Deconstructing the Myth American Public Schooling Inefficiency

Bruce B. Baker

Mark Weber

Rutgers University, Graduate School of Education

WORKING PAPER 2016_01

Abstract

In this paper, we begin by classifying the arguments that assert American schools are relatively inefficient into two categories: the *long-term trend argument* and the *international comparison argument*. Our focus herein is on the latter of these two. We then describe two frameworks for approaching either of these arguments: *cost efficiency* and *production efficiency*. We find that the typical spending/outcome model used to make the case that the United States is a relatively inefficient nation is wholly unsuitable for drawing these or any conclusions. Accounting for differences in student populations is helpful, but still inadequate for building a model that can be used to assess a country's relative efficiency. Evaluating educational inputs such as teacher wages and class sizes can further refine comparisons between nations; however, it is unlikely that even these refinements are enough to conduct analyses that can credibly back claims about the relative efficiency of America's education system.

1.0 American Public Schooling Inefficiency

Persistent arguments that school spending has little or no relation to school quality are often buttressed by two evidentiary claims:[1]

1) Education spending in the United States has doubled if not tripled over the long term yet commonly measured educational outcomes have remained "virtually flat";
2) Education spending in the United States is among the highest in the world, but our commonly measured outcomes lag well behind other nations.

We describe these as the *long-term trend argument*[2] and the *international comparison* argument.[3] Both tend to dominate the public discourse, even as more nuanced assessments of how and why money matters exist in academic literature.[4]

The validity of these arguments is suspect even at face value if only because accurately measuring the changes in test-based outcomes attributable to the effectiveness and efficiency of a nation's educational system or the value of the education dollar over time and across contexts is, at best, a complicated endeavor. Measuring changes in test scores, across changing cohorts and contexts to determine whether performance has remained "virtually flat" is, by itself, no simple task. But even if we set aside these difficulties, the claim of stagnant national progress is not held up by the evidence. Rothstein (2011) shows that, in fact, "On these exams [National Assessment of Educational Progress], American students have improved substantially, in some cases phenomenally." (p. 1) Other research confirms that, when accounting for differences in student disadvantage, US students perform much better than what is suggested by commonly cited,

[1] In a 2011 Opinion piece in Huffington Post titled *Flip the Curve: Student Achievement vs. School Budgets* Bill Gates of Micrsoft used both of these claims. Haddad (2014) also uses both claims.

[2] Borrowing from Jackson, Johnson and Persico (2015)

[3] For a discussion of the sources and recurrence of these arguments, see Baker (2016) and Baker and Welner (2011).

[4] Baker (2016) identifies three additional common arguments made by those wishing to either downplay or deny outright the role of funding in determining schooling quality: 1) tallies, or "vote counts" of correlational studies between spending and outcomes, without regard for rigor of the analyses and quality of the data on which they depend; 2) Anecdotal assertions that states such as New Jersey and cities such as Kansas City provide proof positive that massive infusions of funding have proven ineffective at improving student outcomes; and 3) The assertion that how money is spent is much more important than how much is available.

unadjusted rankings that fail to account for changes in subgroup proportions when aggregating test results (Carnoy and Rothstein, 2013). Similarly, as we shall see below, international comparisons of educational spending that use simple mean per pupil spending figures compiled by the Organization for Economic Cooperation and Development are wholly inadequate to support any claims of the US education system's alleged profligacy.

Regarding the *long term trend* argument, Jackson, Johnson and Persico (2015) rebut the facile nature of assertions built exclusively on upward trending spending and supposed "virtually flat" test scores with a clever analogy:

> "…consider the following true statistics: between 1960 and 2000 the rate of cigarette smoking for females decreased by more than 30 percent while the rate of deaths by lung cancer increased by more than 50 percent over the same time period. An analysis of these time trends might lead one to infer that smoking reduces lung cancer. However, most informed readers can point out numerous flaws in looking at this time trend evidence and concluding that 'if smoking causes lung cancer, then there should have been a large corresponding reduction in cancer rates so that there can be no link between smoking and lung cancer."

The unconditional, contextually detached logic of this example is analogous to the assertion that increased spending over time, coupled with flat test scores, proves that school spending matters little. In addition, at least the second of these assertions (flat test scores) is false to begin with, and the first (increased spending) is far more complicated than it appears.

The *long term trend* argument has been thoroughly rebutted on numerous occasions (Baker, 2015, Baker & Welner, 2011, Jackson, Johnson and Persico, 2015; Rothstein, 2011). Yet except for Carnoy and Rothstein's (2013) dissection of the test score comparisons, less attention has been paid to the *international comparison* argument. Essentially, this argument contends that the United States is spending more and performing worse than other countries simply because it is less efficient than those countries. We must be doing something wrong, and as a result should take efficiency lessons from those who "do more with less."[5] The assertion that the American public education system is less efficient than nearly any other education system in the world often goes something like this:

> *"Compared to other countries, America has spent more and achieved less. We need to build exceptional teacher personnel systems that identify great teaching, reward it, and help every teacher get better."*[6]
>
> *Bill Gates, Microsoft co-founder*

The evidentiary basis for this claim is most often either built on simple references to US per Pupil Spending on primary, lower secondary and upper secondary education (Indicator B1) compared with other nations as reported in the OECD *Education at a Glance* series, coupled with scale scores of US students on the Program for International Student Assessment (PISA).[7] We spend high according to OECD, score low according to PISA, and thus are inefficient. Marginally more rigorous analyses plot PISA scores against national average per pupil spending estimates (annual, or cumulative for ages 6 to 15), showing the US among the world's higher spending nations, with no additional yield in PISA scores, and an overall weak relationship between spending and test scores across nations (See OECD, 2012).[8]

2.0 Efficiency Analysis in Education

There exists a relatively large body of empirical literature on applied efficiency analysis in education (Grosskopf, Hayes & Taylor, 2014); that is, the very kind of methods and models that

[5] More extreme versions of this argument go so far as to assert that "The United States Spends More on Schools Than Any Society In Human History" https://higheredrevolution.com/the-united-states-spends-more-on-schools-than-any-society-in-human-history-d5988649d73e#.1jk6bfskq

[6] Gates (2011) http://www.huffingtonpost.com/bill-gates/bill-gates-school-performance_b_829771.html The graph Gates includes with this piece, which purports to show the US has increased education spending dramatically with little improvement in test scores, was critiqued here: "Junk Charts (2011) Bill Gates should hire a statistical advisor." http://junkcharts.typepad.com/junk_charts/2011/04/bill-gates-should-hire-a-statistical-advisor.html

[7] See for example: http://www.ncee.org/2015/01/statistic-of-the-month-education-performance-equity-and-efficiency/, which explains "Several countries have high scores on PISA and achieve high levels of equity in student performance while spending relatively little per student" as a basis for comparing "efficiency."

[8] In a review of 43 "high quality" studies of schooling resources in developing nations, Glewwe, Hanushek, Humpage and Ravina (2011) suggest that perhaps among less developed nations increases in basic resources (roof over head, desk in which to sit) do associate with improved school quality, but among more developed higher spending nations, the spending to outcome relationship breaks down for a variety of reasons.

might be used to distill whether one institution, organization or jurisdiction more efficiently produces educational outcomes than another. The same basic principles and methods apply whether we are evaluating individual schools, local education agencies (public school districts in the US model of schooling), state school systems, or national education systems. The main difference is that as we move to more and more geographically, economically and culturally distal systems constructing consistent measures of schooling inputs and outcomes, while considering confounding factors, becomes far more difficult yet far more important.

