PEGASUS INSTITUTE

Test Scores, Education Spending, and Productivity in Kentucky Public Education 1990-2015

WHITE PAPER

Educational Test Scores, Education Spending, and the Productivity of Public Education in Kentucky, 1990 – 2015

John Garen

Senior Fellow, Pegasus Institute

Steven Gordon

Associate Fellow, Pegasus Institute

July 2018

Executive Summary

This report examines the National Assessment of Educational Progress (NEAP) test scores for Kentucky students for the roughly 25-year time span from the 1990s forward, as well as changes in funding (inclusive of state, local, and federal) of K-12 in the Commonwealth during that period. Changes in the racial gap in test scores is also examined, along with the gap between the scores of students from low-income families and students from other families.

The main findings are:

- There has been some improvement in Kentucky's grade 4 and grade 8 reading and math test scores since the 1990s.
- Reading test score changes have been quite modest, where the grade 4 score rose by 5.2% from 1992 to 2017 and the grade 8 score rose by just 1.1% from 1998 to 2017.
- The grade 4 math test score improved by 11.2% over the 25 year span between 1992 and 2017, while the grade 8 math score rose by 8.2% from 1990 to 2017.
- Inflation adjusted, per pupil funding of K-12 rose very substantially throughout most of this roughly 25-year period, with the exception of 2011 to 2014 in the aftermath of the recent recession. Overall, real, per pupil funding rose by 45% between 1991 and 2015.
- The productivity of K-12 funding, calculated as NAEP test score points per $1,000 of funding, fell almost continuous during the study time period, dropping for all NAEP test by at least 20%. This implies that taxpayer funds received a lower and lower "bang per buck" in terms of test score achievement.
- The only persistent period of increases in productivity was 2009 to 2013. This was mostly due to the reduction in real, per pupil funding during this time and not to any substantial rise in test scores.
- There was essentially no reduction in the gap between the African American student and white student test scores during the study period.
- The gap between test scores of students from low-income families, as proxied by national-school-lunch-program eligible, and students from other families showed no reduction during the study period.

Though there are other sides to educational success – such as gaining creativity, perseverance, and better employment – cognitive skills as measured by the NAEP test scores are important aspects of educational achievement. By this measure, over the past two-and-a-half decades, Kentucky has improved modestly, has not reduced achievements gaps, though has almost continuously and substantially increased per pupil funding of K-12.

Introduction

Uneasiness about the performance and funding of traditional public schools have been with Kentucky, as well as most other states, for quite some time. Disputes regarding school funding emerged in the most recent legislative session and have occurred regularly in the past. Moreover, there is long-standing concern about the performance of public schools, i.e., how are the school children doing and what are the taxpayers getting out of their education funding?

This study provides an overview of three big-picture aspects of these issues. One is Kentucky's record since the 1990s regarding student achievement as measured by the National Assessment of Education Progress (NAEP) test scores. Second, we document Kentucky schools' inflation-adjusted, per pupil funding over the same time period. Third, we calculate the "productivity" of K-12 education for this time period. Productivity is defined as NEAP test score outcome (e.g., NAEP test points) per $1,000 of funding.[1] We compare these to the nation as a whole and for certain subgroups (e.g., minorities, children of low-income families).

Generally, Kentucky has experienced some increases in NAEP tests scores since 1990. However, there has been a much larger increase in per pupil funding over this time period, with the exception of the decline in funding during the most recent recession. As a result, the test score attainment per $1,000 of funding has sharply declined during most all of this time. We find similar results regarding test scores for minority students and those from low-income families, implying that the increased funding has not served to reduce inequality in achievement on test scores.

Measuring Educational Outcomes and Funding

We follow many other analysts in using the National Assessment of Educational Progress (NAEP) test scores as our measure of educational outcomes.[2] We utilize the reading test scores and the mathematics 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, so enable comparisons over time. The main assessments began in the 1990s and 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).

There are criticisms of test scores as measures of educational outcomes, however. Ultimately, the desired outcome from education is its enabling people to improve their lives in a variety of ways such as school completion, improved earnings, higher levels of employment, better health, and other outcomes. Regarding employment outcomes, there are studies that relate the subsequent labor market experiences of people to their school environment when young,

[1] Hoxby (2004) uses this measure.

[2] See Garen and Bray (2018) for a short summary of related studies.

though these studies are often limited in scope.[3] Also, while the NAEP tests, and others like them, measure important aspects of cognitive skills, it is increasingly recognized that non-cognitive skills, such as persistence, motivation, and dependability, have great importance in determining success in life.[4] NAEP test scores 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.[5] Moreover, research shows that U.S. state GDP growth is related to the state's average NAEP test score.[6] The NAEP test score achievement measure something that matters.

Regarding education funding, we collect data on education revenue and spending and on school enrollment from the National Center for Educational Statistics' (NCES) *Digest of Education Statistics*. This is available from the NCES through 2015. We express educational revenue and spending on a per pupil basis and adjust for inflation by converting all dollars figures to constant 2017 dollars. All sources of revenue are included; state, local, and federal.[7]

Tests Scores, Funding, and Productivity

Test Scores

Figure 1a presents a graph of the average NAEP test score for grade 4 reading from 1992 through 2017. Kentucky is represented by the solid blue line the entire nation by the dashed red line.[8] The underlying data for this and several following figures is in Appendix A, Table A1.

For Kentucky, there has been some improvement on this test from the late 1990s through 2009. Since then, there have been ups and downs, but little net change. From the earliest test in 1992 to the latest in 2017, this score has increased by 5.2%. Though this improvement is modest, it is more than that for the nation as a whole. For the U.S., this score improved by 2.3% and Kentucky went from being a little below the national average to a little above.

---

[3] See, for example, Jackson, Johnson, and Persico (2016), Fredriksson, Ockert, Oosterbeek (2013), Card and Krueger (1992), and Betts (1995),

[4] See Heckman, Stixrud, and Urzua (2006).

[5] Currie and Thomas (2001) is an example.

[6] Hanushek, Ruhose, and Woessmann (2017).

[7] In supplemental analyses, we account for changes in the demographics of the student population over time by controlling for possible effects of changes in students eligible for free or reduced priced lunch, students with disabilities, students who are English language learners, the testing and accommodation of the former two groups, and the racial composition of the students. They are collected from various years of National Center for Educational Statistics, *The Nation's Report Card*.

[8] The scale of this and other test score graphs is about -30 to +30 around the mean because the standard deviation of individual NAEP test scores is roughly 30 each year.

