Expert Report

Maisto v. State of New York

Eric A. Hanushek

11/13/2014

I have been retained by counsel for the defense to provide expert opinion in this action. The following describes my opinions and the background for offering those opinions. The conclusions in this report are directly related to the supporting charts that follow.

My testimony will provide evidence about the impact of school funding on achievement, covering both the existing body of research on the topic and the specific case of resources for New York State schools. As developed below, both the broad research base and the detailed analysis of New York indicate that how money is spent is much more important than how much is spent. Relatedly, it analyzes that impact of court ordered spending in other jurisdictions has had on student outcomes and concludes that dramatic increases in spending have not resulted in improved achievement.

My testimony will also discuss the importance of focusing on teacher effectiveness. Substantial evidence shows that teacher quality is the most important aspect of schools but that it is unrelated to the salaries of teachers. It is, however, under the control of the districts.

Finally, my testimony will discuss the role of preschool education. This testimony will indicate that there are apparent gains, particularly for disadvantaged students, from high quality preschool, but how to provide such high quality preschool effectively and efficiently is currently uncertain.

The basis for my opinions is my own research on each of the elements of the testimony and my general understanding of the state of knowledge based on my evaluation of research by others.

Background and Qualifications

I am the Paul and Jean Hanna Senior Fellow at Stanford University's Hoover Institution, as well as a Professor (by courtesy) of Education and Economics at Stanford University. I am the Chairman of the Executive Board of the Texas Schools Project at the University of Texas at Dallas, a Senior Research Fellow at the University of Texas at Dallas, and a Member of the Koret Task Force on K-12 Education at the Hoover Institution. I am a Research Associate at the National Bureau of Economic Research, and a Member of the Management Team of the Center for Analysis of Longitudinal Data in Education Research ("CALDER"). I am a research professor at the ifo Institute for Economic Research of the University of Munich, a research fellow of the Institute for the Study of Labor (IZA), and the area coordinator for the economics of education of the CESifo Network.

Prior to joining the Hoover Institution, I was a Professor of Economics and Political Science at the University of Rochester from 1978 to 2000, Associate Professor of Economics at Yale University from 1975-1978, and Associate Professor of Economics (and other ranks) at the U.S. Air Force Academy from 1968-1973. I also served in the U.S. Air Force from 1965-74.

I am a distinguished graduate of the U.S. Air Force Academy where I received a B.S. in Economics. I received a Ph.D. in Economics from the Massachusetts Institute of Technology.

I am a member of several learned societies, such as the American Economic Association, the American Education Finance Association, the Association for Public Policy Analysis and Management, the Econometric Society, the Society of Labor Economists, the American Educational Research Association,

and the International Institute of Public Finance. I am a Fellow of the International Academy of Education, the Society of Labor Economists, and the American Educational Research Association, and a Member of the National Academy of Education. I received the Fordham Prize for Excellence in Education in 2004.

I have significant experience working with federal and state governments on education policy. From 2011 to 2013, I was a Commissioner of the Equity and Excellence Commission with the U.S. Department of Education. From 2004 to 2010, I was a Member of the Board of Directors of the National Board for Education Sciences, and was its chair from 2009 to 2010. From 2008 to 2010, I was a Member of the Council of Economic Advisors in California. From 2005 to 2008, I was a Member of the Governor's Committee on Education Excellence (California). And in 2007, I was a Member of the Governor's Commission for a College Ready Texas. I was the deputy director of the Congressional Budget Office from 1983-1985 and Senior Staff Economist at the President's Council of Economic Advisers from 1971-1972.

I also have editorial responsibilities at several publications. I am on the Editorial Board of *Education Finance and Policy*, *Education Next*, and *Economics of Education Review*. I am a Co-Editor for *CESifo Economic Studies.* In addition, I am an Associate Editor of *Journal of Human Capital*, and an Advisory Editor for *Social Science Research*. And I am on the Advisory Board of the National Council on Teacher Quality.

I have published more than twenty books, including four volumes of the *Handbook of the Economics of Education*. My most recent book entitled *Endangering Prosperity: A Global View of the American School*, which I co-authored with Paul Peterson and Ludger Woessmann, documents the economic costs of continuing to have mediocre public schools in the United States. Along with Alfred Lindseth, I published an extensive review and analysis of school finance litigation, *Schoolhouses, Courthouses, and Statehouses: Solving the Funding-Achievement Puzzle in America's Public Schools.* I edited a separate book of analyses of various aspects of school finance litigation: *Courting Failure: How School Finance Lawsuits Exploit Judges' Good Intentions and Harm our Children.*

I have also published over two hundred highly cited articles in the education field. My works cover a broad range of topics including the measurement and impact of teacher quality and the impact of high-stakes accountability on student achievement. My publications also include extensive analysis delving into the importance of school resources, teacher salaries, and class size for achievement. I have also published a variety of articles on the economic returns to high quality education.

I have previously testified as an expert witness for both plaintiffs and defendants in over twenty court cases about school finance, school desegregation, and teacher employment laws. This testimony has covered most of the education topics in which I have researched and published.

Attached as Exhibit A is a true and correct copy of my current Curriculum Vitae.