Efficiency analysis in education (or any sector) may be framed from a production perspective or a cost perspective. Production efficiency and cost efficiency are flip sides of the education spending coin. Each involves identifying outcomes, spending toward achieving those outcomes, and various observable conditions – cost factors – that affect the spending → outcome relationship. However, each seeks to answer different questions:

- Cost Efficiency: Given the outcomes a unit (school) currently achieves, compared to the lowest spending unit achieving the same (all else equal), how much does the given unit spend?

- Production Efficiency: Given the current spending levels (and other factors), how do the outcomes of the unit compare to the maximum outcomes achieved with comparable spending (and other factors)?

Figure 1 presents a view of *cost efficiency*, across schools of varied scale, or enrollment size. The "cost" of producing a given level of outcomes varies by the scale of the school, where very small schools face much higher costs to achieve the same outcomes as scale efficient schools. A subset of schools may fall along the lower boundary, the underlying "cost frontier" or minimum expense, at a given school size, at which the desired outcomes can be produced. Most schools will fall some distance above that frontier, or be spending somewhat more than the minimum to achieve the same outcomes. Some of those differences may be attributed to factors we've simply missed in our data and models – because they weren't measured or were measured poorly; this is partly why we shouldn't live by the data and models alone. Some of those differences may also include expenditures deemed valuable by constituents, but which don't contribute directly to the outcome measures in question. Valuable lessons may be learned from exploring both those schools that fall along the cost frontier and those that deviate from it.

Figure 1. Cost Efficiency Framework

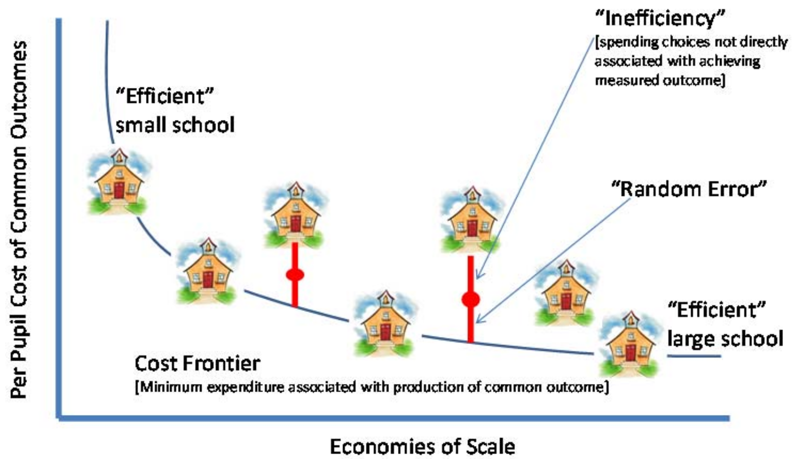
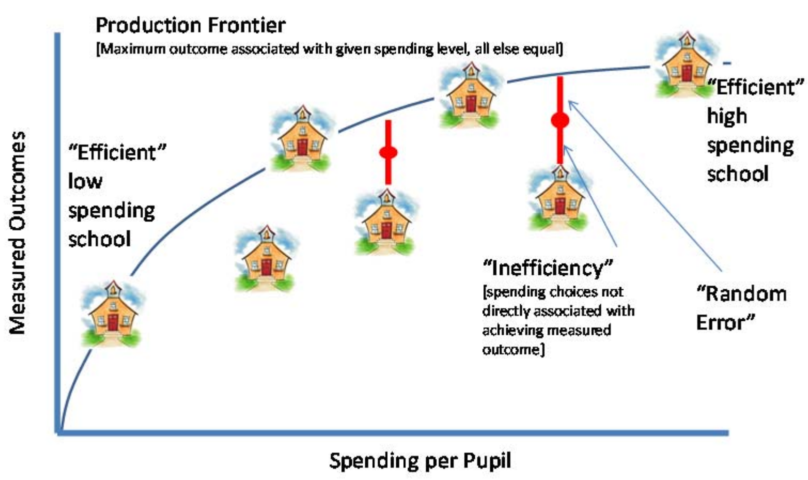


Figure 2 presents the *production efficiency* perspective. Here, it is assumed that for each additional dollar input to schooling, there is a commensurate gain in student outcomes, with diminishing returns. Some schools will fall along – or define – the frontier, or maximum output for any given level of expense, and others will fall below that frontier, for the same types of reasons that schools deviate from the cost frontier.

Figure 2. Production Efficiency Framework



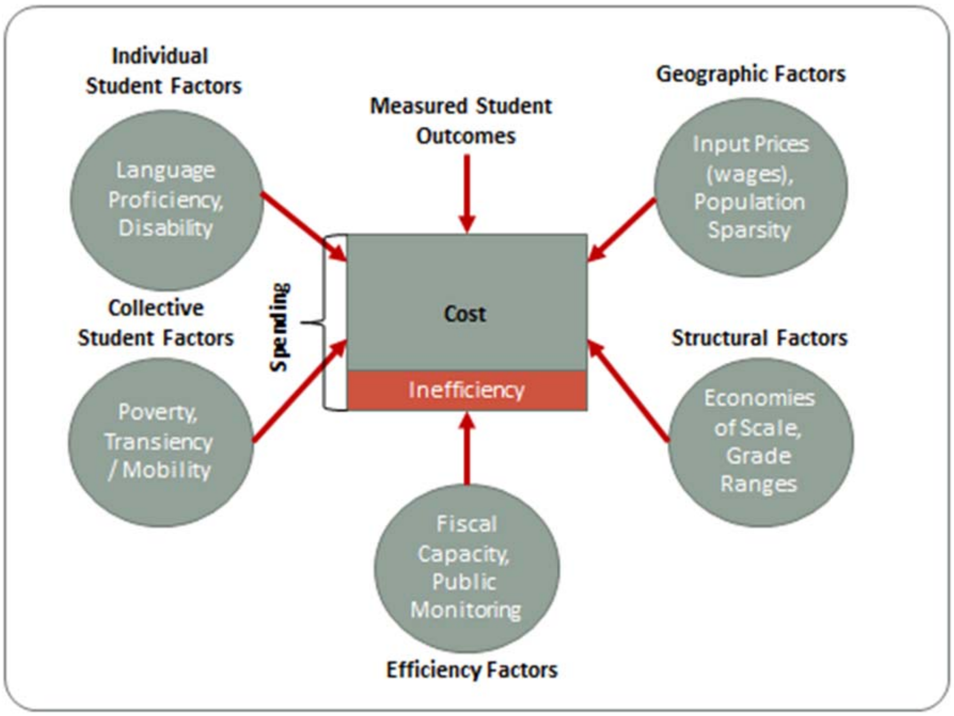
The visuals here are stylized for simplicity, reducing efficiency analysis to two dimensions. No credible scholar or analyst would consider a simple, unconditional cross-sectional analysis of the relationship between nominal per pupil spending and average/aggregate test score levels as legitimate basis for making "efficiency" comparisons. That is, a simple scatterplot of the relationship between per pupil spending (nominal, unadjusted in time or space across units) on the horizontal axis and test scores on the vertical axis is relatively useless for inferring the spending to outcomes relationship or making judgements about relative efficiency. This is true for comparisons across schools within a jurisdiction, across local public school districts within a state and across US states. In addition, it is certainly true for cross-national comparisons.