Figure 1a: 4th grade reading scores

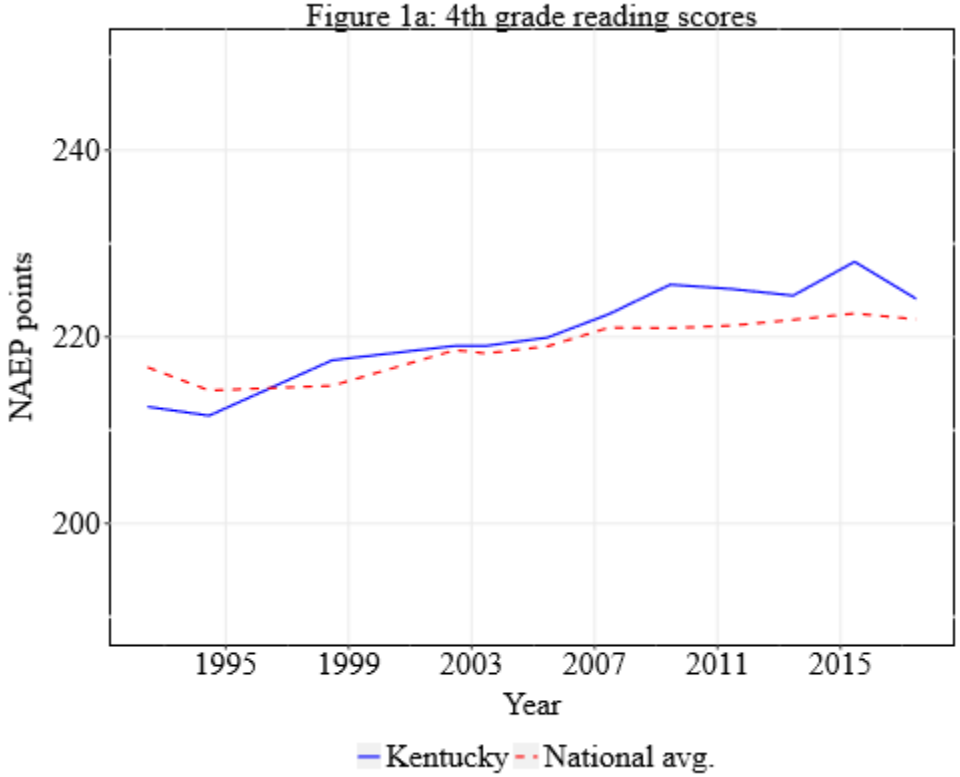


Figure 1b: 8th grade reading scores

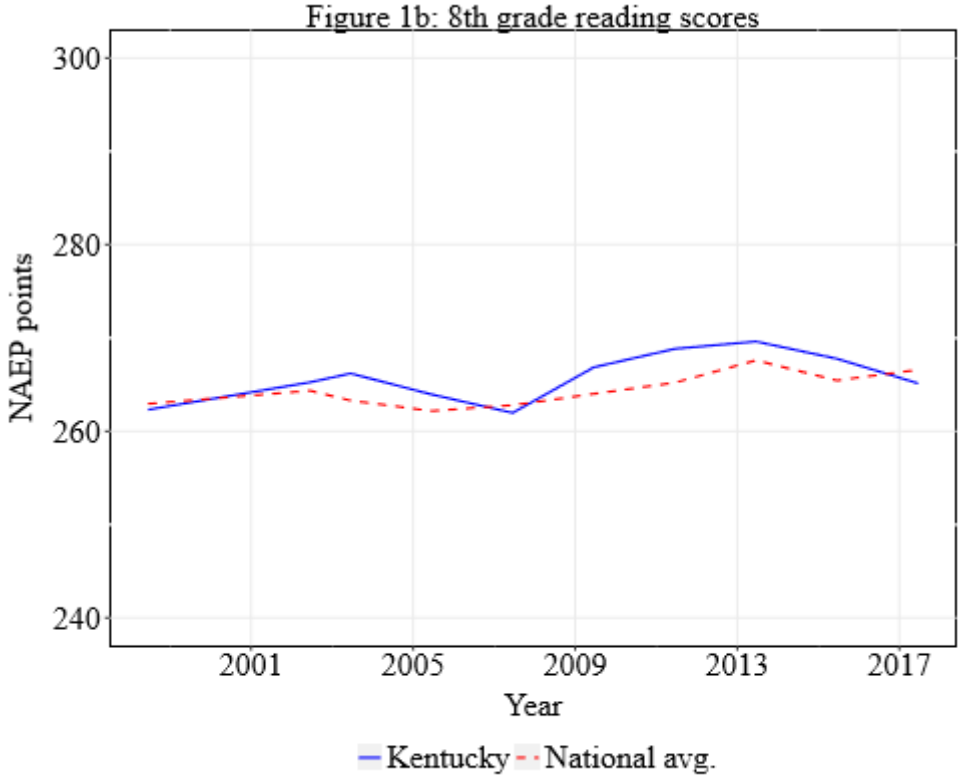


Figure 1b displays the same graph for grade 8 reading scores, available only since 1998. For Kentucky (the solid blue line), this score has seen periods of some improvements and periods of decline. Over the 1998 to 2017 period this test score has improved, on net, by just 1.1%. The scores for the entire U.S. (the dashed red line) have shown some slight ups and downs over this period as well, with a net improvement of 1.5%. Kentucky's scores were initially below the U.S., were above for several time periods and now have fallen below.

The mathematics test scores show more improvement than the reading scores and are presented in Figures 2a and 2b, with former for grade 4 and the latter for grade 8. For grade 4 scores in Kentucky, progress in the 1990s was modest, but bigger gains occurred in the 2000s up until 2009. Since then there has been little change. From 1992 to 2017, scores rose on net by 11.2%. Compared to the U.S. as a whole, Kentucky scores were lower until 2009 and have remained about even since then. U.S. scores have increased by 9.1% since 1992.

Regarding the grade 8 math scores displayed in Figure 2b, the Kentucky score rose incrementally but persistently until 2011, then began to turn down. A similar pattern emerges for the U.S. scores, though they are above the Kentucky scores by about the same amount for the entire time period. For the 1990 to 2017 time span, Kentucky scores increased by 8.2% and those for the nation by 7.6%.

We also examined the possibility that demographic changes may have been a contributor to the changes in the test scores. A change in the composition of the student population toward demographic groups that typically score lower on NAEP tests will lower a state's average test score and the opposite demographic changes will raise the scores. We do not want to conflate changes the demographic composition of the state with changes in the effectiveness of the state's

Figure 2a: 4th grade math scores

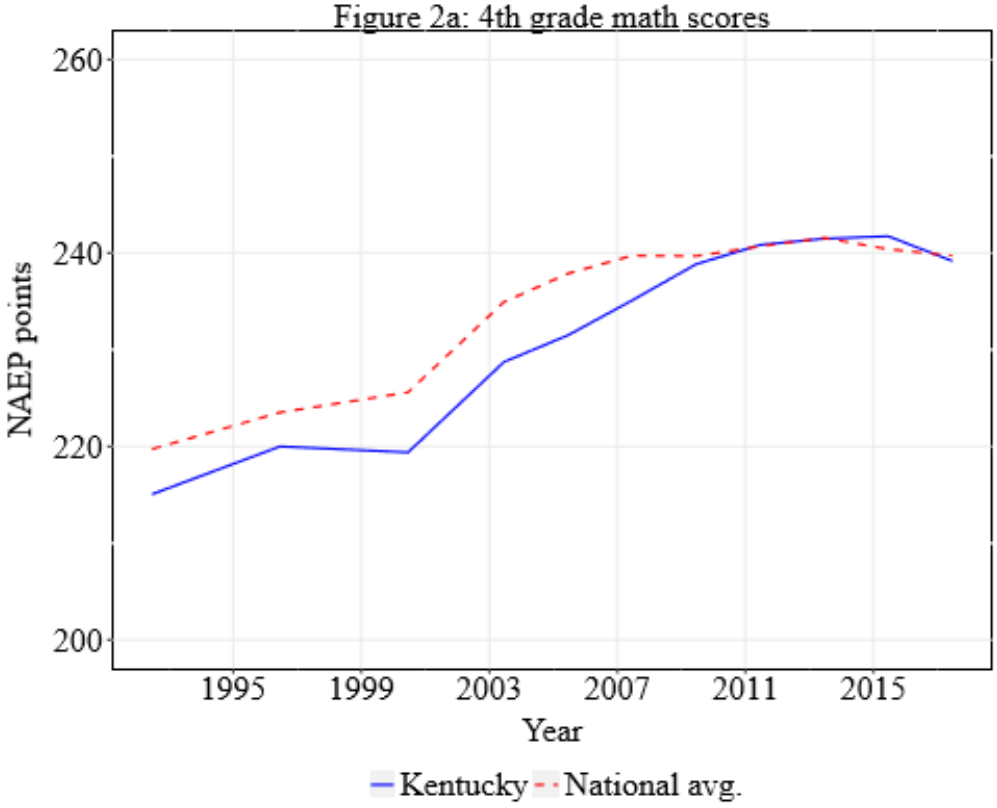
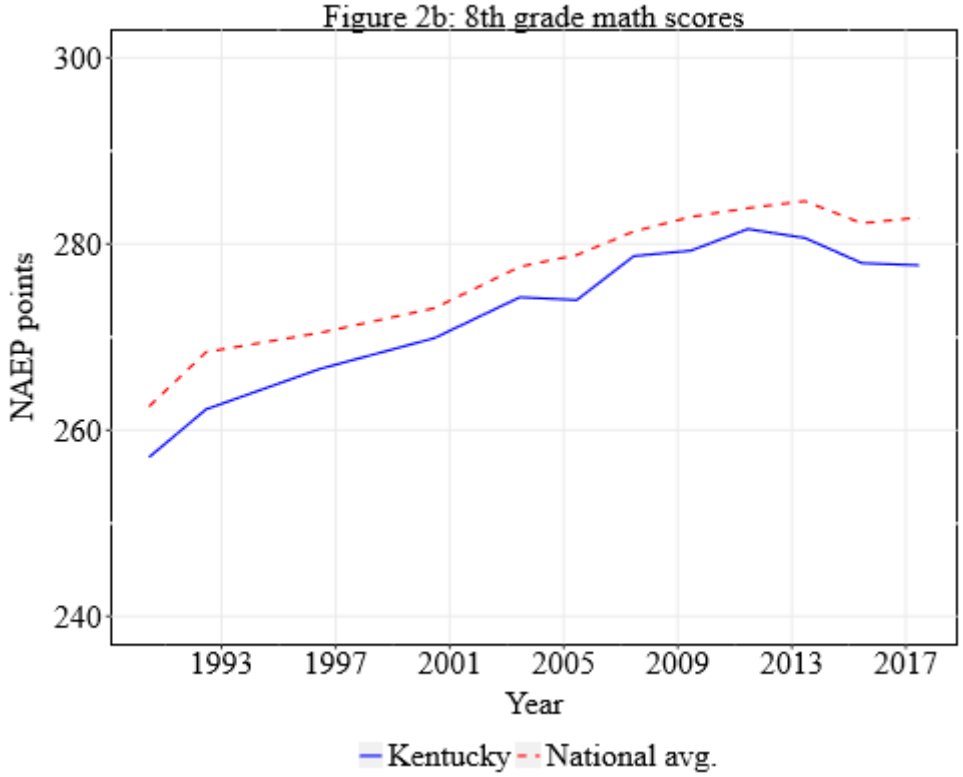


Figure 2b: 8th grade math scores



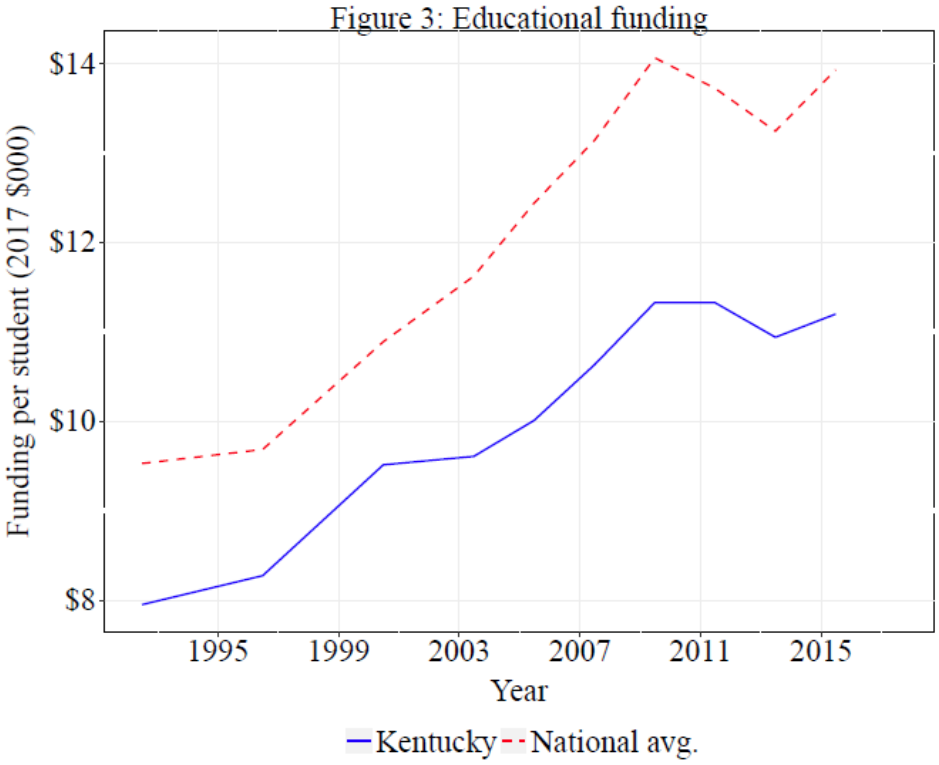
K-12 educational system. Using regression analysis, we examined all states' test score outcomes as dependent on the demographics of the student test takers. We then adjust the test scores to simulate scores as if no demographic changes had occurred. We find very little difference in the pattern of Kentucky adjusted scores relative to the raw scores. More details are in Appendix B.