Spending and Achievement

An enormous amount of scientific analysis has focused on how spending and resources of schools relates to student outcomes. It is now commonly believed that spending on schools is not

systematically related to student outcomes. For example, over fifteen years ago the National Research Council was asked to address issues of productivity within the system of funding schools. Its summary statement at that time was "money can and must be made to matter more than in the past if the nation is to reach its ambitious goal of improving achievement for all students." [1] And, in a recent assessment of school finance debates, Norton Grubb says simply, "School finance as a field of study and practice – what I'll call the conventional or 'old' school finance – is concerned almost entirely with revenues and expenditures rather than with broader conceptions of resources or the effects of resources on schooling outcomes."[2] Further, "most analyses of school resources still concentrate on dollars spent rather than on how resources are used." Systematic review of available studies also reaches this conclusion.[3]

The overall truth of this disconnect of spending and outcomes is easiest to see by looking at the aggregate data for the United States over the past half century. Since 1960, pupil-teacher ratios fell by one-third, teachers with master's degrees over doubled, and median teacher experience grew significantly (Chart 1).[4] Since these three factors are the most important determinants of spending per pupil, it leads to the quadrupling of spending between 1960 and 2009 (after adjusting for inflation). At the same time, plotting scores for math and reading performance of 17-year-olds on the National Assessment of Educational Progress (NAEP, or "The Nation's Report Card") shows virtually no change since 1970 (Charts 2 and 3).[5]

There has been substantial econometric evidence that supports this lack of relationship. [6] This analysis has been conducted for a half century since the major government study of the "Coleman Report".[7] In this, statistical methods, generally some form of regression analysis, are used to separate the impact of school factors on student achievement from the impacts of families, neighborhoods, and other factors on achievement.[8] Overall, when the results of the many studies are combined, there is

[1] Ladd and Hansen (1999), p. 1.
[2] Grubb (2009), p. 26.
[3] For example, Betts (1995); See the review in Hanushek (2006).
[4] Note that the chart number appears in the bottom right hand corner of each chart. Data on spending and resources is found in U.S. Department of Education (2014) and various prior volumes.
[5] http://nationsreportcard.gov/
[6] Hanushek (2003). See also Hanushek (1981, (1986, (1989). The statistical analyses focus on the independent impact of resources on performance after allowing for differences among families, peers, and neighborhoods. A variety of sophisticated approaches have been applied to schooling situations across the countries, and the reviews summarize these studies. The aggregate results of the most sophisticated of these studies are shown below.
[7] Coleman et al. (1966). The "Coleman Report" was a response to a Congressional mandate in the Civil Rights Act of 1964. It was the first major study that attempted to identify the determinants of achievement differences across students. It used statistical methods (analysis of variance) to assess the importance of the various inputs into achievement. While heavily criticized on methodological grounds, it began the large research stream that is discussed here. For criticisms, see Bowles and Levin (1968) and Hanushek and Kain (1972).
[8] Regression analysis is designed to uncover the independent effects of several factors that affect achievement. The explanatory factors affecting student achievement tend to be correlated; for example, parental incomes and education are correlated with spending and other characteristics of schools. Looking at each factor individually will make it difficult to tell which is more important or what the effect of each is because any simple correlation will implicitly incorporate all of the other (correlated) factors.

little reason to believe with any degree of certainty that just providing more funds to schools will lead to higher student achievement.[9]

While there have been hundreds of studies showing the basic facts, it is clearest from the best econometric studies.[10] These studies are based on value-added models of individual student achievement. Value-added models are designed to isolate the influence of schools and teachers after statistically adjusting for differences in prior academic history and achievement, parental background, and in other factors outside the control of teachers and schools that affect achievement. Simply put, it would be inappropriate to judge a teacher who had a class of under-prepared students as ineffective just because she could not lift students to high levels. Similarly, it would be inappropriate to judge a teacher who had a class of well-prepared students as effective just because she maintained the students to above average levels. In essence, value-added models for a teacher are designed to focus on the *growth* in student achievement that is related to the teacher and the resources relevant to the individual classroom. By looking at achievement growth (along with other factors such as language difficulties, poverty, and special education), value-added models isolate the contribution of teachers and schools after allowing for different entering achievement levels.

While teacher effectiveness is discussed in detail below, it is useful to begin with results about overall school resources. These value-added studies show that most estimates of the effects of teacher-pupil ratios, teacher education levels, and teacher experience are either statistically indistinguishable from zero or significantly negative (Chart 4). The studies of teacher experience that find positive and significant effects generally reflect the fact that there is typically improvement in teacher effectiveness during the first year or two of experience, but these experience improvements are not seen after the initial gains.[11]

Over time, there have been a number of academic controversies over the impact of resources on achievement.[12] Virtually all analysts now realize that *how* money is spent is much more important than *how much* is spent. In the 2005 Yearbook of the American Education Finance Association, the editors summarize this as "school performance and efficiency measurement have taken center stage in much of the debate and research in education policy since at least the mid-1990s."[13]This finding is particularly true at the upper levels of current U.S. spending. It also underscores how calculations of equity gaps in spending, of costs needed to achieve equity, or of costs needed to obtain some level of student performance are vacuous, lacking any scientific basis.

At times, this debate has been disingenuously labeled as a debate about 'does money matter?' with an attempt to trivialize the findings. Of course, money matters. It is not possible to run schools without buildings, teachers, textbooks, and other factors. The issue, however, is whether additional

[9] When there are many studies addressing the same issues, it is necessary to have a methodology for aggregating the results (which differ across studies for a variety of reasons). This combining of the results of different studies is frequently called "meta-analysis." See, for example, Hedges and Olkin (1985).
[10] Hanushek (2003)
[11] Rivkin, Hanushek, and Kain (2005); Hanushek and Rivkin (2012).
[12] See, for example, Burtless (1996).
[13] Stiefel et al. (2005), p. 1.

resources systematically lead to improved student achievement. Further, the findings do not say that added resources never matter or that added resources cannot matter. The findings simply say that, given current incentives and spending behavior, there is no consistent improvement in achievement with added resources.

The debates have also at times focused on specific resource issues instead of just aggregate spending. Perhaps the most long lasting focus has been on class size. The results about achievement effects for class size differences have been consistent with the lack of impact in the aggregate. Nonetheless, these findings have been very contentious, with some people relying heavily on the STAR experiment from Tennessee in the mid-1980s.[14] The STAR experiment was not a very good experiment from a scientific perspective, but even if taken at face value does not support the large reductions in class size that have occurred over the past two decades. The results from the STAR experiment show that there appeared to be some small improvements during kindergarten, but the small classes did not continue to have an impact on achievement in the later grades of the experiment – even though that would have been expected if small classes had an impact across grades (Chart 5). Moreover, the reductions in STAR were very large (moving from 23 to 15 students per class), making this an extraordinarily expensive policy. No comparable studies even exist for later grades. These modest results must be balanced by the large number of studies showing no effect from smaller classes and by the expense of pursuing the policies.