Figure 3 summarizes factors that affect education costs from one school to the next, one district or state to the next and one nation or continent to the next. Once we step outside comparisons of schools operating within a single labor market of consistent geography (no huge swings in population density, major geographic barriers, etc.) simple comparisons of education spending from one unit to the next are no-longer simple. The value of the education dollar varies toward purchasing even the same quality and quantity of educational inputs, where the key inputs to education are teachers, and where the wage required to recruit and retain teachers of specific qualifications varies from one location and setting to the next. One of the most significant factors affecting the "cost" of providing comparable educational services across schools and districts is economies of scale, or average school size, which may be constrained by geography and population density (Andrews, Duncombe and Yinger, 2002). Further, when comparing education outcomes, one must consider various attributes of the student populations being served, to the extent that those attributes correlate with outcomes.

Credible models comparing school or district efficiency across US schools and districts typically consider student disability status (total numbers and severity levels), language proficiency status, indicators of child/family income and poverty and sometimes race/ethnicity of student population (Baker, 2011).

Figure 3, which presents the "cost model" perspective, also includes a reference to "inefficiency." As suggested in Figures 1 and 2, inefficiency is that difference between what was actually spent, versus what needs to be spent at a minimum to achieve a given level of outcomes. A separate body of literature has addressed attributes of public jurisdictions (jurisdictions with control over public budgets, taxing and spending) that may be associated with inefficiency (Borge, Falch & Tovmo, 2008, Grosskopf, Hayes, Taylor & Weber, 2001). In particular, measures of inter-jurisdictional competition and "public monitoring" (often measured by extent of proximal/local involvement in finance and decision making) have been identified as leading to greater efficiency. Inefficiency, on the other hand, is more likely to be associated with greater fiscal capacity; in other words, those who can spend more (even for constant measured outcomes) are more likely to do so.[9]

Figure 3. Factors Affecting Education Costs



Source: Baker & Levin (2014)

A thorough applied production efficiency analysis in education would estimate the *gains in student outcomes* – in other words, the value added over what students enter with[10] -- as a function of a) the resources expended on comparable/relevant services, b) the geographic factors affecting structural costs and input prices, and c) the student characteristics that might have exerted

[9] Duncombe and Yinger (2007) explain that the "cost model" perspective is more useful for sorting out cost vs. inefficiency because it permits (more logically than the production model) the inclusion of measures of fiscal capacity and public monitoring to more precisely isolate "cost."

[10] Alternatively, one could focus on levels of student outcomes, to the extent that one sufficiently captures student background characteristics predictive of students' initial performance

exogenous influence on achievement gains, including disability status, language proficiency and child poverty. By contrast, in cost modeling, we predict the spending levels (as the dependent variable) associated with achieving given levels of student outcomes, controlling for factors that affect the value of the education dollar toward contributing to outcomes, and additionally correcting for factors in Figure 3 that may explain differences in inefficiency across institutions or jurisdictions (Baker, 2016, pp 12-14). Typically the models used would be estimated to schools or districts as units using multiple years of annual data, to ensure stable, reliable estimates (Gronberg, Jansen & Taylor, 2012). Even then, our ability to precisely, consistently identify more and less efficient schools, districts, states, or even countries is suspect due to the imprecision of the data used to create the models, and the many omitted variables that might bias those models (see Bifulco & Bretschneider, 2001, Duncombe & Bifulco, 2000).

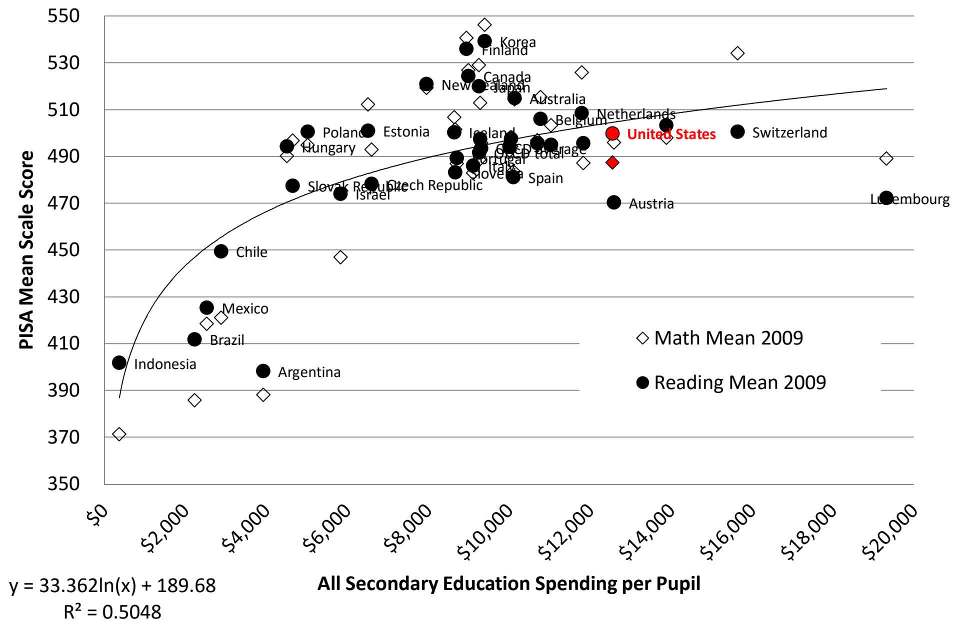
3.0 The Oft-invoked Unconditional PISA Production Function

Now that we've established what a credible production or cost analysis should include, let's take a look at commonly cited evidence behind the claim of US inefficiency, starting with what is actually one of the more rigorous analyses. In their report *Does Money Buy Strong Performance in PISA?* (OECD, 2012), the authors use two simple frameworks to evaluate a nation's educational efficiency: the relationship between both national Gross Domestic Product and PISA scores, and measures of per pupil expenditure and PISA scores. They find that for lower spending nations, there does indeed appear to be a relationship between GDP and outcomes, as well as a relationship between spending and outcomes. For wealthier, higher spending nations, however, this relationship falls apart.

Figure 4 presents our own recreation of the cross-national spending to PISA score relationship, fit with a log-linear (diminishing returns) curve (we present alternative versions in Appendix A). Note that this is the very kind of graph/analysis we earlier characterized as insufficient because it does not take into account student factors (individual and collective), geographic factors, or structural factors. But even when using this simplistic model, it would appear that the United

States is only slightly below the curve and certainly not the inefficient standout of Austria or Luxembourg. [11]

Figure 4. The Unconditional PISA Production Function



OECD (2012), Education at a Glance 2012: OECD Indicators , OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en
Indicator B1.1a: http://dx.doi.org/10.1787/888932665867
OECD (2010), PISA 2009 Results: What Students Know and Can Do – Student
Performance in Reading, Mathematics and Science (Volume I) http://dx.doi.org/10.1787/9789264091450-en