Per Pupil Funding

Using data on school funding and enrollment from the National Center for Educational Statistics (NCES), we compute both total school revenue (inclusive of state, local, and federal) and total school expenditures, adjust for inflation, and express on a per pupil basis. We do this for every year from 1991 to 2015, the last year that NCES data are available. Revenue and expenditures follow one another very closely, so we present the plot for school revenue in Figure 3. As before, the solid blue line represents Kentucky and the dashed red line is the U.S.

School funding for Kentucky and the U.S., with some exceptions, follow a very similar pattern. Both increased steadily and substantially until 2010 in the aftermath of the recent recession. Per pupil funding (adjusted for inflation) fell for a few years, then resumed its rise. For Kentucky, funding rose from $7,722 per pupil in 1991 to $11,196 in 2015. This translates to a 45% increase. For the U.S. as a whole, funding rose from $9,491 per pupil in 1991 to $13,924 in 2015, a 46.7% increase. The level of Kentucky funding has been below that for the nation each year, though the gap narrowed somewhat in the 1990s.

Figure 3: Educational funding



Relative Changes in Scores and Funding

The next set of figures presents a visual comparison of the changes in test scores and of funding by displaying them on a comparable scale. Figure 4a does so for grade 4 reading. This test was first given in 1992, so the score in Kentucky in that year is indexed to 100. Later years' test scores are expressed relative to the 1992 score. The plot in the figure is interpreted as the percentage above the 1992 score, i.e., value of 110 indicates a 10% higher score than in 1992.

Similarly, we index per pupil funding in 1992 to 100 and express later years' funding relative to its 1992 level. Each plot point is interpreted similarly as the percentage above the

Figure 4a: 4th grade reading scores and funding

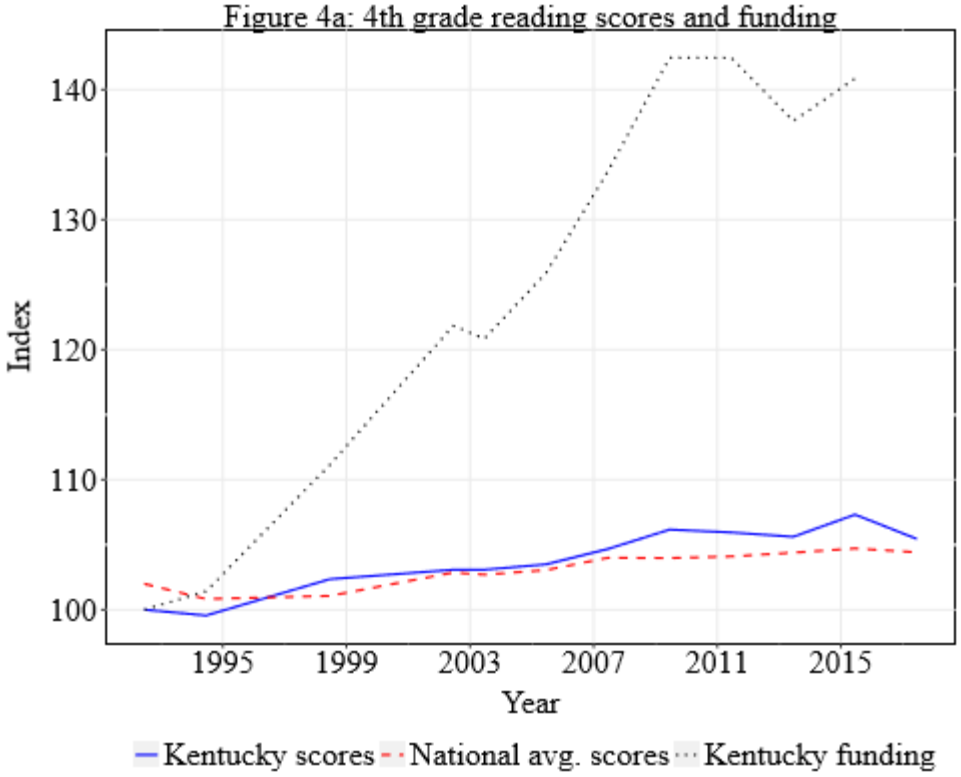
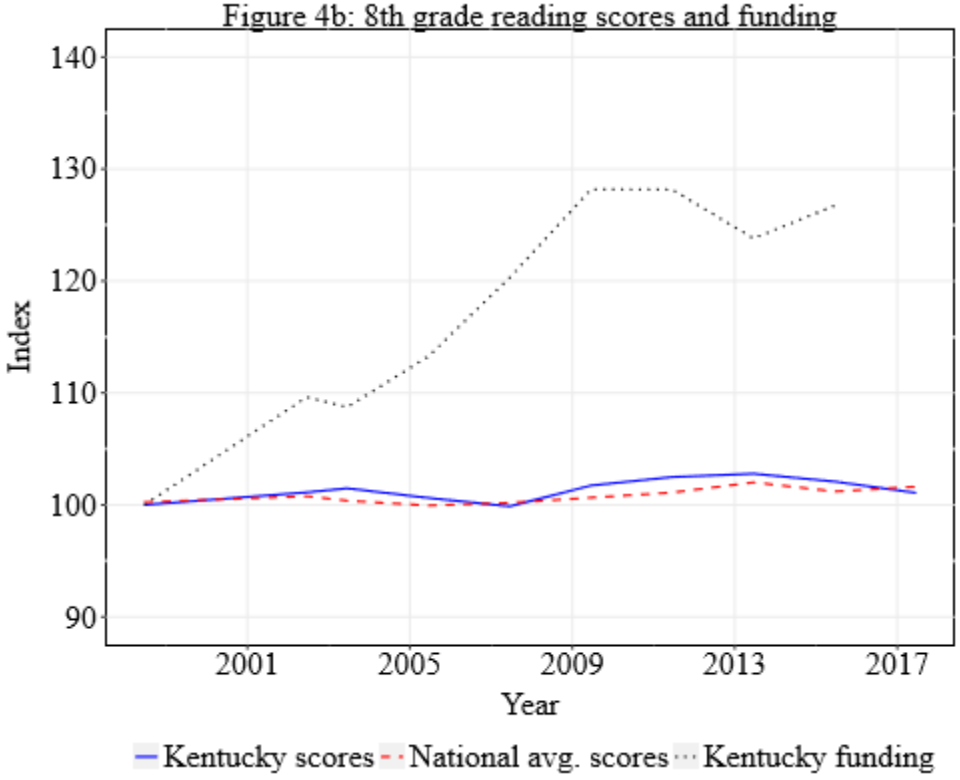


Figure 4b: 8th grade reading scores and funding



1992 level. The solid blue is Kentucky's grade 4 reading test score index, the dotted black line is Kentucky's funding index, and the dashed red line is the U.S. relative to the Kentucky 1992 score.

As is quite evident, since 1992 funding has increased very dramatically relative to grade 4 reading scores; a 41% increase in the former with a 5.2% increase in the latter. A very similar story holds for grade 8 reading as shown in Figure 4b. Values in the graph are indexed to 1998 since this is the first year that the grade 8 reading test was administered. Since 1998, this score has increased by 1.1% in Kentucky while funding rose by 26.7%.

Figures 5a and 5b display similar graphs for grade 4 and grade 8 math, respectively. As noted above, math scores have improved more than reading scores and the figures 5a and 5b reflect this. However, their increases and much smaller than the increase in funding. For grade 4 math, since 1992 scores have increased by 11.1% while funding rose by 41%. For grade 8 math, scores rose by 8.2% since 1990 while funding rose by 45%.