The studies of school resources also make it clear why there is so little consistent impact of just spending more. The major determinants of spending per pupil are teacher salaries and pupil-teacher ratios – but these are unrelated to teacher effectiveness and student achievement. Thus, if added spending simply goes toward teacher salaries and class size, there is no reason to expect improved achievement. This is the story of the large increases in spending documented in Chart 1.[15]

Aggregate New York Expenditure and Performance

The New York legislature currently provides resources for schools that exceed those in every jurisdiction except the District of Columbia. New York State spending in 2011 was $19,708 per pupil (in 2013 dollars), while the national average was $11,153 – making New York spending 77 percent higher than the rest of the nation.[16] Moreover, New York spending per pupil in 2011 was 46 percent higher (adjusted for inflation) than spending in 2000. This spending increase in New York was fifth highest over the decade, and dramatically above the national growth in spending of 18 percent. Only Wyoming, the District of Columbia, New Hampshire, and North Dakota had more rapid increases in spending per pupil.

Chart 6 provides comparisons of state spending and shows that New Jersey and Wyoming – two states with extensive court intervention in appropriation decisions –remain below New York in

[14] See Word et al. (1990). For part of the controversy, see Mishel and Rothstein (2002) and Krueger (1999).
[15] See also Hanushek and Rivkin (1997).
[16] U.S. Department of Education (2014), table 236.65.

spending. Wyoming had a significant increase in spending after its court decision in *Campbell v. Wyoming*.[17] New Jersey has had continuing expenditure increases under the rulings in *Abbott v. Burke*.[18]

In short New York State has aggressively increased spending on schools.

Charts 7-10 provide comparative achievement of New York students and of those in other states. Student performance data come from the National Assessment of Educational Progress (NAEP).[19] NAEP regularly assesses the performance of a representative sample of students in each state. These charts show the relative performance of New York State students in fourth and eighth grade reading and math. Except for fourth grade reading, NY students consistently performed below the national average in 2013. As is clear from Chart 6, all states outperforming New York on these assessments spent less per pupil on their schools.

While New York has a more challenging student population than some other states, Charts 11-14 show that low income students still do not do relatively well when compared to other states. The charts compare performance of students eligible for free or reduced price lunches across states. Fourth grade reading performance is somewhat better than eighth grade reading or than math performance, but neither compare favorably to the spending shown in Chart 6. Except for fourth grade reading, New York performance is below the median for states.

Chart 15 indicates that New York does better in college readiness, although it still lags behind states with lower spending. New York ranks relatively high on advanced placement performance but still lags many other states in college attendance.

Overall, the data indicate that New York has not shirked in spending on schools, but the districts have not used the funds very effectively. There is little reason to believe that simply providing more funds to New York schools would lead to consistent gains in student achievement. In simplest terms, students in states that spend less than New York consistently perform better than New York students.

District Performance in New York State

The prior analysis has shown that the national picture is one of substantial increases in school spending without commensurate increases in student performance. This finding comes both from aggregate data and from detailed econometric studies. It also shows that the national picture is duplicated in New York State, where nation-leading spending has still left New York students below the national average in achievement.

New York State districts also show how achievement is not strongly driven by spending but instead how effectively funds are used. It is possible to see this by tracing how spending is related to

[17] Hanushek and Lindseth (2009). *See also, e.g., Campbell County School District, et al. v. State of Wyoming, et al.*, 907 P.2d 1238 (Wyo. 1995) (*Campbell I*); 19 P.3d 518 (Wyo. 2001) (*Campbell II*); 32 P.3d 325 (Wyo. 2001) (*Campbell III*); 181 P.3d 43 (*Campbell IV*).

[18] Hanushek and Lindseth (2009). *See also, e.g., Abbott, et al. v. Burke, et al.*, 100 N.J. 269 (N.J. 1985) (*Abbott I*); 119 N.J. 287 (N.J. 1990) (*Abbott II*); 153 N.J. 480 (N.J. 1998) (*Abbott V*); 163 N.J. 95 (N.J. 2000) (*Abbott VI*); M-1293-09 (N.J. May 24, 2011) (*Abbott XXI*).

[19] http://nces.ed.gov/nationsreportcard/

student performance across New York State districts. The lack of relationship between spending and achievement holds also for districts within New York State.

To demonstrate this, the following analysis considers increasingly stringent controls for factors other than district spending – specifically families and student backgrounds. In the end, the prior picture of little or no relationship of spending and achievement is duplicated for New York State.

It has been well-documented that disadvantaged students, as measured by such things as poverty or lack of intact families, generally come to school less prepared than more advantaged students. This relationship with poverty and other factors outside of the schools has been known since the well-known "Coleman Report" of 1966.[20] The question of importance for public policy is "how much do schools add to student achievement?" and "is the value-added of schools related to the resources that are available to the school?" In order to address these issues, it is necessary account for differences in achievement from families and from factors outside the control of schools so that we can obtain clear answers to the questions about the value-added of schools. The subsequent analysis uses increasingly stringent statistical controls in order separate the independent impact of schools and their resources from these other factors.

Charts 16-18 show how student achievement in grades 4, 6, and 8 relates to district spending per student in 2012 for math achievement. (Spending and achievement in the graphs are compared to the average in the state; i.e., the zero point for expenditure per pupil is the state average in 2012). Charts 19-21 show similar figures for reading achievement in 2012. These are raw relationships, unadjusted for any differences among students and districts. New York City and all districts spending greater than $35,000 per student are eliminated from the analysis, although retaining the high spending districts has virtually no impact on the results.[21] All statistical results are weighted by the number of tested students in each grade. The simple relationship between spending and performance is statistically significant by conventional standards in each of the six charts, but the magnitude of the relationship is small. For example, in grade 4, $1,000 per pupil in additional spending would on average be associated with a 0.8 point gain in student achievement (where the standard deviation of scores across districts is 11 points.