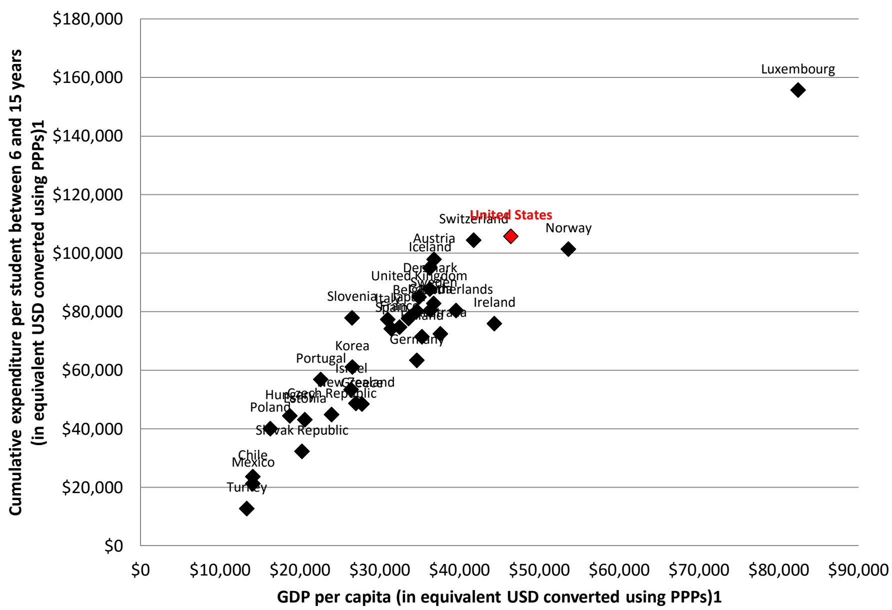
Figure 4 above appears on the surface to confirm the classic economic expectation of positive but diminishing returns in educational outcomes to marginal increases in educational spending. But this seemingly logical pattern is bound up in complex, circular or "endogenous" relationships. Notably, the OECD (2012) report finds nearly the same pattern for the relationship between GDP and PISA as for per pupil spending and PISA. This is unsurprising because, as shown in Figure 5, wealthier nations spend more on education. So children in wealthier nations score better on PISA (up to a point), while children in higher spending nations also score better on PISA (up to a point). But because education spending and GDP are correlated, the seemingly logical

[11] Alternatively, if we really wanted to make the US look bad, we could fit a linear model to the data, which would produce a straight trendline, rather than the better fitting and more appropriate log-linear model (which produces a curved trendline) that we use here. Using a linear model increases the distance from the trendline for the US and makes the country appear to be even more inefficient. Numerous such examples exist in the blogosphere, including: http://excelined.org/2013/09/20/things-heaven-earth-dr-ravitch-dreamt-ideology-ravitch-vs-reality-part-iii/ or here: http://bpr.berkeley.edu/2014/10/29/a-finnish-ed-model-for-national-education/ These models are made even more useless to support the claim of alleged US inefficiency by the exclusion of low-spending, low-outcome countries, which create the initial "bend" in the log-linear trendline.

relationship – the "production curve" shown in Figure 4 – actually tells us little to nothing about a nation's relative efficiency.

What we see here is not really a production function at all, but rather a pattern that appears because of all of the coinciding relationships underlying the data. For this reason, measuring "efficiency" against this curve is a suspect endeavor. Unlike legitimate production, cost and efficiency analysis, these analyses consider only a single year of cross-sectional data, and only levels of student outcomes; they do not adequately model the contribution of a nation's education policies and practices to gains in their students' outcomes.

Figure 5. National Fiscal Capacity and Education Spending



OECD (2010), Table I.2.20. Socio-economic indicators and the relationship with performance in reading, in *PISA 2009 Results: What Students Know and Can Do*, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789264091450-table76-en

4.0 Contexts & Covariates

If we wish to truly measure a country's educational efficiency, we must move beyond simple spending/outcomes correlations. For many reasons, however, it is nearly, if not entirely, impossible to estimate a legitimate cross-national education production or cost model. It's difficult enough, if not implausible, to estimate such a model for America's individual states given the vast structural and geographic differences between them. There are also significant problems with consistently measuring the value of the education dollar from New York to New Mexico (Taylor & Fowler, 2006) and problems with consistently measuring student need factors, including child poverty, from one region to the next (Baker, Taylor, Levin, Chambers & Blankenship, 2013). The difficulties in addressing these issues are compounded when the frame of comparison moves to the international level.

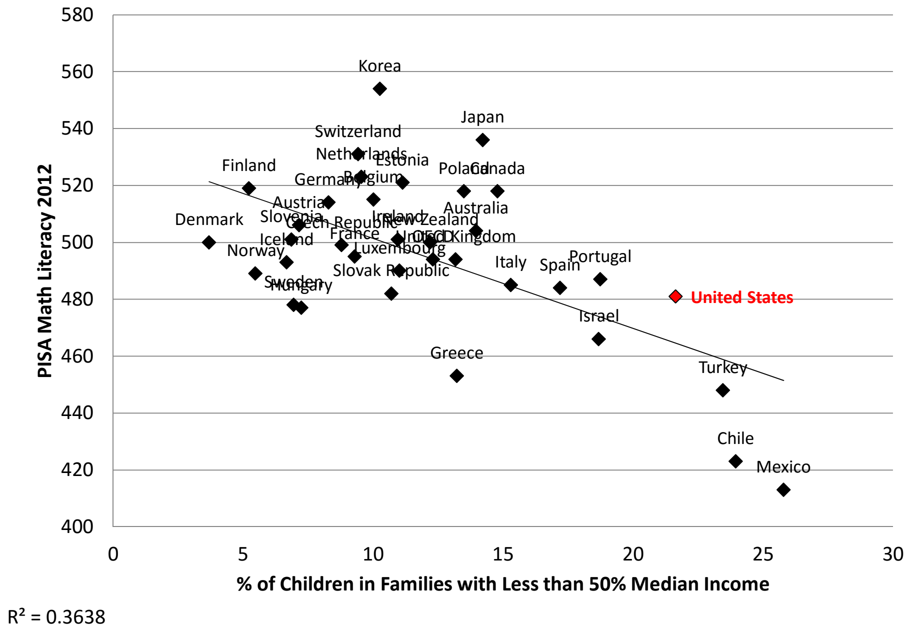
But the impossibility of achieving an adequate cross-national cost or production model should not impede us from at least exploring additional covariates and patterns that raise even more questions about common assertions of American inefficiency. For example how much do available (albeit limited) measures of student need, including economic status, affect national average/aggregate PISA scores? To the extent that they do at all, they must be considered when making judgments about the relative efficiency of US schools.

Socio-economic context

When constructing cost or production models for evaluating school or district efficiency, the goal is to find those measures of student attributes that are sufficiently exogenous – that is, do not have a circular relationship with the outcomes measured – and that, with reasonable causal explanation, are substantively correlated with the educational outcome measure under investigation. There are, for example, many approaches to measuring poverty or socio-economic status. Some measures, like the USDOE's "free or reduced-price lunch" metric, are crude, categorical variables. Other dichotomous measures, such as under/over the poverty line, will have differing thresholds, but the same limitations. Some measures characterize poverty in relative terms, with respect to the income distribution or some point within it; others measure poverty with respect to specific, absolute income levels (Coley & Baker, 2013). Due to the complexities of establishing specific, comparable income thresholds associated with "poverty" status (which are necessary for "absolute" poverty measurement) cross-national comparisons often use *relative* poverty measures, such as the share of children in families with less than half the median income.