Figure 5a: 4th grade math scores and funding

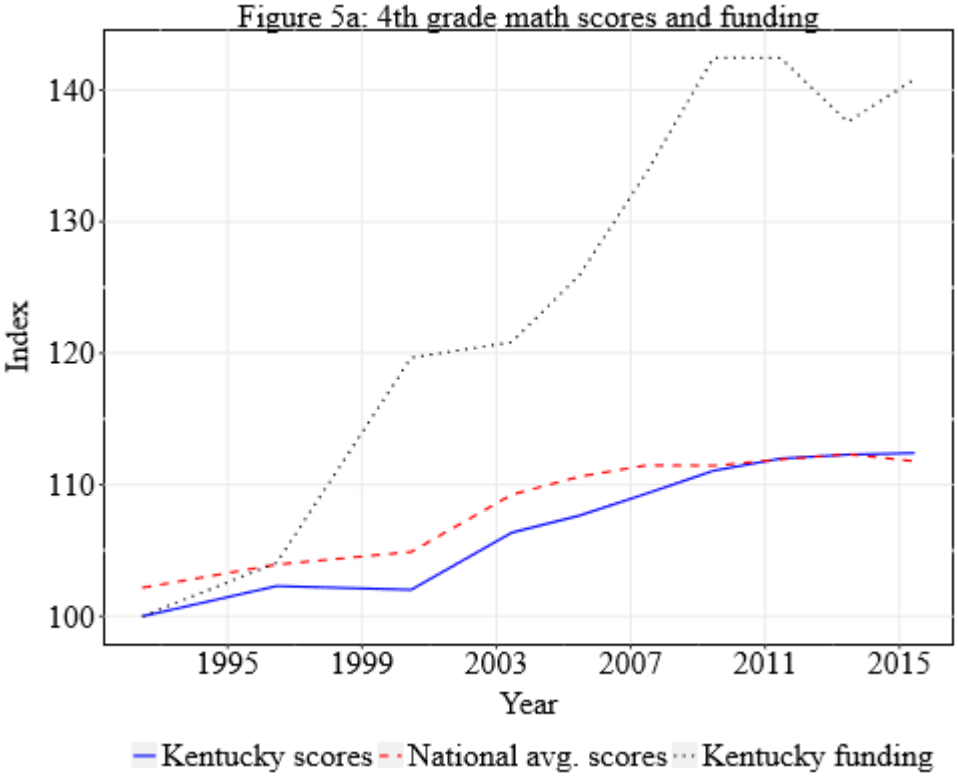
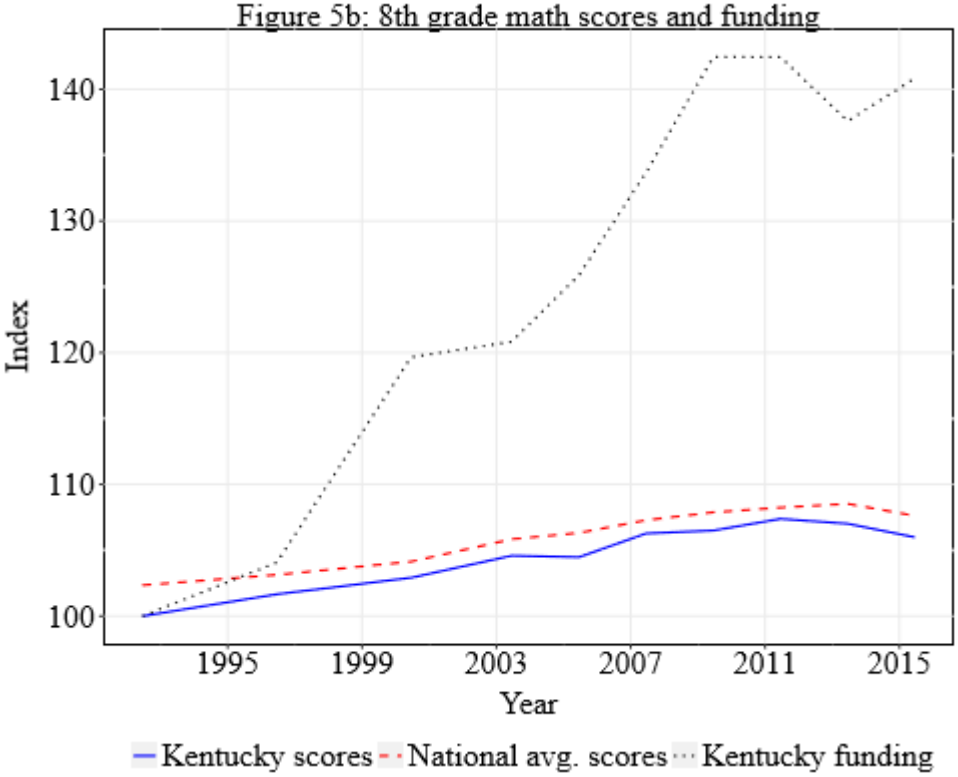


Figure 5b: 8th grade math scores and funding



Productivity

This subsection presents the data in a different way. Here, our figures display the productivity of K-12 funding, defined as NAEP test score points per $1,000 of funding. This may be coined "bang per buck," i.e., educational achievement generated (as measured by NAEP test scores) per $1,000 of taxpayer funds. Figures 6a and 6b shows productivity for grade 4 and grade 8 reading, respectively. Figures 7a and 7b display the values for grade 4 and grade 8 math. Along with the plots for Kentucky (solid blue), those for the U.S. (dashed red) are shown.

The most striking feature of all of these figures is the nearly continuous downward movement in productivity until the aftermath of the recent recession, followed by a brief upturn, then a return to the downward trend. For example, points per $1,000 of funding for grade 4 reading fell from 1992 to 2009 from 26.8 to 20; at 25% reduction. For grade 8 reading, the decline between 1998 and 2009 was from 29.6 to 23.6 (a 20% decline). The productivity of grade 4 math fell from 27 to 21.1 (a fall of 22%) from 1992 to 2009. The decline for grade 8 math from 1992 to 2009 was 33 to 24.6 (25%).

Moreover, the uptick in productivity just after 2009 is almost entirely due to the reductions in per pupil funding rather than to the any substantial increase in test scores. For example, between 2009 and 2013, the grade 4 reading test score fell by 2 points and funding fell by $387 per pupil, but points per $1,000 rose due to the relatively larger reduction in funding. For grade 4 math, scores did rise by 2 points from 2009 to 2013 (a 0.8% gain) and funding fell by 3.4%, implying that most of the 4.2% rise in productivity was due to the funding reduction, not the score improvement. Following 2013, funding per pupil increased and test scores changed little, causing a return to the pattern of declining productivity.