More importantly, however, raw comparisons of student performance across districts are likely to be very misleading, because the student populations differ dramatically across districts. The simplest approach is to adjust districts for differing compositions of student bodies. In all of the subsequent analyses here of student outcomes, district performance in 2012 is adjusted for levels of student background including poverty, proportion with Limited English Proficiency (LEP), and proportion special

[20] Coleman et al. (1966).

[21] New York City does not report spending for its individual community school districts, and would have a disproportionate impact on the results if it could be included because of its size. The 13 districts with spending greater than $35,000 are all small. Because the statistical results are weighted by students, these small districts have little impact on the results even though they are outliers in spending.

education. This adjustment comes from using statistical analysis, here some form of regression analysis, in order to find the independent effects of school and non-school factors.[22]

There is a positive relationship between spending and achievement in some grades and subjects, but the charts show that the effect of added spending is always very small.[23] While some of the relationships are statistically significant, the magnitude of the impact of different spending levels is always small.[24] For example, we plot the estimated impact of per pupil expenditures on achievement in grade 4 and show that the upward sloping line indicates a significant relationship – but an added $5,000 would only yield 2.0 points in fourth grade math (Chart 22) and 1.5 points in fourth grade reading (Chart 25).[25] These are 0.15-0.18 standard deviations for a roughly 25 percent increase in average district spending. Similar results are found for performance in grades 6 and 8 (Charts 23, 24, 26, and 27). Moreover, such gains could easily be obtained for little or no added spending through retargeting existing spending toward things that are known to matter, such as described below.

On these charts and all subsequent charts of New York district achievement, the essentially flat line relating the average achievement that is associated with more spending indicates that it is not resources that are having the biggest effect. At any given spending level, the difference between high achieving districts and low achieving districts is much larger and more important than the impact of spending per se.

On each of the charts, there is a huge difference in student outcomes after allowing for demographic factors (including poverty, special education, and LEP) and spending. For example, for fourth grade math (Chart 22) near the average spending (the 0 point on the horizontal axis) the highest achieving district is some 40 points above the lowest district. This difference is adjusted for demographics and again should be compared to the standard deviation of 11 points. In other words,

[22] The techniques are described in my textbook (Hanushek and Jackson (1977)), although the techniques are thoroughly covered in many textbooks and are often the subject of undergraduate courses in the social sciences.
[23] Called an added-variable plot, these graphs show the association between two variables after the influences of other control variables are taken out. Thus, for all remaining New York State district plots, both achievement and spending are first regressed on the other controls (in this case, percent free or reduced lunch, percent black, percent Hispanic, percent LEP, and percent special education). Only the residuals of these regressions, which comprise the part of the variation in the two variables that cannot be accounted for by the controls, are used in the graph. Thus, the graph makes sure the depicted association between the two variables is not driven by the control variables. The procedure is numerically equivalent to including the other controls in a multivariate regression of the dependent variable (achievement) on the independent variable under consideration in the graph.
[24] Statistical significance indicates the likelihood that any estimate was found by chance. It says nothing about the magnitude of any estimated effect, and it is magnitude of effect that is important for policy deliberations.
[25] All analyses use weighted regressions where the weights are the number of tested students. This special form of regression analysis gives more emphasis to the evidence from large districts than from small for two reasons. First, with small districts random measurement errors in student achievement become more important and could unduly influence the results. Second, the analysis should more closely match the impacts of larger districts that are simply serving more of New York State students. The small districts that spend more than $35,000 per student are excluded to ensure that the results are not driven by outliers, i.e., not driven by unusual circumstances. The exclusion has virtually no impact on the relevant results because they have relatively few students and thus have limited impact on the weighted results.

district performance is much more strongly dictated by factors other than spending per pupil. Districts that use their resources well get gains that are much larger than districts that do not.

The adjustments for demographic differences of districts may still not capture all of the differences across districts in preparation of students. This leads me to go even further to ensure that the results are not simply reflecting differences in families, neighborhoods, and other factors that also affect achievement. In other words, we want to make sure that we have really identified any impact of resources and are not just being confused by differences in the students themselves. Thus, we track the achievement path of individual students to see if these paths are affected by resources of the schools. In other words, by looking at the performance over time for individual students, the influences of families and other factors are accounted for in a very general way. This approach is very standard in the most rigorous research about student achievement.[26] We, for example, find fourth grade students who were in New York in third grade the previous year and calculate the third grade test scores for each district for those students. We then look at fourth grade performance after allowing for the previous demographics and after allowing for third grade performance. At this point expenditure has an even smaller impact on reading and math achievement. A $5,000 increase in district spending would be estimated to increase fourth grade achievement by just 0.85 points in math and 0.65 points in reading) – Charts 28 and 31. Similar results are found for grades 6 and 8 (Charts 29, 30, 32, and 33).

Poverty concentration has an important impact on achievement, so it is useful to look at the impact of spending just on poor children. Relying on just the most sophisticated approaches as described above that look at individual student gains in achievement, the impact of added spending on performance of poor children in New York State is even less than that reported for all students. If we look at the performance of poor children after allowing for their third grade scores, the estimated effect of spending is virtually identical to that for all students in the district – i.e., more spending is predicted to increase poor children's performance by less than a point for either math or reading in fourth grade (Charts 34 and 37). The effect of spending on achievement is qualitatively the same for poverty students in grade 6 and 8 (Charts 35, 36, 38, and 39).

Two conclusions stand out.

- When looking across New York schools, there is little to no relationship between how much money per pupil is spent and student results, either for all students or for students from poor families.
- Additionally, at any level of spending, one sees a very wide difference in how well students in different districts perform – a range that is much larger than the observed impact of tens of thousands of dollars.