Figure 6 shows that even these seemingly less precise measures of relative poverty are reasonably strongly associated with PISA mathematics literacy in 2012. The cross-national correlation is greater than .60 (r-squared =.3638). In Figure 6 it would appear that the US (on this one outcome measure) does better than expected, given its relative child poverty rate: the nation is above the trendline, indicating it outperforms this simple model's prediction.

Figure 6. PISA and Relative Poverty



PISA Data: http://nces.ed.gov/pubs2014/2014024_tables.pdf (table M4)
OECD Relative Poverty: Source: Provisional data from OECD Income distribution and poverty database (www.oecd.org/els/social/inequality).

Table 1 presents the correlations across a handful of measures compiled for an OECD (2010) report exploring relationships between economic indicators and reading performance. The two strongest correlations with reading outcomes are for the "proportion of the population in the age group between 35 and 44 years with tertiary education," and "Share of students in their country whose PISA index of economic, social and cultural status is below -1."[12] Notably, adult education levels, and child disadvantage levels are also associated with GDP (.41 and -.48), and accordingly with the spending measure (in this case, cumulative spending per student from age 6 to 15). In other words, a multitude of economic capacity, education spending, and student/family background characteristics are modestly related across nations (where the number of nations included in these calculations is only 34).

[12] "The Programme for International Student Assessment (PISA) index of economic, social and cultural status was created on the basis of the following variables: the International Socio-Economic Index of Occupational Status (ISEI); the highest level of education of the student's parents, converted into years of schooling; the PISA index of family wealth; the PISA index of home educational resources; and the PISA index of possessions related to "classical" culture in the family home." See: https://stats.oecd.org/glossary/detail.asp?ID=5401. See also Ganzenboom (2010) for information on the ISEI.

Table 1. PISA and other SES measures

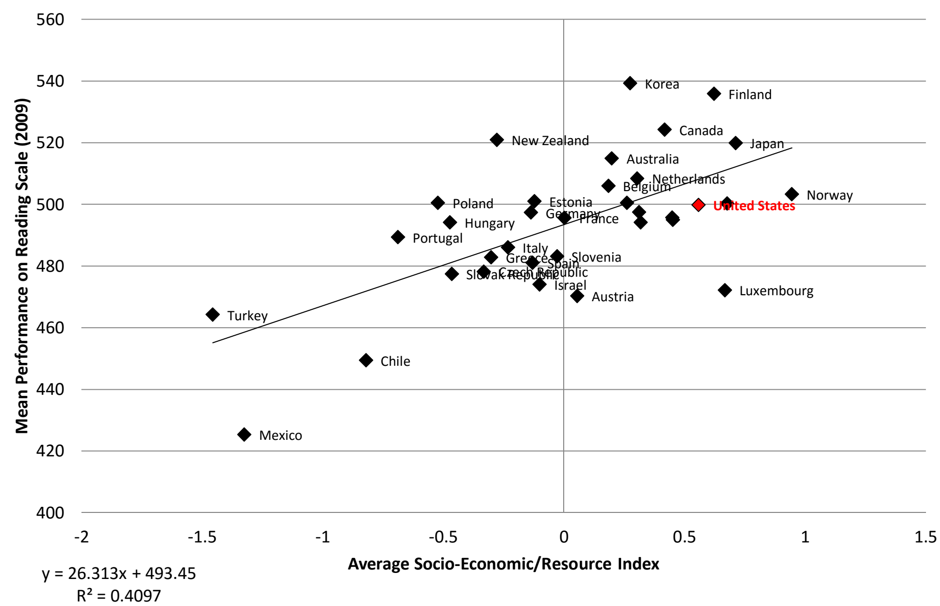
Indicator	Mean performance on the reading scale	GDP per capita (in equivalent USD converted using PPPs)1	Cumulative expenditure per student between 6 and 15 years (in equivalent USD converted using PPPs)1	Percentage of the population in the age group 35-44 years with tertiary education1	Proportion of 15-year-olds with an immigrant background	Share of students in their country whose PISA index of economic, social and cultural status is below -1
GDP per capita (in equivalent USD converted using PPPs)1	0.24					
Cumulative expenditure per student between 6 and 15 years (in equivalent USD converted using PPPs)1	0.30	0.93				
Percentage of the population in the age group 35-44 years with tertiary education1	**0.67**	0.41	0.43			
Proportion of 15-year-olds with an immigrant background	0.11	0.69	0.63	0.38		
Share of students in their country whose PISA index of economic, social and cultural status is below -1	**-0.68**	-0.48	-0.55	-0.58	-0.33	
Size of the 15-year-old student population	-0.05	0.01	0.03	0.12	0.01	0.21

[1] Source: OECD, *Education at a Glance 2010: OECD Indicators.*

OECD (2010), Table I.2.20. Socio-economic indicators and the relationship with performance in reading, in *PISA 2009 Results: What Students Know and Can Do*, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789264091450-table76-en (I'm unclear on the sourcing here.)

Figure 7 shows the relationship between reading performance and the combined SES index, revealing a correlation of .64 (r-squared of .4097), or slightly stronger than the relationship between national relative poverty levels and math performance in 2012. Here, the US falls slightly below expectations but is certainly no underperforming (or over-performing) outlier. In any case, because measures of socio-economic status are significantly correlated to student outcomes, failure to consider these SES index or relative poverty measures when making assertions of relative efficiency across nations is a major oversight.

Figure 7. PISA Reading and Combined SES Index



OECD (2010), Table I.2.20. Socio-economic indicators and the relationship with performance in reading, in *PISA 2009 Results: What Students Know and Can Do*, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789264091450-table76-en

Heterogeneity of U.S. States

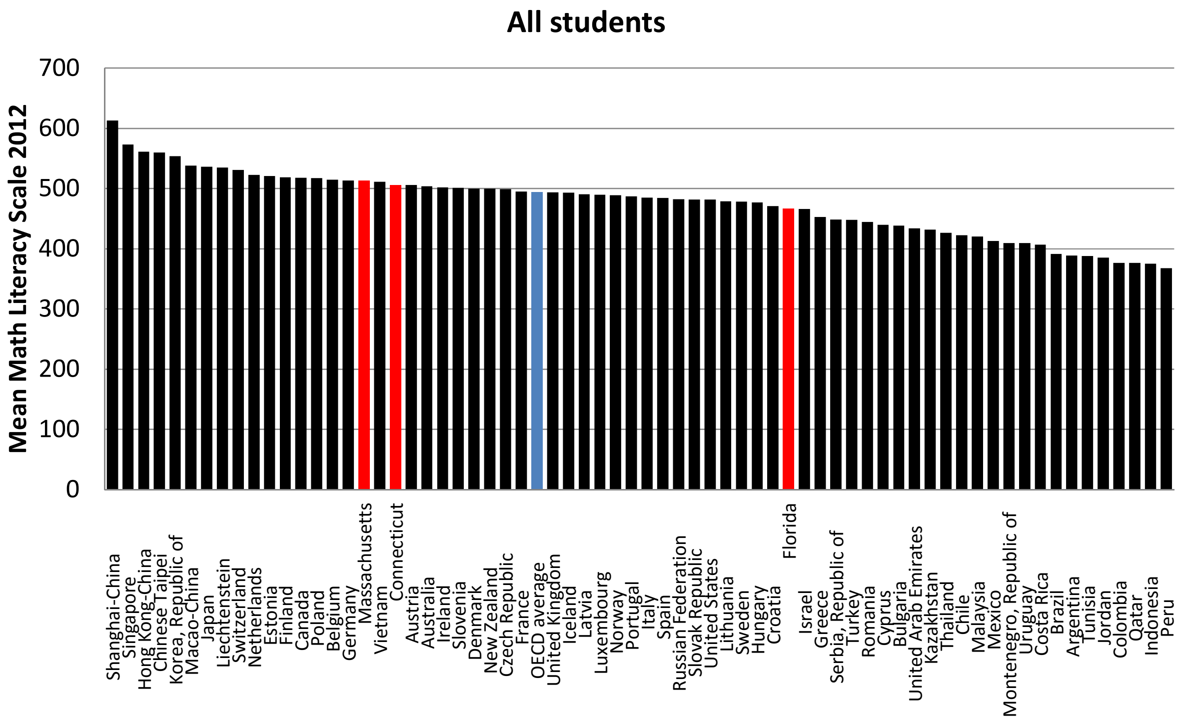
It is also important to recognize that the United States education system is, in fact, 51 largely independent education systems, where the majority of funding comes from state and local sources, and where accountability systems are adopted by states in compliance with federal statutes and regulations that allow for significant differences in funding, policies, and practices. US states also vary widely in measures of socio-economic status, including child poverty rates. The sheer size of the United States alone contributes to the heterogeneity of the country's student population. In the OECD 2010 report, for example, the US education system is reported as serving by far the largest total number of 15 year olds (at nearly 3.4 million) with Mexico second (at just over 1.3 million). Frequently cited "high performing" nations like Finland serve only 61,000 15 year olds (1.8% of US 15 year olds); Korea enrolls 630,000 15 year olds, 18.7% the size of US. These other systems are much smaller in magnitude and tend to be more highly centralized. There are also significant differences in the cultural, racial, and linguistic diversity of different countries; arguably, the United

