755 (11.39)	-8.0966 (7.12)	-11.5434 (14.83)
2011	-6.4292 (6.95)	-8.5372 (11.07)	-8.4706 (7.45)	-12.5275 (14.44)
2013	-5.2793 (5.70)	-7.6944 (9.15)	-7.2230 (6.35)	-11.763 (12.19)
2015	-5.3076 (5.73)	-7.8027 (8.95)	-7.2777 (6.40)	-12.0277 (11.75)
Demographics	no	yes	no	yes
State effects	no	yes	no	yes
R^2	.1632	.9357	.2416	.9328
N	484	484	517	517

[a]Baseline year for each test.

V. Conclusion

It is clear that educational productivity, as measured by NAEP test scores per dollar of funding, continues to decline. Though test scores did improve during the 2000s, this improvement was far outstripped by increased educational funding, causing further reductions in productivity. Our analysis of the cross-state panel shows positive and statistically significant associations of school funding to NAEP test scores, but with very small magnitudes. Our preferred estimates imply that the magnitude is so small that $1,000 more of per pupil funding is associated with less than one point higher NAEP test scores (where scores average over 200). This is consistent with the continued decline in productivity that we verify in the cross-state data. Similar results hold for black student scores, implying that the increased funding has not served to raise minority outcomes or to reduce racial inequality in test scores. We examine the influence of federal, state, and local funding separately, and generally find the effect of local funding to be the largest and most robust. However, the magnitude of the effects remain very small.

While it is appropriate to be cautious regarding causality with respect to funding and outcomes, the continued decline in educational productivity seems very evident. Schools are predominantly government-sponsored organizations, often with little competition, and the literature is quite clear regarding incentive problems in these types of organizations. The lack of efficiency should not be surprising. We suggest looking for improvements in educational performance by turning to alternatives that embrace or emulate private-sector, competitive organizations.

References

Betts, Julian B., "Does School Quality Matter? Evidence From the National Longitudinal Survey of Youth," *Review of Economics and Statistics*, 77(2), May 1995, pp. 231-250.

Card, David and Krueger, Alan B., "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100(1), February 1992, pp. 1-40.

Coleman, James S., *Equality of Educational Opportunity*, U.S. Department of Health, Education, and Welfare, 1966.

Currie, Janet and Thomas, Dunca, "Early Test Scores, School Quality and Socioeconomic Status: Long Run Effects on Wage and Employment Outcomes," in *Research in Labor Economics: Worker Wellbeing in a Changing Labor Market*, v .20, 2001, p. 103-132.

Fredriksson, Peter; Ockert, Bjorn; and Oosterbeek, Hessel, "Long-Tern Effects of Class Size," *Quarterly Journal of Economics*, February 2013, pp. 249-285.

Garen, John, "Assessing the Literature on School Reform From an Entrepreneurship Perspective," *Journal of Entrepreneurship and Public Policy*, 5(3), 2016, pp. 383-403.

Glewwe, Paul; Hanushek, Eric A.; Humpage, Sarah D.; and Ravina, Renato, "School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010," in Glewwe, Paul (ed.), *Education Policy and Policy in Developing Countries*, Chicago: University of Chicago Press, 2013, pp. 13-64.

Gustafsson, Jan-Eric, "What Do We Know About Effects of School Resources on Educational Results?," 10, *Swedish Economic Policy Review*, January 2003, pp. 77-110.

Hanushek, Eric A., 'The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature*, 24(3), September 1986, pp. 1141-1177.

Hanushek, Eric A., "The Failure of Input-Based Schooling Policies," *Economic Journal*, 113, February 2003, pp. 65-98.

Hanushek, Eric A.; Ruhose, Jens; and Woessmann, Ledger, "Economic Gains from Education Reform in the US States," *Journal of Human Capital*, 11(4), Winter 2017, p. 447-486.

Heckman, James; Stixrud, Jora ; and Urzua, Sergio (2006), "The Effect of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,"*Journal of Labor Economics*, Vol. 24 No. 3, pp. 411-482.

Hoxby, Caroline M., "Productivity in Education: The Quintessential Upstream Industry," *Southern Economic Journal*, 71(2), 2004, pp. 209-231.

Jackson, C. Kirabo; Johnson, Rucker C.; and Persico, Claudia, "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms," *Quarterly Journal of Economics*, 131(1), February 2016, pp. 157-218.