The plaintiff districts look no different from the remaining districts in the state in terms of performance and the effect of resources. In the charts relating spending to achievement, by looking at the individual district points it becomes clear that some of the plaintiff districts do better than expected while others do worse. In other words, as a group they are neither performing better nor worse on

[26] Hanushek and Rivkin (2012)

average than would be expected based on their spending and demographics. There is no reason to believe that just providing them with additional funds will lead them to do significantly better than they are now doing. If they were adept at spending funds wisely, they would appear at the top of the achievement bubbles shown in the charts – not at a point in the middle as represented by the average spending-performance lines in the charts.

The New York district analysis reinforces the previous conclusions from the national analyses, as summarized in Charts 1-5, about the lack of a systematic relationship between spending and student performance. Just adding more money is unlikely to lead to improved student outcomes.

Court Ordered Spending: New Jersey

Since the early 1970s, many state courts have attempted to improve schools through rulings that called for extra spending. More than any other state, courts in New Jersey have had a dominant role in the determination of spending on schools.[27] Up through today, with more than 20 N.J. Supreme Court rulings, a number of the largest and most impoverished school districts have been permitted to spend almost anything they requested, including expanded preschool and a variety of other specialized programs. It has, unfortunately, not shown up in student outcomes.

While New Jersey has always been a high spending state, over the last decade its spending per pupil increased dramatically in comparison to the increases seen by other states (Chart 40). (As noted previously, New York State has increased spending more rapidly than New Jersey over this period). Given the dramatic increases in spending in New Jersey compared to the nation, one would expect by the presumptions of the plaintiffs that student scores would improve relative to those seen in other states. This has not been the case. On NAEP fourth grade reading, New Jersey students actually progressed at the same rate as other states (Chart 41). This lack of expected performance is also seen for black fourth graders in New Jersey where New Jersey actually lags the improvement in the rest of the nation (Chart 42). Precisely the same pattern of lagging the nation also holds for fourth grade math of all students and of black students (Charts 43 and 44). The only slightly different pattern is found for eighth grade math where New Jersey students do better than the rest of the nation, although the gains are very small relative to the overall spending increases (Charts 45 and 46).

New Jersey is a useful comparison. It is geographically and demographically close to New York State. It has instituted very large spending increases because of court orders. And it has little to show for the added spending in terms of achievement of students.[28]

Low- to No-Cost Options for Student Improvement

Because there is currently great inefficiency in spending – a fact detailed for New York districts – it is important to consider how the state can contribute to improved performance. Clearly, the state is

[27] Hanushek and Lindseth (2009)

[28] Similar conclusions come from the dramatic increases in spending (due to court orders) in Kentucky and Wyoming, although these states are not so demographically similar to New York State.

not in a position to dictate the nature of the educational process within schools and districts. It does not have sufficient knowledge of the capacities of districts or of the demands on them to prescribe how education is conducted across the heterogeneous districts of the state. Moreover, as has been demonstrated, simply providing extra money to districts has proven to be a very unsuccessful strategy. Nonetheless, the state does have a variety of low- or no-cost options that, according to existing research, can increase student performance. These include strong accountability (something New York has implemented and is advancing), more parental choice of schools (something New York is moving toward[29]), and better alignment of rewards with school and teacher performance (something New York is moving toward).

The common theme of these options is establishing an incentive structure so that districts will use available funds in a more efficient and productive manner. Each of these options is designed to identify and reward successful performance by districts and thus to encourage them to re-adjust their programs in order to improve performance of students. And each approach (contrary to just providing more funds) is built on a research base that indicates the likelihood of improvement in student outcomes.

District Actions for Student Improvement

The most fundamental actions that will improve student performance, however, clearly lie with the district. And, in this, recent analysis of student outcomes emphasizes the pivotal role of effective teachers, something that is entirely under the control of local districts. The existing evidence on the impact of teachers suggests that improving the teaching force would have enormous impacts on students' outcomes and additionally on the state and national economy. This research, under the general heading of value-added analysis, has used sophisticated statistical analysis to identify the variations in teacher quality that exist.[30] These studies, as mentioned before, separate the influence of teachers from other influences on student achievement, but instead of just looking at characteristics of teachers (such as experience or education) they focus on the total impact of teachers on student learning growth. We now have quite good estimates of how different teachers are in terms of their influence on student achievement growth.[31] Similarly, we have good estimates of the impact of achievement differences on the lifetime earnings of individuals.[32]

[29] In 2012, 3.9 percent of public schools in New York State were charter schools that served 2.3 percent of the student population; U.S. Department of Education (2014), table 216.90.

[30] The study of value-added of teachers began with Hanushek (1971), Murnane (1975), and Armor et al. (1976). There have been extensive analyses along similar lines such as Rockoff (2004), Aaronson, Barrow, and Sander (2007), Kane, Rockoff, and Staiger (2008), and Chetty, Friedman, and Rockoff (2013). There has also been considerable investigation of the properties of these estimates as summarized and evaluated in Hanushek and Rivkin (2012) and Chetty, Friedman, and Rockoff (2014a).

[31] Hanushek and Rivkin (2010) provide a summary of the quantitative estimates of the variation in teacher quality through 2010. Chetty, Friedman, and Rockoff (2014a) discuss various challenges to existing estimates and show that the existing models survive the multiple questions that have been raised.

[32] Hanushek (2011), Hanushek et al. (2013), Chetty, Friedman, and Rockoff (2014b).

Chart 48 provides estimates of the present value of future student incomes according to where the teacher falls in the distribution of teacher effectiveness.[33] The chart highlights the fact that the full impact of a teacher depends on two things: the effectiveness of the teacher and the number of students that she directly affects (i.e., the number of students in the class). A good teacher with 15 students in her class will have less overall impact than the same teacher with 20 students simply because with a larger class she is able to expand the learning of a greater number of students. Thus, the horizontal axis in the chart is class size.[34] The vertical axis shows that present value of added incomes compared to an average teacher (the zero point on the vertical axis).