States is more diverse than many "high performing" nations (Alesina et al., 2002; Fearon, 2003), although determining the extent of these differences is a complex endeavor.[13]

Because the states vary so significantly, we gain additional insights from those few reports which make efforts to compare individual US States to foreign nations. Figure 8 is based on data from a 2012 OECD report which included PISA data and constructed the Economic, Social, and Cultural Status index for nations and three US states – Massachusetts, Connecticut and Florida. Massachusetts and Connecticut are two relatively affluent Northeastern states with relatively high average per pupil spending. Florida is a higher poverty, much lower spending Southern state (Baker, Farrie, & Sciarra, 2015). Figure 8 shows that for all students, on average, Massachusetts and Connecticut students beat OECD averages while Florida students do not. Massachusetts and Connecticut students on average perform more similarly to Germany, Belgium and Austria, whereas Florida students perform more similarly to Croatia.

[13] Some have argued that the United States is relatively *less* diverse than many other OECD countries; see: http://educationbythenumbers.org/content/top-us-students-fare-poorly-international-pisa-test-scores-shanghai-tops-world-finland-slips_693/ Determining the relative diversity of different countries is, however, quite complicated, as our sources in this paragraph demonstrate. At the very least, any attempt to introduce diversity as a variable in a model of international comparisons on test scores should explain *how* diversity indices were calculated.

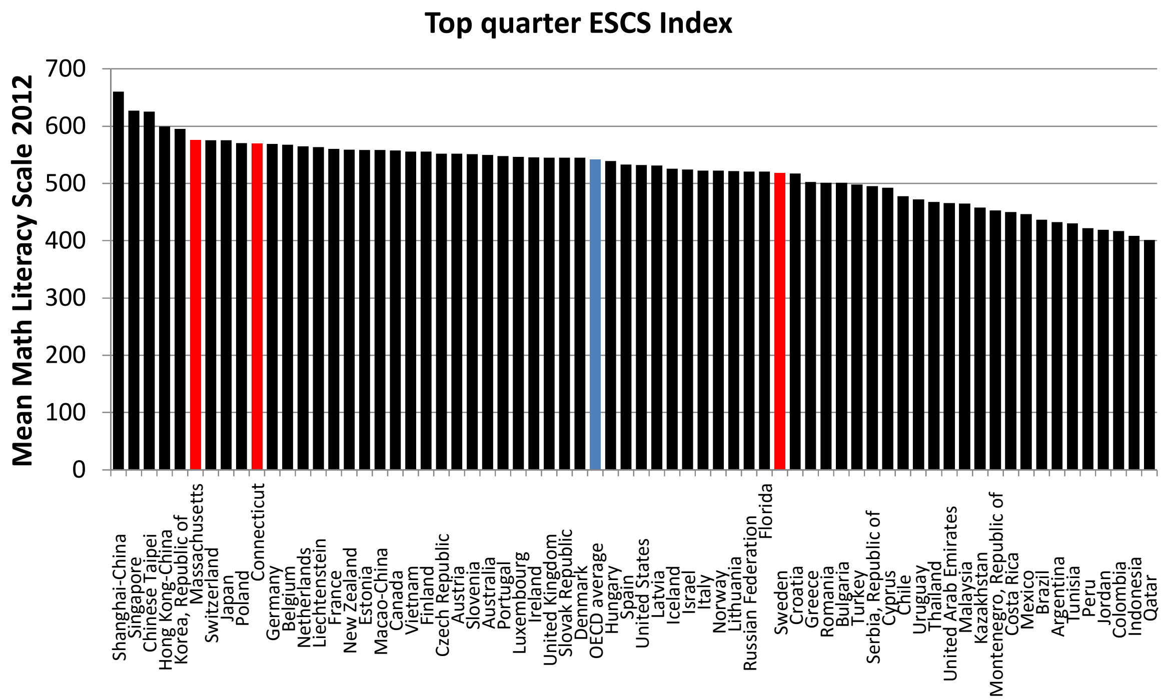
Figure 8. Select U.S. States and National PISA Comparisons [All Students]



NOTE: The PISA index of economic, social and cultural status (ESCS) was created using student reports on parental occupation, the highest level of parental education, and an index of home possessions related to family wealth, home educational resources and possessions related to “classical” culture in the family home. The home possessions relating to "classical" culture in the family home included possessions such as works of classical literature, books of poetry, and works of art (e.g. paintings). The OECD average is the average of the national averages of the OECD member countries, with each country weighted equally. Standard error is noted by s.e. Italics indicate non-OECD countries and education systems. Results for Connecticut, Florida, and Massachusetts are for public school students only.
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2012.

Even within relatively high-performing states like Massachusetts, however, there is significant variation in the socio-economic status of the student population. Figure 9 compares the math literacy scale scores for children in the top socio-economic quartile within their jurisdiction. Top quartile students in Massachusetts are outperformed by only select Chinese enclaves, Singapore and Korea. Connecticut is not far behind. But Florida, where even the top quartile is less well off, performs similarly to The Russian Federation and Sweden.