Merrifield, John, "The Potential for System-Friendly K–12 Reform," *Cato Journal*, Vol. 29, No. 2, Spring/Summer 2009.

National Center for Educational Statistics, *About NAEP*, last updated May 10, 2017, https://nces.ed.gov/nationsreportcard/about/.

National Center for Educational Statistics, *About State NAEP*, https://nces.ed.gov/nationsreportcard/about/state.aspx, electronic data at https://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx

National Center for Educational Statistics, *Digest of Educational Statistics*, https://nces.ed.gov/programs/digest/, electronic data at https://nces.ed.gov/programs/digest/current_tables.asp.

National Center for Educational Statistics, *The Nation's Report Card: Trends in Academic Progress 2012*, U.S. Department of Education, https://nces.ed.gov/nationsreportcard/subject/publications/main2012/pdf/2013456.pdf, electronic data at https://nces.ed.gov/nationsreportcard/lttdata/.

Appendix

Table A1: NAEP National Reading and Mathematics Scores

Panel (a): Reading NAEP Scores and Gains from 1971

	(1)	(2)	(3)	(4)
Year	**Gr 4 Read**	**Gr 4 Gain from 1971**	**Gr 8 Read**	**Gr 8 Gain from 1971**
1971	208	0	255	0
1975	210	2	256	1
1980	215	7	258	3
1984	211	3	257	2
1988	212	4	257	2
1990	209	1	257	2
1992	211	3	260	5
1994	211	3	258	3
1996	212	4	258	3
1999	212	4	259	4
2004	216	8	257	2
2008	220	12	260	5
2012	221	13	263	8
Mean	212.92		258.09	

Panel (b): Math NAEP Scores and Gains from 1978

	(5)	(6)	(7)	(8)
Year	**Gr 4 Math**	**Gr 4 Gain from 1978**	**Gr 8 Math**	**Gr 8 Gain from 1978**
1978	219	0	264	0
1982	219	0	269	5
1986	222	3	269	5
1990	230	11	270	6
1992	230	11	273	9
1994	231	12	274	10
1996	231	12	274	10
1999	232	13	276	12
2004	239	20	279	15
2008	243	24	281	17
2012	244	25	285	21
Mean	230.80		274.08	

Source: National Center for Educational Statistics, *The Nation's Report Card,* https://nces.ed.gov/nationsreportcard/lttdata/.

Table A2: Real School Revenue and Expenditure, U.S., 2016 $

Year	Total Revenue	Revenue, Federal	Revenue, State	Revenue, Local	Total Expenditure	Current Expenditure	Instructional Expenditure
1971	5748	485	2248	3015	5748	4990	
1972	6232	557	2385	3291	5938	5211	
1973	6161	535	2447	3179	6034	5368	
1974	6238	528	2583	3127	6056	5361	
1975	6381	575	2680	3126	6351	5610	
1976	6706	595	2976	3134	6596	5844	
1977	6731	592	2907	3229	6628	5976	
1978	6880	650	2958	3272	6779	6171	
1979	6837	668	3118	3050	6709	6134	
1980	6776	665	3172	2940	6670	6084	
1981	6837	630	3238	2968	6679	6092	
1982	6848	509	3259	3080	6850	6280	
1983	7156	508	3428	3220	7148	6594	
1984	7418	505	3545	3369	7428	6791	
1985	7811	518	3821	3471	7708	7187	
1986	8284	554	4089	3640	8156	7619	
1987	8425	539	4190	3696	8440	7779	4793
1988	8598	543	4260	3795	8744	7966	4955
1989	9248	573	4420	4255	9171	8337	5074
1990	9446	575	4450	4421	9637	8526	5143
1991	9549	589	4503	4457	9809	8638	5226
1992	9544	630	4426	4488	9807	8593	5227
1993	9606	670	4399	4537	9812	8571	5236
1994	9693	683	4377	4633	9884	8626	5277
1995	9752	663	4560	4528	9961	8707	5375
1996	9815	652	4662	4501	10016	8702	5372
1997	10006	659	4803	4544	10271	8861	5482
1998	10406	709	5032	4665	10672	9113	5635
1999	10753	759	5241	4753	11015	9376	5782
2000	11091	806	5491	4794	11358	9632	5948
2001	11523	835	5730	4957	11794	10001	6153
2002	11740	928	5780	5032	12184	10309	6343
2003	11910	1015	5798	5096	12310	10489	6433
2004	12093	1097	5690	5306	12403	10558	6477
2005	12284	1129	5756	5399	12582	10705	6549
2006	12620	1153	5870	5598	12817	10898	6636
2007	13048	1107	6190	5751	13200	11196	6825
2008	13224	1079	6392	5752	13492	11462	6971
2009	13453	1287	6279	5886	13859	11784	7177
2010	13323	1695	5793	5835	13540	11718	7163
2011	13028	1629	5752	5647	13031	11370	6955
2012	12676	1289	5730	5657	12703	11126	6776