Compared to an average teacher, in terms of individual incomes a 75th percentile teacher annually can produce over $400,000 added lifetime income for a class of 30. This is the annual impact of a 75th percentile teacher, and it would be repeated as long as the teacher continued to teach.

On the other hand, a 10th percentile teacher annually subtracts some $800,000 in present value (seen at the bottom in red on Chart 48). In other words, retaining ineffective teachers is very costly to students. [35]

It is also possible to see the economic impact of teacher effectiveness by concentrating on the least effective teachers.[36] If the least effective 5-8 percent of teachers could be brought up to the average teacher, U.S. achievement could rise to the level of Canada, or possibly Finland (Chart 49).[37,38]

From a policy perspective, districts should be trying various approaches to bring up their ineffective teachers, including mentoring, professional development, and other improvement programs. But past evidence indicates that these are not always successful. If these prove ineffective, districts need to find other ways to remove ineffective teachers from the classroom and to bring in more typical teachers.

These findings are important because they link to actions that individual districts can take to improve their teachers. The evidence on teacher effectiveness indicates common measures of "teacher

[33] The present value calculations sum up the added income over the working life of a student and weights all incomes later in a work career less than those immediately at graduation. All future income gains due to differences in achievement are discounted at three percent per annum.

[34] Note, however, that this is not the commonly discussed impact of larger or smaller class sizes on student achievement. As noted earlier, there has been a long standing argument that having fewer students in a classroom makes a teacher more effective because she can better judge the needs of individual students, can give more individualized attention to students, and the like. But, as discussed, extensive empirical analyses show that the overall impact of class size per se is small to nonexistent. Even looking at just the largest estimates of the impact of class size would not come close to matching the impact of differences in teacher quality as shown in Chart 48.

[35] Moreover, other recent research provides similarly large impacts on students' future income; Chetty, Friedman, and Rockoff (2014b).

[36] Hanushek and Woessmann (2012)

[37] See Hanushek (2009) and Hanushek (2011).

[38] Historical impacts of this on the U.S. economy, based on the analysis of how achievement affects economic growth, suggest that added growth could produce more than $100 trillion of added GDP (present value over 80 years); Hanushek and Woessmann (2011) and Chart 50. The current U.S. Gross Domestic Product is $17 trillion, so the gains from improving achievement amount to some five times our current GDP.

quality" including experience past the third year and teacher education are not closely related to teacher effectiveness.[39] Therefore, calls to invest in teachers with more of these attributes are unlikely to improve student achievement if done in typical ways. (Chart 51) Moreover, maintaining the single salary schedule where all teachers of a given experience level and amount of graduate education receive the same salary does not further higher achievement but instead locks in the inefficiencies in spending previously described.[40]

It is difficult to overstate the importance of these findings. School districts in general do not align salaries with teacher effectiveness. A highly effective teacher is generally just as likely to have a low salary as a high salary. The existing evidence suggests that teacher salaries are almost completely unrelated to effectiveness.[41] Said differently, raising the salaries of all teachers according to existing salary schedules would raise the salaries of both effective and ineffective teachers but would be unlikely to affect student achievement. [42]

Pre-School in New York

Various people argue that pre-school programs can be particularly beneficial to disadvantaged students.[43] These programs have the advantage of starting learning early in life, enhancing the benefits of later learning.[44] They also provide the potential for closing achievement gaps. These are exactly the reasons why the Governor and the State of New York have focused on providing quality pre-school programs.

There are two places, however, where the existing evidence is quite weak, making it difficult to legislate any universal preschool program. First, there is virtually no evidence about how different program designs (and expenses) are related to the effectiveness of the program.[45] For example, even though many people want to declare what the characteristics of a high quality program might be, little to no evidence links programs to outcomes. [46] Thus, with no evidence about how teacher salaries relate to student outcomes, it is not possible to relate enhanced salaries to the effective operation of a program. A similar lack of evidence plagues declarations about class size for preschool programs.

Second, there has been no evaluation of how best to design preschool programs in terms of the financial arrangements such as degree of means testing or how different fees affect the enrollment rates and the distribution of participants. Any preschool program, faced with the budgetary demands on states, must consider how program design affects the costs and the amount of substitution of private for public expenditure.

---

[39] Hanushek and Rivkin (2012)
[40] Hanushek (2007)
[41] Hanushek and Rivkin (2004), Hanushek (2011).
[42] Some have argued that paying higher salaries would expand the pool of applicants for teaching jobs and would thereby allow a district to increase quality over the long run. This, however, would only be the case if districts make good guesses of which people in the applicant pool will be highly effective – something that is not clear from past evidence. See, for example, Ballou (1996), Ballou and Podgursky (1997).
[43] Heckman et al. (2010), Witte (2007), Kirp (2007)
[44] Cunha and Heckman (2007)
[45] Hanushek and Lindseth (2009), Besharov et al. (2011), Whitehurst (2013), Fuller (2007)
[46] Barnett et al. (2014)

New York has been proceeding in a responsible way – acknowledging and supporting the importance of preschool while phasing in its scope and application.[47] This approach permits the state to evaluate key issues in program design without locking in expensive and ineffective features. Further, through its mixed provider strategy it permits the broad development of cost-effective programs. Just as with the funding for K-12 schools discussed above, simply providing more money to pre-schools without understanding how and when it is effective would be irresponsible. Therefore, a thoughtful expansion with experimentation in different structures is the best way to proceed.