Figure 9. Select U.S. States and National PISA Comparisons [Top National/State Quartile ESCS Index]



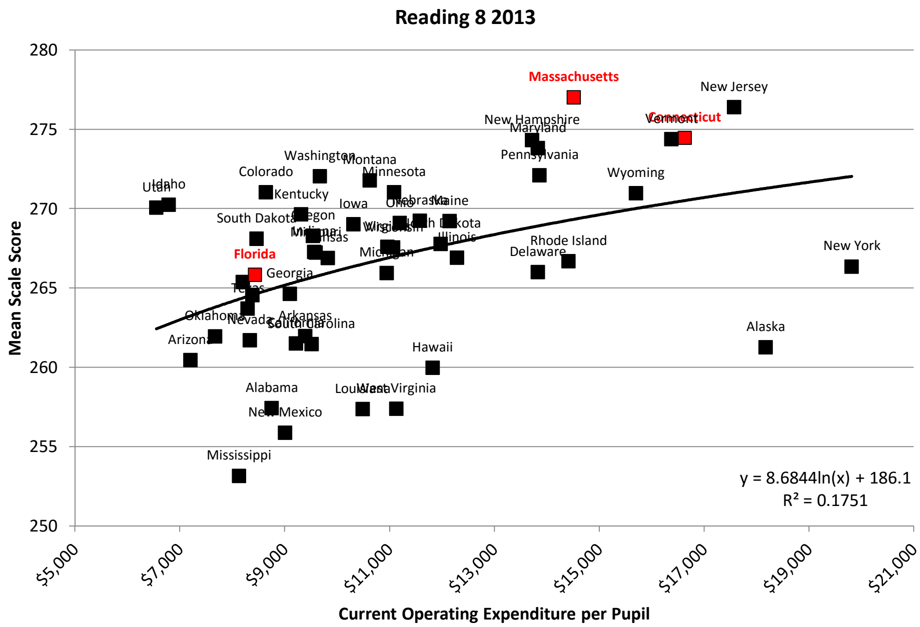
NOTE: The PISA index of economic, social and cultural status (ESCS) was created using student reports on parental occupation, the highest level of parental education, and an index of home possessions related to family wealth, home educational resources and possessions related to "classical" culture in the family home. The home possessions relating to "classical" culture in the family home included possessions such as works of classical literature, books of poetry, and works of art (e.g. paintings). The OECD average is the average of the national averages of the OECD member countries, with each country weighted equally. Standard error is noted by s.e. Italics indicate non-OECD countries and education systems. Results for Connecticut, Florida, and Massachusetts are for public school students only.
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2012.

The next few figures put these three US states into context among the other states using analyses similar to the previous simple cross-national pseudo-production function analyses. Again, these analyses simply compare spending and outcomes; they do not account for student, geographical, or structure differences. Figure 10 shows per pupil spending and National Assessment of Educational Progress (NAEP) Reading Grade 8 and Figure 11 shows per pupil spending and NAEP math Grade 8 for US states in 2013. Adopting the facile logic of the common cross-national comparison, one might assert that Alaska and New York are woefully inefficient, whereas Massachusetts and New Jersey are far more efficient, as are Utah and Idaho. Clearly though, there are as many missing pieces to this relationship as there were to the cross-national patterns. The math and reading pseudo-production function curves are similar, and the states are in similar positions.[14]

---

[14] In another useful exposition found on the Economic Industry USA View blog, http://economyindustryusa.blogspot.com/2011/01/relationship-between-education-spending.html, the author

Figure 10. The Unconditional Production Function for U.S. States [Reading Grade 8]

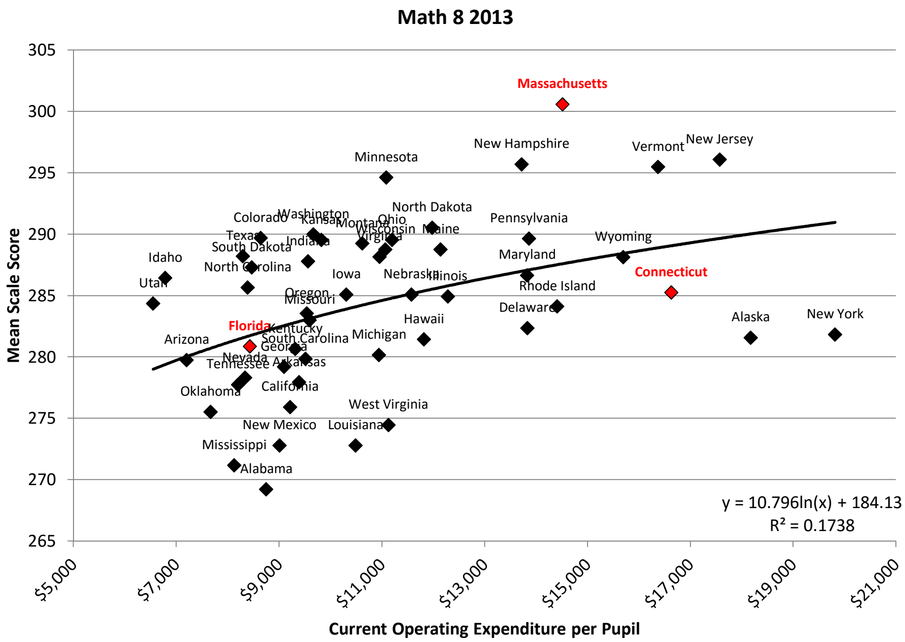


NAEP Data: http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx
State Spending Data: http://www.census.gov/govs/school/

---

combines student population adjustment, and US state and OECD spending and outcome measures to show the position of a handful of US states placed in international context.

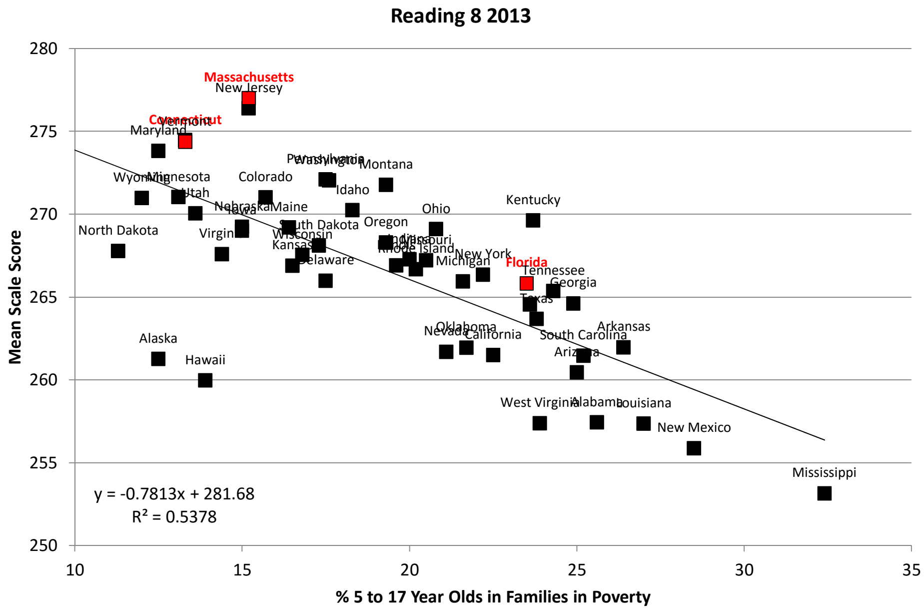
Figure 11. The Unconditional Production Function for U.S. States [Math Grade 8]



NAEP Data: http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx
State Spending Data: http://www.census.gov/govs/school/