Source: National Center for Educational Statistics, *Digest of Educational Statistics*, https://nces.ed.gov/programs/digest/current_tables.asp.

Table A3: Productivity. NAEP Points per $1000 of Education Revenue

Year	Grade 4 Read	Grade 8 Read	Grade 4 Math	Grade 8 Math
1971	36.2	44.4		
1975	32.9	40.1		
1978			31.8	38.4
1980	31.7	38.1		
1982			32.0	39.2
1984	28.4	34.7		
1986			26.8	32.5
1988	24.7	29.9		
1990	22.1	27.2	24.3	28.6
1992	22.1	27.2	24.1	28.6
1994	21.8	26.6	23.8	28.3
1996	21.6	26.3	23.5	27.9
1999	19.7	24.1	21.6	25.7
2004	17.9	21.2	19.8	23.1
2008	16.6	19.6	18.4	21.3
2012	17.4	20.7	19.2	22.5

Source: National Center for Educational Statistics, *The Nation's Report Card,* https://nces.ed.gov/nationsreportcard/lttdata/ and *Digest of Educational Statistics*, https://nces.ed.gov/programs/digest/current_tables.asp.

Table A4: Years of State-Based NAEP Assessments

Year	Grade 4 Reading	Grade 8 Reading	Grade 4 Math	Grade 8 Math
1990				selected states
1992	selected states	selected states	selected states	
1994	selected states			
1996			selected states	selected states
1998	selected states	selected states		
2000			selected states	selected states
2002	selected states	selected states		
2003	all states	all states	all states	all states
2005	all states	all states	all states	all states
2007	all states	all states	all states	all states
2009	all states	all states	all states	all states
2011	all states	all states	all states	all states
2013	all states	all states	all states	all states
2015	all states	all states	all states	all states

Source: National Center for Educational Statistics, *The Nation's Report Card.*

Table A5: Coefficient Estimates, Grade 4 Reading Equation*

Dependent Variable: NAEP Grade 4 Reading Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil (lagged one year)	.0006119 (6.16)	.0002567 (2.94)	.0006992 (6.51)	.0002549 (2.18)
Percent White		.1639922 (7.92)	-.2041423 (3.79)	-.0031506 (0.06)
Percent Black		-.0901961 (3.92)	-.573056 (7.91)	-.4595293 (6.53)
Percent Hispanic		.1033978 (4.00)	.0380793 (1.31)	.0245305 (0.94)
Percent School Lunch Eligible		-.156467 (6.45)	-.0027509 (0.09)	-.13664 (3.42)
Pct. English Language Learners		-.3307864 (2.63)	-.0077471 (0.08)	-.02928 (0.32)
Pct. English Language Learners assessed		.1026955 (0.67)	-.1103212 (0.97)	-.1582443 (1.44)
Pct. English Language Learners accommodated		.3362454 (3.24)	.1116681 (1.32)	.2991223 (2.87)
Pct. Students with Disability		.4376456 (3.39)	.3860687 (3.91)	.2853661 (2.86)
Pct. Students with Disability assessed		-.8676102 (5.28)	-.7096491 (6.14)	-.5043579 (4.60)
Pct. Students with Disability accommodated		1.17883 (8.11)	.4411311 (3.83)	-.06893 (0.58)
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.0680	0.6432	0.8853	0.9108
N	523	523	523	523

*For sources and variable definitions, see Table 2.