Projecting Spending, Partial Revenues, and Gaps

When considering the relationship between resources and student outcomes, one needs to look at how the spending of a district relates to its students' performance, as seen in the prior exhibits. For that reason, analyses such as by Bruce Baker do not make any sense.[48] He presents an analysis of state and local revenues – excluding any federal revenues and not ever considering what is actually spent. Nationally, the federal government provided 12.5 percent of total revenue in 2011.[49] In New York State, the federal government contributed 9.3 percent of total revenues.[50] If one traces dollars into the schools and the classroom, it is impossible to tell the difference between a dollar from the state government and a dollar from the federal government, since all funds are fungible. Thus, trying to exclude federal money from a district's spending is simply not feasible. Neither can one talk about the effectiveness or impact of state and local revenues as distinct from federal revenues. It is again impossible to segregate the impact of state and local revenue from federal because all sources of funds go into the overll spending decisions of a district. Further, if revenues in any year are not the same as the actual expenditure of funds in the schools (as they typically are not), one would want to look at what the districts actually spend.

A related point comes from the projections of various "revenue gaps" or "adequacy gaps." All of Baker's analysis rests on projecting various district revenue numbers from past cost studies or analyses . These analyses are then projected to the current day to obtain his estimates of what is needed (or supposedly promised). But these projections assume an updating of numbers that bears no relationship to the reality of recent years. In particular, the recession of 2008 led to dramatic declines in the revenues of states, forcing them to reconsider all aspects of their budgets and spending. Related to this, the federal government dramatically increased its education spending, in part thrugh the stimulus package (American Recovery and Reinvestment Act of 2009). This spending was designed in part to offset any budgetary shortfalls of states in the wake of the recession. As a result federal revenues as a proportion of total revenues for K-12 schools grew from 8.2 percent in 2008 to 12.7 percent in 2010. As shown in chart 52, the balance of revenues shifted significantly. But this is entirely ignored by considering just state and local revenues and making simplistic projections. Any projections based upon just a portion of total revenues and ignoring this shift in funding responsibilities makes no sense.

[47] See the discussion of approaches in Hanushek and Lindseth (2009).
[48] Bruce D. Baker, "Evaluation of New York State's School Finance System," October 23, 2013.
[49] U.S. Department of Education (2014), Table 235.20 [accessed July 25, 2014].
[50] U.S. Department of Education (2014), Table 235.20 [accessed July 25, 2014].

Conclusion

Improving achievement is very important for the children of New York State, for incomes of New York residents, and for the well-being of the United States. Extensive evidence shows, however, that achievement is unlikely to improve simply by providing more resources. In fact, New York State has pursued such a spending program over the past decade – with no results to show for it.

Districts do, however, have the possibility of dramatically improving student performance by instituting better teacher selection, evaluation, and retention policies. Without improving how money is spent by the districts, New York State has little chance of improving.

The key decisions are made by individual school districts. If a district chose to align salaries with effectiveness and worked to keep the most effective teachers, it could by existing evidence dramatically improve student performance. Moreover, by making decisions based on existing evidence, the district could take this step with *no additions in resources*.[51] The fact that some districts make better choices than others lies behind the previous spending graphs. Some districts clearly get and retain more effective teachers than others while spending the same amount. And, there is no reason to believe that just providing more money to districts will lead them to use their funds more wisely. A district that currently gets low achievement from its spending is unlikely to make significantly better decisions with any additional funds.

In sum, there is reliable research to show that increasing teacher quality can lift achievement and can be sufficient to close achievement gaps. At the same time, there is no reliable research that indicates just providing increased funding would close achievement gaps. Providing extra funding for compensatory purposes has been done over a long period of time – both with federal and with state funding – but the achievement gaps have persisted and overall achievement has been below desired levels.

Said differently, the gains from improved personnel and management of local districts are enormous. And closing achievement gaps through improved teachers is much more likely than simply providing more funds to schools serving disadvantaged students and hoping that something new produces better results.

Eric A. Hanushek

November 13, 2014

[51] The spending-achievement effects of layoffs based on performance for New York are analyzed in detail by Boyd et al. (2011). A similar set of findings is found in Goldhaber and Theobald (2011).

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and student achievement in the Chicago public high schools." *Journal of Labor Economics* 25, no. 1 (January): 95-135.

Armor, David J., Patricia Conry-Oseguera, Millicent Cox, Niceima King, Lorraine McDonnell, Anthony Pascal, Edward Pauly, and Gail Zellman. 1976. *Analysis of the school preferred reading program in selected Los Angeles minority schools*. Santa Monica, CA: Rand Corp.

Ballou, Dale. 1996. "Do public schools hire the best applicants?" *Quarterly Journal of Economics* 111, no. 1 (February): 97-133.

Ballou, Dale, and Michael Podgursky. 1997. *Teacher pay and teacher quality*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Barnett, W. Steven, Megan E. Carolan, James H. Squires, and Kirsty Clarke Brown. 2014. *The State of Preschool 2013 : First Look*. Washington, DC: National Center for Education Statistics (May).

Besharov, Douglas J., Peter Germanis, Caeli Higney, and Douglas M. Call. 2011. "Summaries of twenty-six early childhood evaluations." Welfare Reform Academy College Park, MD: University of Maryland (July).

Betts, Julian R. 1995. "Does school quality matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics* 77, no. 2 (May): 231-247.

Bowles, Samuel, and Henry M. Levin. 1968. "The determinants of scholastic achievement--an appraisal of some recent evidence." *Journal of Human Resources* 3, no. 1 (Winter): 3-24.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2011. "Teacher Layoffs: An Empirical Illustration of Seniority versus Measures of Effectiveness." *Education Finance and Policy* 6, no. 3 (Summer): 439-454.

Burtless, Gary, ed. 1996. *Does money matter? The effect of school resources on student achievement and adult success*. Washington, DC: Brookings.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014a. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Revew* 104, no. 9 (September): 2593–2632.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014b. "Measuring the impacts of teachers II: Teacher value-added and the student outcomes in adulthood." *American Economic Revew* 104, no. 9 (September): 2633–2679.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2013. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." Mimeo. Cambridge, MA: Harvard University.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Cunha, Flavio, and James J. Heckman. 2007. "The technology of skill formation." *American Economic Review* 97, no. 2: 31-47.