Figure 6a: 4th grade reading score productivity

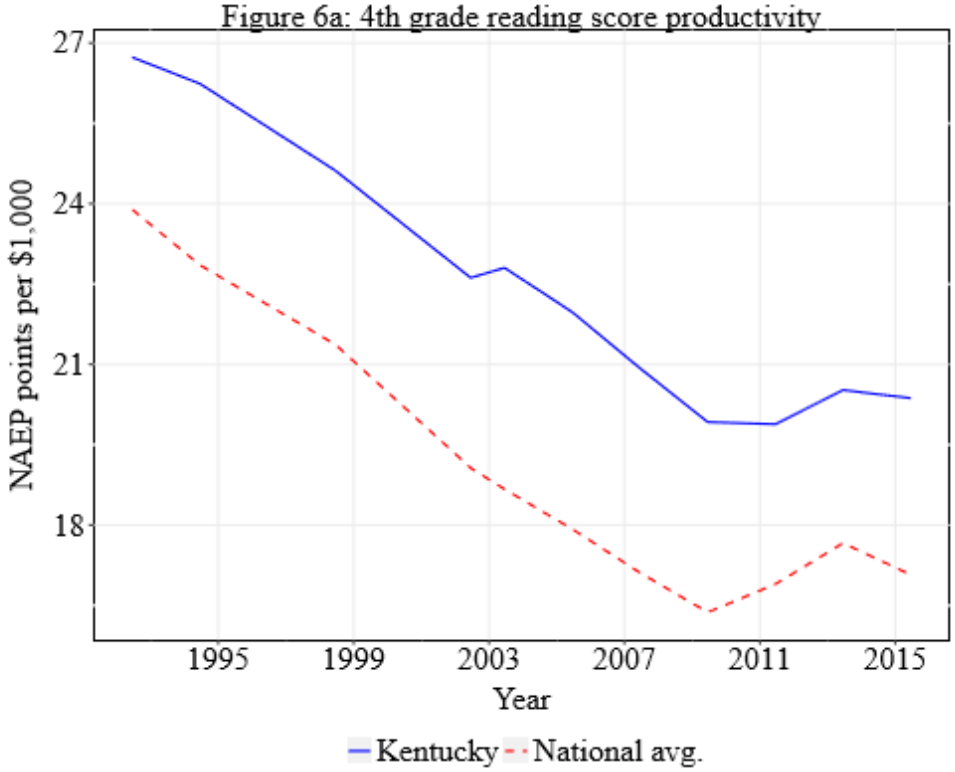


Figure 6b: 8th grade reading score productivity

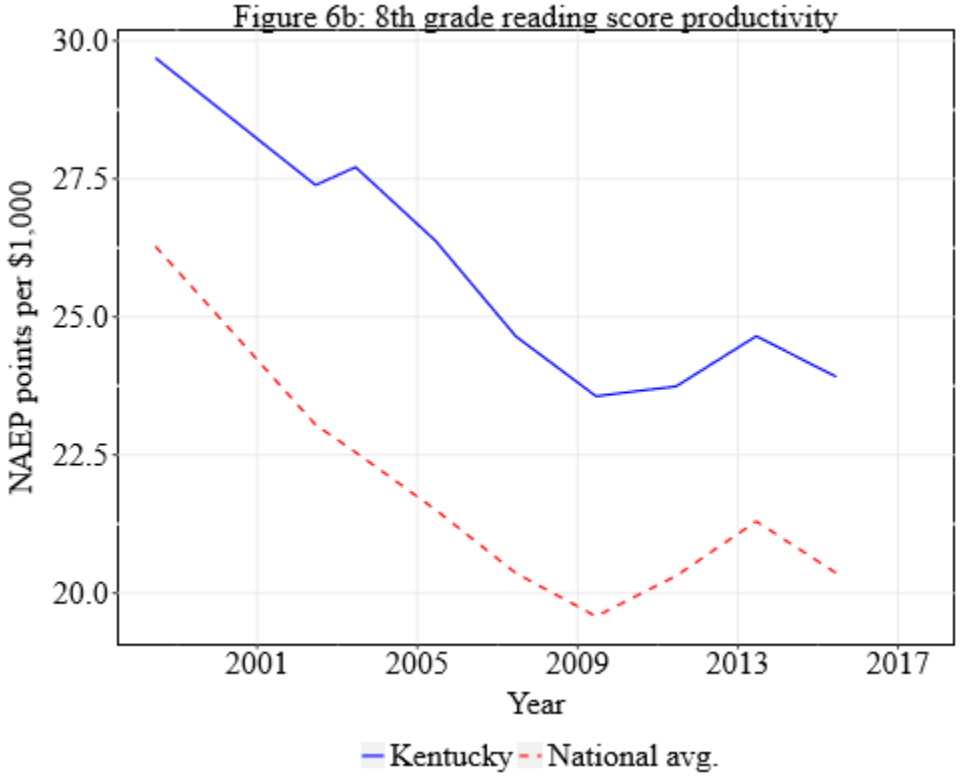


Figure 7a: 4th grade math score productivity

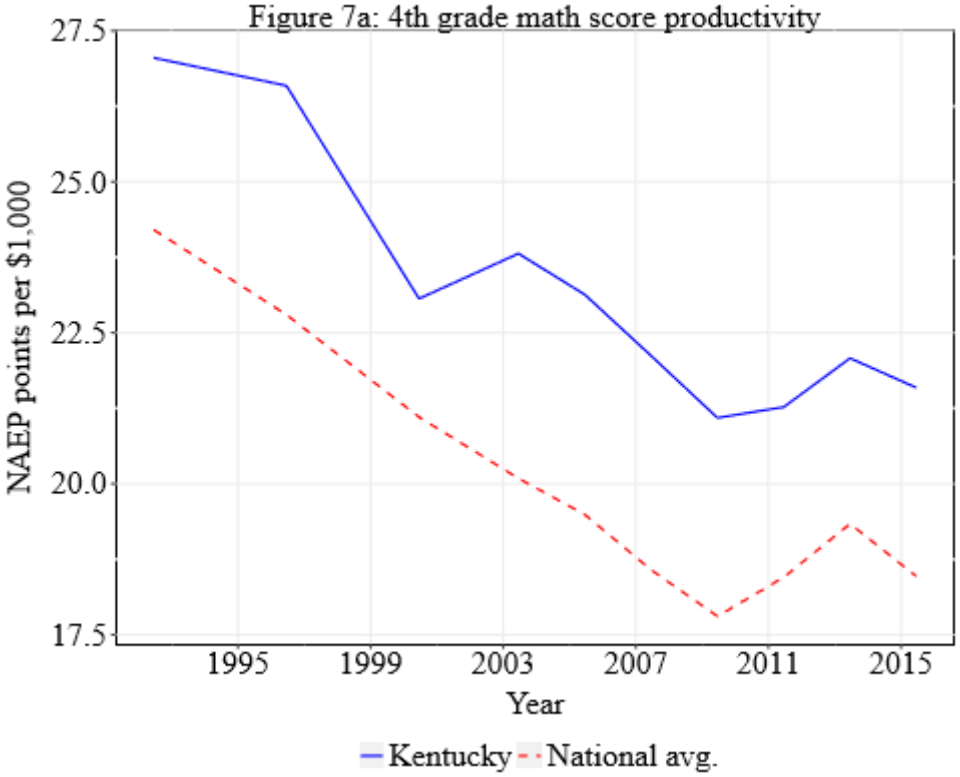
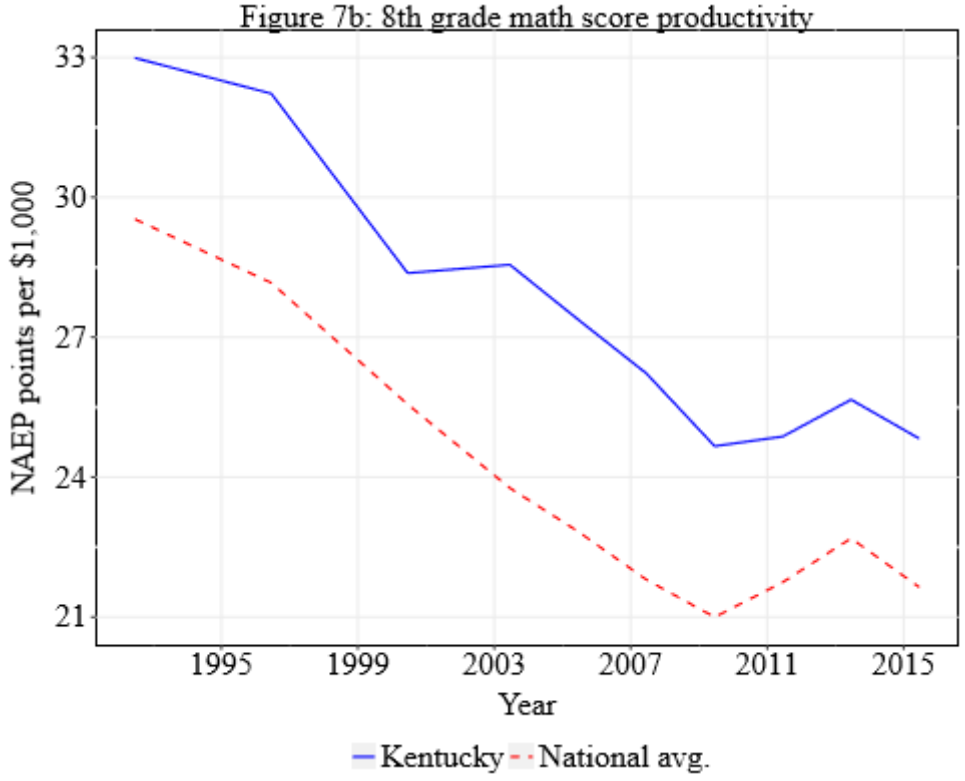


Figure 7b: 8th grade math score productivity



Comparing Kentucky productivity to that of the U.S., we see that the time trends are nearly identical; a nearly continuous decline until just after the recent recession, followed by a brief uptick then a return to the decline. However, Kentucky's productivity is consistently 3 to 4 points per $1,000 higher than that for the nation as a whole. This is due the fact the Kentucky test scores are fairly comparable to the nation's average, while funding per pupil has averaged about $2,000 less.

Examining Subgroups

The overall picture that emerges from the above examination is long term and large increases in K-12 funding, but very modest increases in tests scores. This suggests that the funding has not been very effective in improving overall NAEP scores. However, another important question that is often asked is whether the increased funding has served to improve the outcomes for disadvantaged groups and thereby reduce test score achievement differences. This section studies this question by examining trends and test score gaps for racial groups and income groups, with the latter proxied by eligibility for the national school lunch program.

Racial Trends and Differences

Figures 8a and 8b present the grade 4 and grade 8 reading scores for African American (black) students and white students.[9] Since all of the data are drawn from the National Commission of Educational Statistics (NCES), we utilize the NCES racial designation of "black" instead of "African American."

For the grade 4 reading scores, there are modest improvements over time for both blacks and whites. Likewise for whites regarding grade 8 reading but for black students, this score has changed little. Overall, the reading test score gap remains large – around 20 points – and has changed little over the past two or two-and-a-half decades, despite the 45% increase in per pupil funding during this time.

Figures 9a and 9b present the plots for grade 4 and grade 8 math. The patterns by race are similar to those for overall math scores shown in figures 2a and 2b. For both blacks and whites, there was more substantial improvements in math test scores relative to the reading scores until around 2011, followed by declines. However, there is no sign of any reduction of the black-white gap in math test scores.

[9] The data are presented in Appendix A, Table A2.