Table A6: Coefficient Estimates, Grade 8 Reading Equation*

Dependent Variable: NAEP Grade 8 Reading Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil (lagged one year)	.0003064 (3.24)	.0001131 (1.57)	.0002478 (3.09)	.0002033 (2.35)
Percent White		.1534081 (9.78)	.0788794 (1.49)	.097364 (1.80)
Percent Black		-.0741953 (4.17)	-.3238809 (4.49)	-.2197445 (3.16)
Percent Hispanic		.1035303 (4.94)	.0542785 (0.94)	.0465352 (0.88)
Percent School Lunch Eligible		-.221742 (10.28)	.1793948 (6.26)	.004827 (0.14)
Pct. English Language Learners		.2975124 (1.51)	.2615761 (2.12)	.045122 (0.37)
Pct. English Language Learners assessed		-.8038768 (3.68)	-.4592573 (3.09)	-.3011538 (2.12)
Pct. English Language Learners accommodated		.7371854 (7.03)	.3659163 (4.68)	.228682 (2.39)
Pct. Students with Disability		.003481 (0.03)	.1903165 (2.11)	.3077137 (3.37)
Pct. Students with Disability assessed		-.7712474 (5.66)	-.7275993 (7.55)	-.8404803 (8.98)
Pct. Students with Disability accommodated		.8434071 (8.32)	.3534601 (3.13)	.2961936 (3.04)
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.0236	0.7636	0.9366	0.9499
N	436	436	436	436

*For sources and variable definitions, see Table 2.

Table A7: Coefficient Estimates, Grade 4 Math Equation*

Dependent Variable: NAEP Grade 4 Math Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil (lagged one year)	.0012101 (9.39)	.0002094 (1.97)	.0012159 (8.14)	.0002902 (2.55)
Percent White		.107246 (3.68)	.2917496 (2.92)	.1228174 (1.64)
Percent Black		-.136853 (4.49)	.0326449 (0.36)	-.2407664 (2.58)
Percent Hispanic		-.0241265 (0.59)	.0686704 (0.67)	.0880234 (1.26)
Percent School Lunch Eligible		-.0874069 (2.75)	.3015466 (6.77)	-.117845 (2.88)
Pct. English Language Learners		-1.571452 (5.42)	-.7669221 (3.09)	-.4131196 (2.47)
Pct. English Language Learners assessed		1.869721 (5.95)	1.037856 (4.21)	.3045977 (1.77)
Pct. English Language Learners accommodated		.2175722 (1.27)	.011931 (0.08)	.103427 (1.19)
Pct. Students with Disability		.0801938 (0.37)	.3870232 (2.11)	.2729143 (2.07)
Pct. Students with Disability assessed		-.648937 (2.35)	-.5717215 (2.64)	-.39699 (2.70)
Pct. Students with Disability accommodated		1.930095 (10.69)	1.145408 (7.24)	.1124339 (0.93)
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.1546	0.6981	0.8916	0.9538
N	484	484	484	484

*For sources and variable definitions, see Table 2.

Table A8: Coefficient Estimates, Grade 8 Math Equation*

Dependent Variable: NAEP Grade 8 Math Score
(absolute value of t-statistics in parentheses)

Variable	(1)	(2)	(3)	(4)
Real revenue per pupil (lagged one year)	.0011562 (8.74)	.000437 (3.41)	.0013896 (10.78)	.0005193 (4.00)
Percent White		.1618097 (5.58)	.3190889 (3.10)	.3438009 (3.67)
Percent Black		-.2167494 (6.67)	.0898037 (0.66)	-.1600979 (1.30)
Percent Hispanic		.0550193 (1.45)	.5961319 (4.66)	.3799881 (3.49)
Percent School Lunch Eligible		-.1101374 (3.16)	.295285 (8.71)	.0235089 (0.49)
Pct. English Language Learners		-.8544714 (2.07)	-.010864 (0.04)	-.1644773 (0.74)
Pct. English Language Learners assessed		1.028798 (2.33)	-.0082677 (0.03)	.0061493 (0.03)
Pct. English Language Learners accommodated		.0870519 (0.23)	-.2693493 (1.07)	.0159452 (0.07)
Pct. Students with Disability		-.2139781 (1.02)	.3261958 (2.41)	.1200237 (0.97)
Pct. Students with Disability assessed		-.9922245 (4.05)	-.7905019 (5.30)	-.4907514 (3.61)
Pct. Students with Disability accommodated		1.852907 (10.33)	.705193 (5.65)	.14871 (1.13)
State effects	no	no	yes	yes
Year effects	no	no	no	yes
R^2	0.1291	0.6638	0.9171	0.9443
N	517	517	517	517

*For sources and variable definitions, see Table 2.