Fuller, Bruce. 2007. *Standardized childhood: The political and cultural struggle over early education*. Stanford, CA: Stanford University Press.

Goldhaber, Dan, and Roddy Theobald. 2011. "Managing The Teacher Workforce in Austere Times: The Implications of Teacher Layoffs." Center for Education Data and Research CEDR Working Paper 2010-07-2. Bothell, WA: University of Washington Bothell (March).

Grubb, Norton W. 2009. *The Money Myth; School Resources, Outcomes, and Equity*. New York: Russel Sage Foundation.

Hanushek, Eric A. 1971. "Teacher characteristics and gains in student achievement: Estimation using micro data." *American Economic Review* 60, no. 2 (May): 280-288.
Hanushek, Eric A. 1981. "Throwing money at schools." *Journal of Policy Analysis and Management* 1, no. 1 (Fall): 19-41.
Hanushek, Eric A. 1986. "The economics of schooling: Production and efficiency in public schools." *Journal of Economic Literature* 24, no. 3 (September): 1141-1177.
Hanushek, Eric A. 1989. "The impact of differential expenditures on school performance." *Educational Researcher* 18, no. 4 (May): 45-51.
Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *Economic Journal* 113, no. 485 (February): F64-F98.
Hanushek, Eric A. 2006. "School resources." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 865-908.
Hanushek, Eric A. 2007. "The Single Salary Schedule and Other Issues of Teacher Pay." *Peabody Journal of Education* 82, no. 4 (October): 574-586.
Hanushek, Eric A. 2009. "Teacher deselection." In *Creating a new teaching profession*, edited by Dan Goldhaber and Jane Hannaway. Washington, DC: Urban Institute Press: 165-180.
Hanushek, Eric A. 2011. "The economic value of higher teacher quality." *Economics of Education Review* 30, no. 3 (June): 466-479.
Hanushek, Eric A., and John E. Jackson. 1977. *Statistical methods for social scientists*. NY: Academic Press.
Hanushek, Eric A., and John F. Kain. 1972. "On the value of `equality of educational opportunity' as a guide to public policy." In *On equality of educational opportunity*, edited by Frederick Mosteller and Daniel P. Moynihan. New York: Random House: 116-145.
Hanushek, Eric A., and Alfred A. Lindseth. 2009. *Schoolhouses, courthouses, and statehouses: Solving the funding-achievement puzzle in America's public schools*. Princeton, NJ: Princeton University Press.
Hanushek, Eric A., and Steven G. Rivkin. 1997. "Understanding the twentieth-century growth in U.S. school spending." *Journal of Human Resources* 32, no. 1 (Winter): 35-68.
Hanushek, Eric A., and Steven G. Rivkin. 2004. "How to improve the supply of high quality teachers." In *Brookings Papers on Education Policy 2004*, edited by Diane Ravitch. Washington, DC: Brookings Institution Press: 7-25.
Hanushek, Eric A., and Steven G. Rivkin. 2010. "Generalizations about using value-added measures of teacher quality." *American Economic Review* 100, no. 2 (May): 267-271.
Hanushek, Eric A., and Steven G. Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics* 4: 131-157.
Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann. 2013. "Returns to skills around the world." NBER Working Paper 19762. Cambridge, MA: National Bureau of Economic Research (December).
Hanushek, Eric A., and Ludger Woessmann. 2011. "How much do educational outcomes matter in OECD countries?" *Economic Policy* 26, no. 67 (July): 427-491.
Hanushek, Eric A., and Ludger Woessmann. 2012. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." *Journal of Economic Growth* 17, no. 4 (December): 267-321.
Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The rate of return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94, no. 1-2 (February): 114-128.
Hedges, Larry V., and Ingram Olkin. 1985. *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. "What does certification tell us about teacher effectiveness? Evidence from New York City." *Economics of Education Review* 27, no. 6 (December): 615-631.

Kirp, David L. 2007. *The sandbox investment: The preschool movement and kids-first politics*. Cambridge: Harvard University Press.

Krueger, Alan B. 1999. "Experimental estimates of education production functions." *Quarterly Journal of Economics* 114, no. 2 (May): 497-532.

Ladd, Helen F., and Janet S. Hansen. 1999. *Making money matter: Financing America's schools*. Washington, DC: National Academy Press.

Mishel, Lawrence, and Richard Rothstein, eds. 2002. *The class size debate*. Washington, DC: Economic Policy Institute.

Murnane, Richard J. 1975. *Impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73, no. 2 (March): 417-458.

Rockoff, Jonah E. 2004. "The impact of individual teachers on student achievement: Evidence from panel data." *American Economic Review* 94, no. 2 (May): 247-252.

Stiefel, Leanna, Amy Ellen Schwartz, Ross Rubenstein, and Jeffrey Zabel. 2005. "Measuring school efficiency: What have we learned?" In *Measuring school performance and efficiency: Implications for practice and research*, edited by Leanna Stiefel, Amy Ellen Schwartz, Ross Rubenstein, and Jeffrey Zabel. Larchmont, NY: Eye on Education: 1-15.

U.S. Department of Education. 2014. *Digest of Education Statistics, 2013*. Advanced Release ed. Washington, DC: National Center for Education Statistics.

Whitehurst, Grover J. 2013. "New Evidence Raises Doubts on Obama's Preschool for All." In *The Brown Center Chalkboard*. Washington, DC: Brookings Institution.

Witte, John F. 2007. "A Proposal for State, Income-Targeted, Preschool Vouchers." *Peabody Journal of Education* 82, no. 4: 617-644.

Word, Elizabeth, John Johnston, Helen Pate Bain, B. DeWayne Fulton, Jayne Boyd Zaharies, Martha Nannette Lintz, Charles M. Achilles, John Folger, and Carolyn Breda. 1990. *Student/teacher achievement ratio (STAR), Tennessee's K-3 class size study: Final summary report, 1985-1990*. Nashville, TN: Tennessee State Department of Education.