Figure 8a: 4th grade reading scores by race

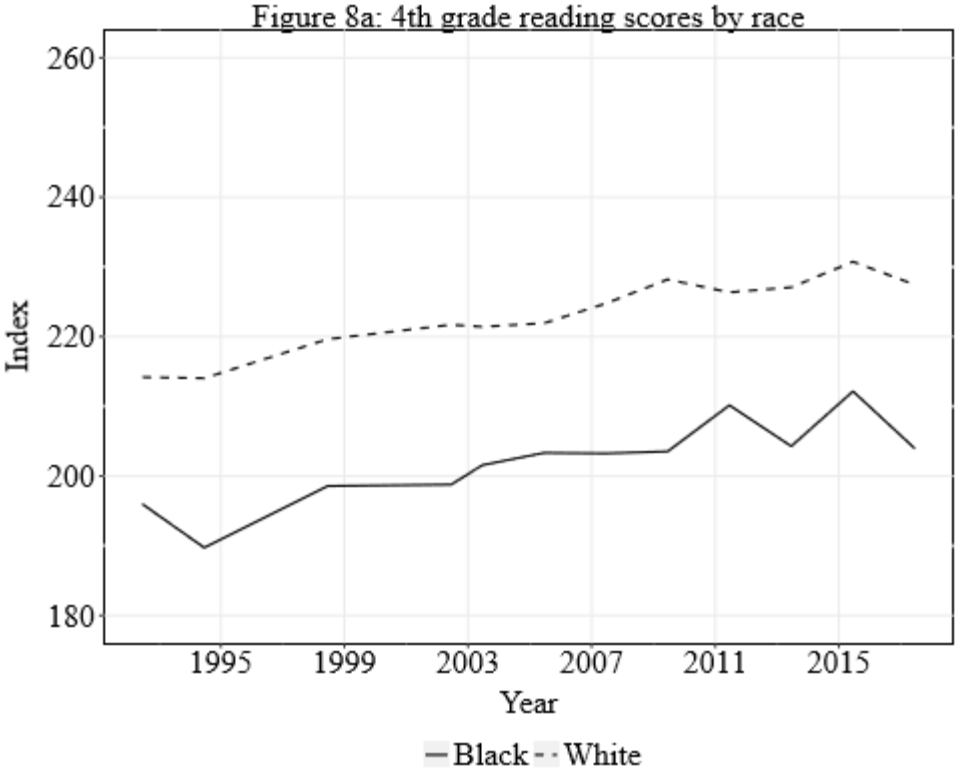


Figure 8b: 8th grade reading scores by race

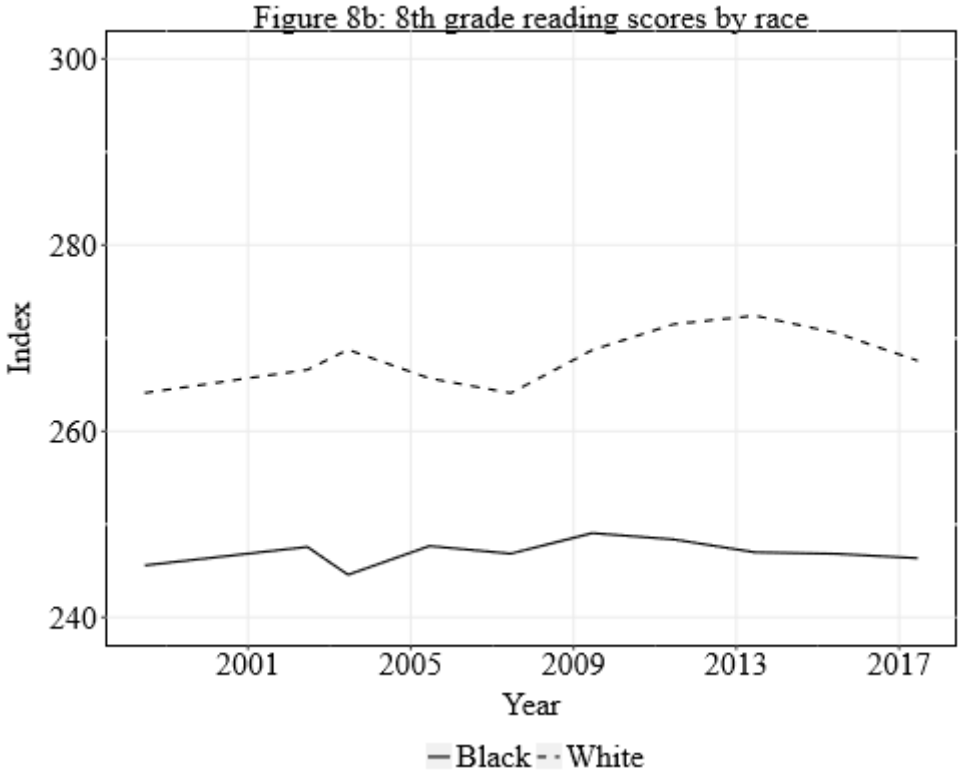


Figure 9a: 4th grade math scores by race

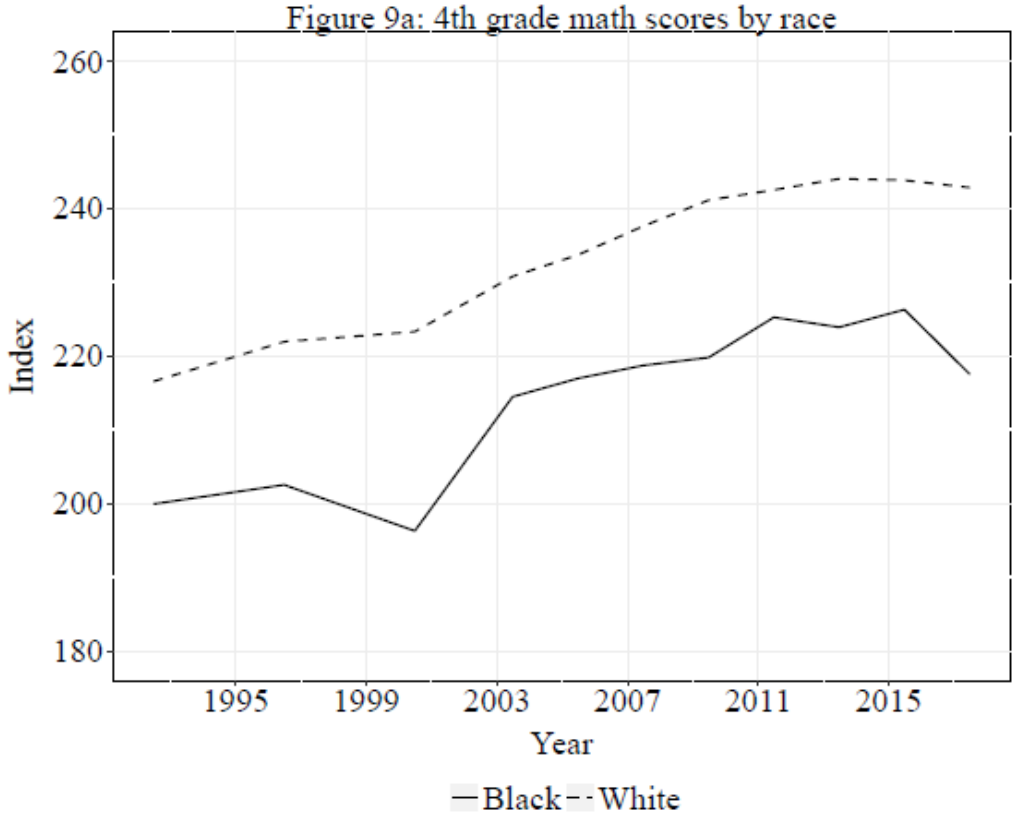
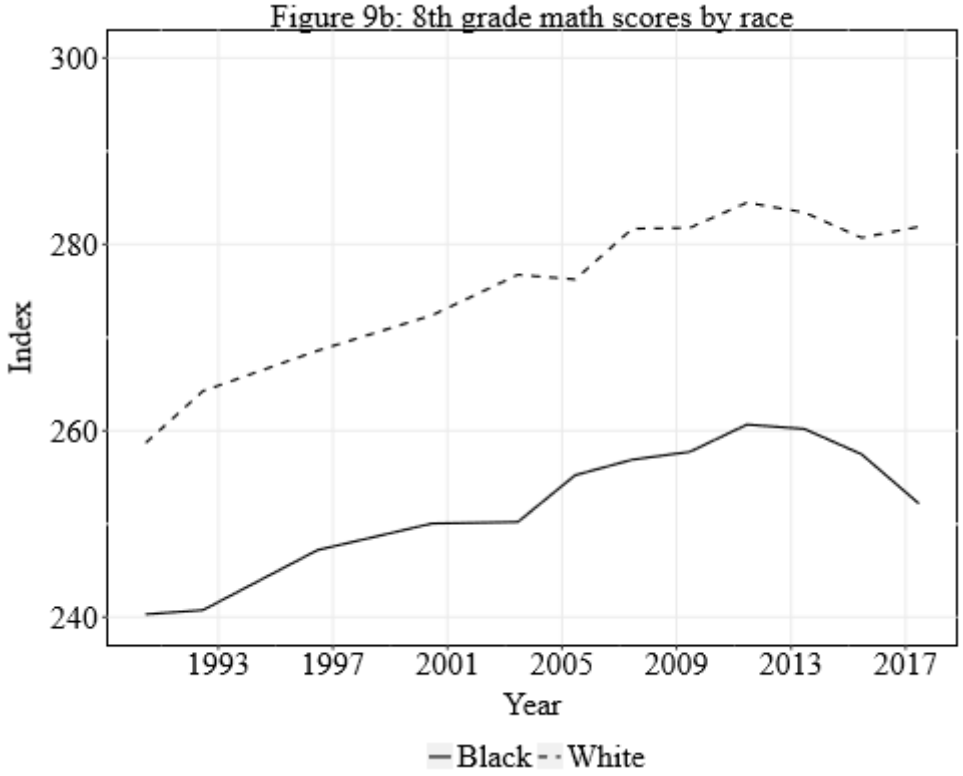


Figure 9b: 8th grade math scores by race



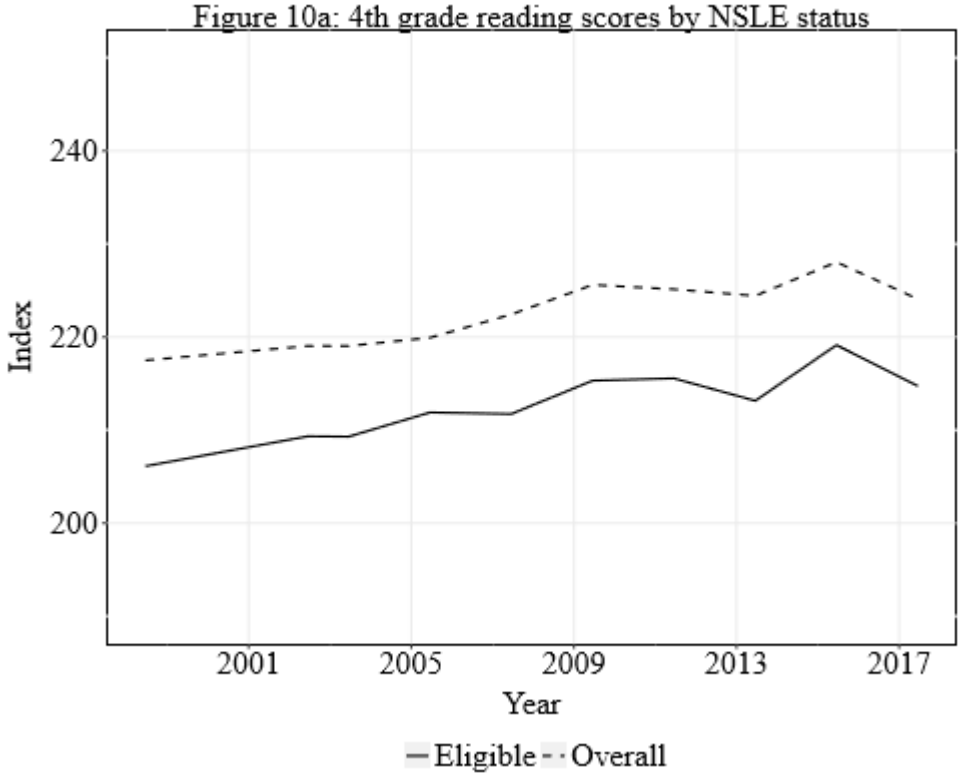
Differences by School Lunch Eligibility Status

Though the data does not include the income of the test taker's family, it does indicate if the student is eligible for the national free-or-reduced-price-lunch program, which we denote as NSLE (national school lunch eligible). Children of low-income families are eligible for this program, so it serves as a proxy for a student being from a low-income household. Thus, we examine the data to see if eligible student scores have improved relative to non-eligible students.

Figures 10a and 10b show the data plots for grade 4 and grade 8 reading.[10] These show modest improvement for both NSLE and non-NSLE students, similar to overall patterns for reading scores. There gap between eligible and non-eligible students has not changed much over the years and there remains about a 20-point gap.

Figures 11a and 11b presents the data for grade 4 and grade 8 math. As with other breakdowns of the math scores, these show more gains that the reading scores for both NSLE and non-NSLE students. Regarding the difference between non-eligible and eligible, the grade 4 math score gap has remained quite persistently at around 20 points. The gap for grade 8 math fell somewhat in the early 2000s to a little less than 20 points, but began to rise to be persistently above 20 points. As with the racial score gap, there is no indication of any reduction in the score gap by income group, as proxied by national school lunch program eligibility.

Figure 10a: 4th grade reading scores by NSLE status



[10] These data are in Appendix, Table A3.

Figure 10b: 8th grade reading scores by NSLE status

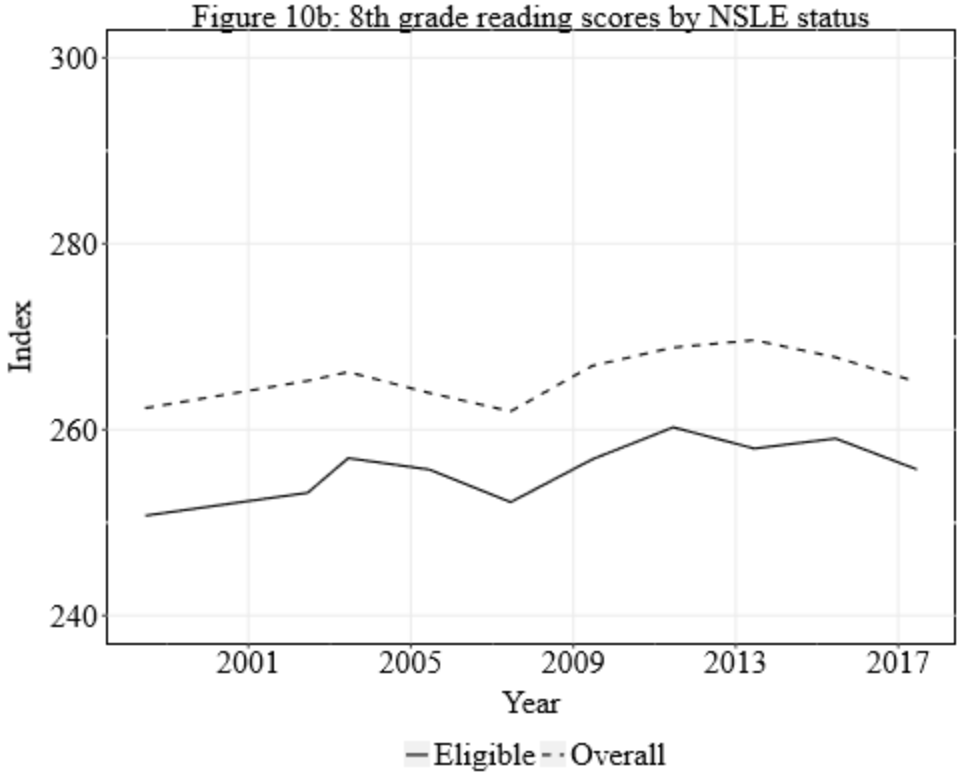


Figure 11a: 4th grade math scores by NSLE status

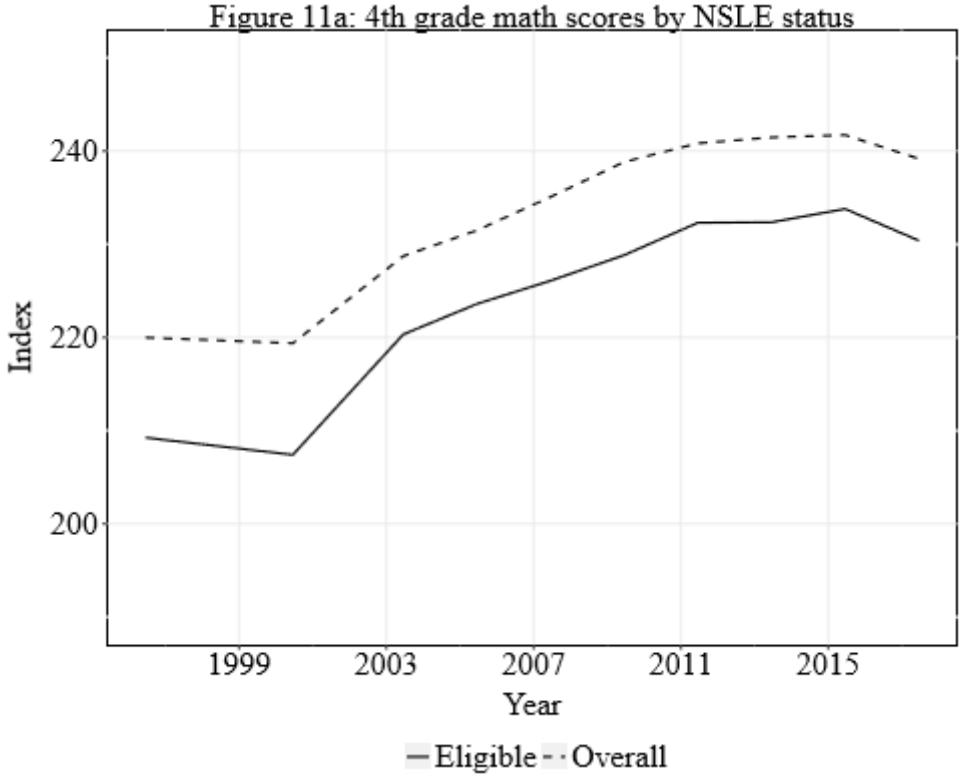
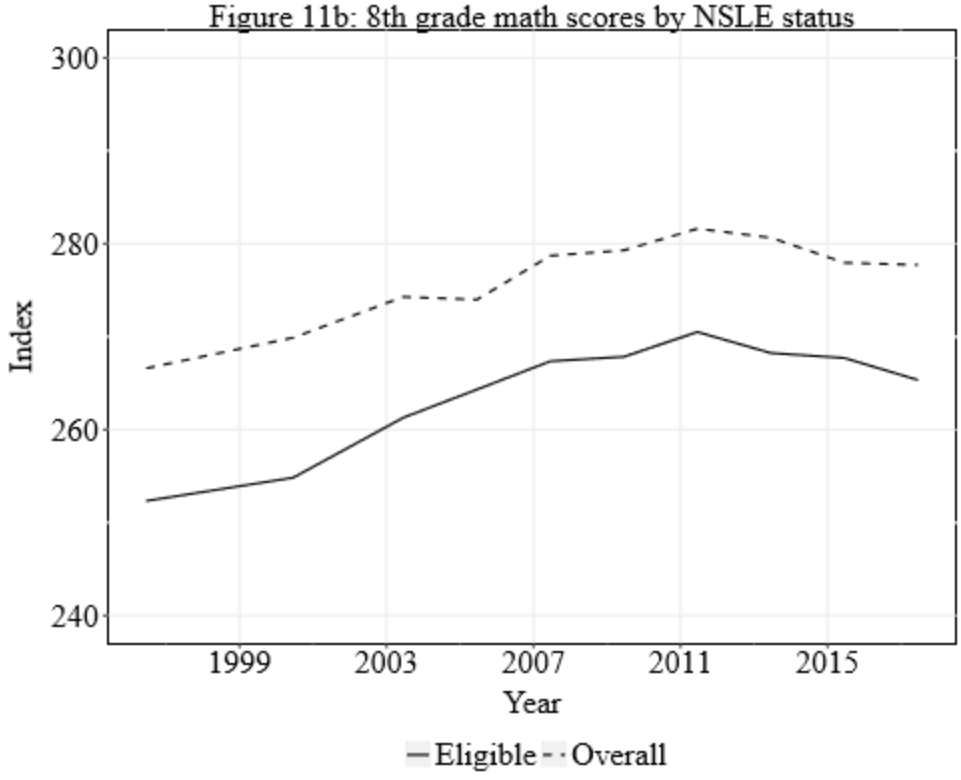


Figure 11b: 8th grade math scores by NSLE status



Conclusion

Kentucky's NAEP test score achievement has shown some improvement. However, any improvement is small relative to the 45% increase in real, per pupil funding that K-12 has received over that past two-and-a-half decades. The taxpayers' "bang per buck," i.e., NAEP achievement per $1,000 of funding, has dropped a great deal during this time period. Moreover, there has been no reduction in the test achievement gap for racial groups, nor for students from low-income families.

NAEP test score success is not the only thing that an educational system ought to strive for. Nevertheless, the cognitive skills as measured by the NAEP test scores are important aspects of educational achievement. Kentucky's modest improvements in this regard during a time period of substantial increases in K-12 funding suggests that concerns continue about the performance of the Commonwealth's public education system.

About the Authors

John Garen, PhD, Senior Fellow, Pegasus Institute

John Garen is the BB&T Professor of Economics in the Gatton College of Business and Economics at the University of Kentucky and is the founding director and an affiliate of the Institute for the Study of Free Enterprise. He has been a member of the Gatton faculty since 1985.

Dr. Garen received his Ph.D. from Ohio State University in 1982 and has served as a Visiting Professor at the University of Chicago, a Visiting Scholar at the Mercatus Center, at National Sun Yat-Sen University, and at National Taiwan Normal University, and is a member of the Mercatus Center's Faculty Network and of the Board of Scholars for the Bluegrass Institute for Public Policy Solutions. From 2005-2009 he served as Department Chair and during 2004-2005 he was Co-Director of the Gatton College's Center for Business and Economic Research.

Dr. Garen's interests have focused on the economics of organizations, labor and human resource economics, and the role of government in society. He has conducted research on a wide variety of economic issues, leading to over thirty-five publications in many of the foremost academic journals, as well as numerous reports and manuscripts. Based on this experience, Dr. Garen has generated a flow of opinion columns, media work, and presentations to the public on the economy and economic issues, and on the importance of economics to good public policy.

Steven Gordon, PhD, Associate Fellow, Pegasus Institute

Steven Gordon has conducted numerous empirical studies on a wide range of topics, including estimating the returns to digital advertising, analyzing the impact of demographic trends on urban economic growth, and using public data sources such as GoogleTrends to forecast consumer demand for products in particular locations. In addition, he has authored policy pieces and academic articles.

The primary focus of his work is to deliver value to governments, businesses, and organizations through the application of logically sound statistical techniques to answer data-driven questions. The hallmark of his methodology is a careful, scientifically rigorous approach that draws on sophisticated econometrics and machine learning algorithms, creative intuition, and a thorough institutional knowledge of a variety of specific contexts.

In addition to his research and consulting work, Steven teaches economics as an adjunct professor for multiple institutions and serves on the board of the Bluegrass Fellows Program.

Steven holds a Bachelor's degree from Virginia Tech, a Master's degree from George Mason University, and a Ph.D. in Economics from the University of Kentucky. He lives in Lexington, Kentucky.

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.

Currie, Janet and Thomas, Duncan, "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 and Bray, Rex, "Educational Test Scores, Education Spending, and Productivity in Public Education: National Trends and Evidence Across States and Over Time, 1990 – 2015," John H. Schnatter Institute for the Study of Free Enterprise, University of Kentucky, January 2018.

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.

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

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,* electronic data at https://nces.ed.gov/nationsreportcard/lttdata/.

Appendix A

Table A1: Raw Data on NAEP Scores and Funding

Year	Grade 4 Reading		Grade 8 Reading		Grade 4 Math		Grade 8 Math		Spending	
	KY	US	KY	US	KY	US	KY	US	KY	US
1990	-	-	-	-	-	213	257	263	-	-
1991	-	-	-	-	-	-	-	-	7722	9491
1992	213	217	-	260	215	220	262	268	7949	9527
1993	-	-	-	-	-	-	-	-	7955	9564
1994	212	214	-	260	-	-	-	-	8062	9560
1995	-	-	-	-	-	-	-	-	7928	9674
1996	-	-	-	-	220	224	267	270	8273	9683
1997	-	-	-	-	-	-	-	-	8831	9858
1998	218	215	262	263	-	-	-	-	8835	10215
1999	-	-	-	-	-	-	-	-	9450	10637
2000	-	213	-	-	219	226	270	273	9512	10887
2001	-	-	-	-	-	-	-	-	9378	11346
2002	219	219	265	264	-	-	-	-	9685	11666
2003	219	218	266	263	229	235	274	278	9606	11618
2004	-	-	-	-	-	-	-	-	9954	12071
2005	220	219	264	262	231	238	274	279	10008	12435
2006	-	-	-	-	-	-	-	-	10572	12648
2007	222	221	262	263	235	240	279	281	10628	13135
2008	-	-	-	-	-	-	-	-	11216	13591
2009	226	221	267	264	239	240	279	283	11324	14059
2010	-	-	-	-	-	-	-	-	11361	14151
2011	225	221	269	265	241	241	282	284	11323	13718
2012	-	-	-	-	-	-	-	-	11095	13416
2013	224	222	270	268	241	242	281	285	10937	13240
2014	-	-	-	-	-	-	-	-	10912	13341
2015	228	223	268	265	242	240	278	282	11196	13924
2017	224	222	265	267	239	240	278	283	-	-

Table A2: Race Differences for NAEP Scores

Year	Grade 4 Reading			Grade 8 Reading			Grade 4 Math			Grade 8 Math		
	White	Black	Diff.	White	Black	Diff.	White	Black	Diff.	White	Black	Diff.
1990	-	-	-	-	-	-	-	-	-	259	240	18
1992	214	196	18	-	-	-	217	200	17	264	241	24
1994	214	190	24	-	-	-	-	-	-	-	-	-
1996	-	-	-	-	-	-	222	203	19	269	247	21
1998	220	199	21	264	246	19	-	-	-	-	-	-
2000	-	-	-	-	-	-	223	196	27	272	250	22
2002	222	199	23	267	248	19	-	-	-	-	-	-
2003	221	202	20	269	245	24	231	214	16	277	250	27
2005	222	203	19	266	248	18	234	217	17	276	255	21
2007	225	203	21	264	247	17	238	219	19	282	257	25
2009	228	204	25	269	249	20	241	220	21	282	258	24
2011	226	210	16	271	248	23	243	225	17	284	261	24
2013	227	204	23	272	247	25	244	224	20	283	260	23
2015	231	212	19	271	247	24	244	226	18	281	257	23
2017	227	204	24	268	246	21	243	218	25	282	252	30

Table A3: Score Differences for NSLE

Year	Grade 4 Reading			Grade 8 Reading			Grade 4 Math			Grade 8 Math		
	Eligible	Not eligible	Diff.	Eligible	Not eligible	Diff.	Eligible	Not eligible	Diff.	Eligible	Not eligible	Diff.
1996	-	-	-	-	-	-	209	230	-21	252	276	-23
1998	206	227	-21	251	270	-20	-	-	-	-	-	-
2000	-	-	-	-	-	-	207	230	-23	255	280	-26
2002	209	229	-20	253	273	-20	-	-	-	-	-	-
2003	209	229	-20	257	273	-16	220	237	-17	261	284	-23
2005	212	228	-16	256	271	-15	224	240	-17	264	283	-18
2007	212	234	-22	252	271	-19	226	245	-19	267	288	-21
2009	215	236	-21	257	276	-19	229	249	-21	268	290	-22
2011	216	236	-21	260	278	-18	232	251	-19	271	294	-23
2013	213	237	-23	258	281	-23	232	251	-19	268	293	-24
2015	219	241	-22	259	279	-20	234	255	-21	268	291	-23
2017	215	237	-22	256	278	-23	230	252	-21	265	293	-28

Appendix B

This appendix explains the regression analysis that controls for differences and changes in student demographics and allows us to adjust the test scores for these differences. The first step is to estimate a regression equation of the following form:

$$T_{jt} = b_0 + b_1X_{jt} + b_2S_{jt} + b_3\,\theta_j \cdot S_{jt} + \theta_j + \delta_t + \varepsilon_{jt}$$

In this equation, 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), X_{jt} = a vector of demographics for state j at time t, S_{jt} = state j's real per pupil funding 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.

The demographic variables include the percent of students eligible for free or reduced priced lunch, the percent of students with disabilities, the percent of students who are English language learners, the testing and accommodation of the former two groups, and the racial composition of the students in the state and year.

This equation is estimated by ordinary least squares. Coefficient estimates are then used to simulate Kentucky's test scores as if its demographics did not change over time. This removes the effect of these changes from the test scores. As noted in the text, these adjustments are very minor. Table B1 presents the raw and adjusted data.

Table B1: Raw and Adjusted Scores

Year	Grade 4 Reading		Grade 8 Reading		Grade 4 Math		Grade 8 Math	
	Raw	Adj.	Raw	Adj.	Raw	Adj.	Raw	Adj.
1990	-	-	-	-	-	-	257.1	254.6
1992	212.5	210.7	-	-	215	211.9	262.2	259.9
1994	211.6	208.7	-	-	-	-	-	-
1996	-	-	-	-	220	217.5	266.6	264.9
1998	217.5	216.4	262.3	262.3	-	-	-	-
2000	-	-	-	-	219.4	218.1	269.9	270.9
2002	219	218.2	265.2	263.8	-	-	-	-
2003	219	219	266.2	266.2	228.7	228.7	274.3	274.3
2005	219.9	219	263.9	262.6	231.5	230.9	274	274.3
2007	222.4	223.4	262	261.2	235.1	235.5	278.7	278.1
2009	225.6	225.8	266.9	266.8	238.8	238.4	279.3	279
2011	225.1	224.9	268.8	267.8	240.8	240.4	281.6	282.2
2013	224.4	226.7	269.6	270.6	241.5	241.3	280.6	281.5
2015	228.1	231.7	267.8	269.4	241.7	242.9	277.9	279.7
2017	224.1	-	265.1	-	239.2	-	277.7	-