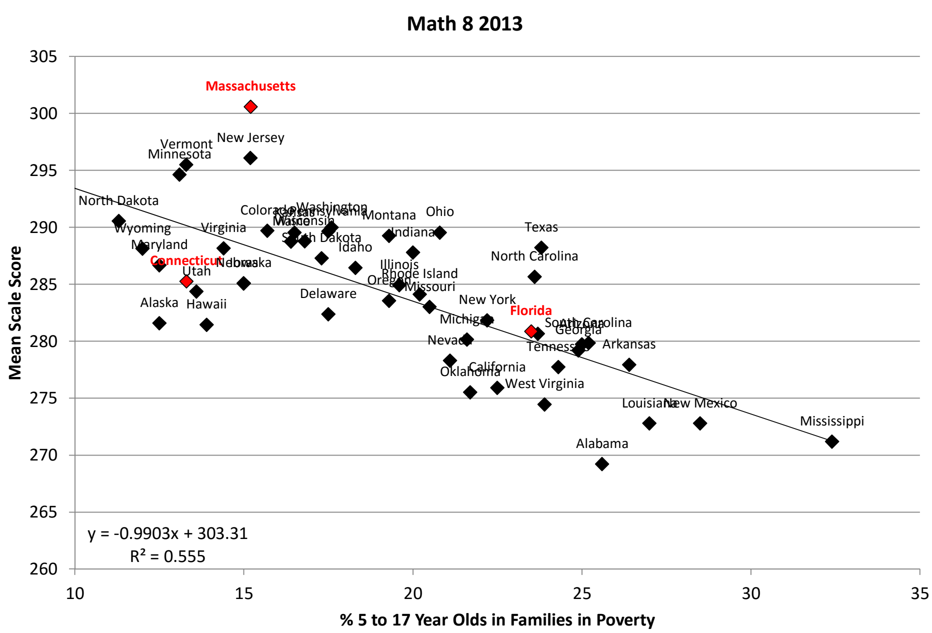
One of the biggest overlooked factors that simultaneously influences both state spending levels and state average outcome levels is the economic status of families and children across states. Figures 12 and 13 show the relationship between state child poverty rates and state average scale scores. Like the cross-national relationship, and even more so, these relationships are strong, and cannot possibly be ignored in making judgments about the relative efficiency of state systems. Connecticut and Massachusetts are relatively low-poverty, high-performing states; Florida, in contrast, is a higher-poverty, lower-performing state that also spends much less on its schools.

Figure 12. Poverty and NAEP Outcomes (Reading Grade 8 2013)



NAEP Data: http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx
Poverty Data: http://www.census.gov/did/www/saipe/data/statecounty/data/index.html

Figure 13. Poverty and NAEP Outcomes (Math Grade 8 2013)



NAEP Data: http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx
Poverty Data: http://www.census.gov/did/www/saipe/data/statecounty/data/index.html

As noted above, Massachusetts and Connecticut compare favorably with high performing OECD nations, whereas Florida does not. But this finding is neither sufficient basis for lauding the achievements of Massachusetts and Connecticut, nor for condemning the failures of Florida, at least with respect to relative efficiency. A model that accounts for differences in student characteristics is certainly an improvement on simple cost/outcome comparisons; however, it is still inadequate to the task of measuring efficiency.

5.0 Comparability of Input and Outcome Measures Matters

A major shortcoming of analyses purporting to evaluate cross-national relative efficiency of education systems regards the measurement of fiscal inputs to the education system. PISA, as well as other assessments including the Trends in International Mathematics and Science Study (TIMSS), provides some reasonable standardization of outcome measures. But national education systems vary widely. Their governance varies widely. The scope of services and related expenditures covered under the umbrella of "education" vary widely. If one wished to do a legitimate evaluation of the *relative efficiency of producing math and reading outcomes,* then one would have to precisely identify the services intended to produce those outcomes, and isolate the expenditures on those services. Generously, if we assume that the *core services* within ours and other national education systems have, as a central objective, improving math and reading outcomes, then we could at least focus on those "core services" whatever they may be.

But even that task is complicated in the human and capital resource intensive process of delivering, at large scale, public (and publicly-subsidized private) primary and secondary education services. Across nations, governance models and financial reporting systems, there exists:

- Inconsistent governance & expenditure of employee health and pension benefits,
- Inconsistent governance & expenditure of related health services and other disability services,
- Inconsistent "coverage" of various other educational (& related) service components (extra-curricular activities, transportation, food, etc.).

So, for example, in nations where employee health care and pensions are nationalized through separate agencies, education spending may appear reduced. Where disability services are covered through other agencies, education spending may appear reduced. The same for transportation, food service or various other activities embedded or separated from "schooling." Further, personal and family expenditures on supplemental programs may affect PISA or TIMSS outcomes, but may not be accounted for as school spending. Heyneman (2013) explains, for example, the amount of time and personal expense incurred by Korean families (not on the school expenses) to support math achievement.[15]

The commonly used OECD per pupil spending figures fail to accurately isolate comparable educational services and relevant, comparable expenditures on those services; they are, therefore, of minimal (if any) use for cross-national efficiency evaluation. The figures are also not appropriately adjusted to account for input price variation; in other words, the purchasing power of the education dollar varies across nations. The most appropriate adjustments would account for the competitive wage required to recruit and retain similarly qualified teachers, as constructed for US states (and labor markets within states) by Taylor and Fowler (2006). The relative competitiveness of teacher wages matters greatly to the overall quality of entrants (and stayers) to the workforce, and thus to the quality of schooling students receive (Baker, 2016). Finally, no attempt is made in the OECD per pupil spending figures to adjust for other geographic factors including population sparsity/remoteness (proportions of populations served under varied conditions). Put bluntly: if we wouldn't compare per pupil spending in Salina, Kansas and New York City without making the full range of appropriate adjustments, then we shouldn't compare Croatia and the US without doing so either.

Deconstructing Educational Services

While per pupil spending figures reported by OECD are especially problematic, some insights may be gained by comparing the core elements of educational service provision – specifically information on teacher compensation, teacher characteristics, teacher quantities (class sizes, overall staffing ratios) and teaching/school time. While the dollar value of teacher compensation should not be compared directly across contexts, the relative position of teachers on the labor market (compared to similarly educated, same age peers) can be a useful indicator of the adequacy of teacher wages for recruiting and retaining a quality workforce.

A substantial body of literature validates the conclusion that teachers' overall wages and relative wages affect the quality of those who choose to enter the teaching profession, and whether

[15] http://educationnext.org/obvious-flaws-obviate-new-education-efficiency-index/ also discussed Korean supplemental spending

they stay once they get in. For example, Murnane and Olson (1989) found that salaries affect the decision to enter teaching and the duration of the teaching career, while Figlio (1997, 2002) and Ferguson (1991) concluded that higher salaries are associated with more qualified teachers. In addition, more recent studies have tackled the specific issues of relative pay noted above. Loeb and Page (2000) showed that:

> "Once we adjust for labor market factors, we estimate that raising teacher wages by 10 percent reduces high school dropout rates by 3 percent to 4 percent. Our findings suggest that previous studies have failed to produce robust estimates because they lack adequate controls for non-wage aspects of teaching and market differences in alternative occupational opportunities."

In short, while salaries are not the only factor involved, they do affect the quality of the teaching workforce, which in turn affects student outcomes.

Research on the flip side of this issue—evaluating spending constraints or reductions—reveals the potential harm to teaching quality that flows from leveling down or reducing spending. For example, Figlio and Rueben (2001) note: "Using data from the National Center for Education Statistics we find that tax limits systematically reduce the average quality of education majors, as well as new public school teachers in states that have passed these limits."

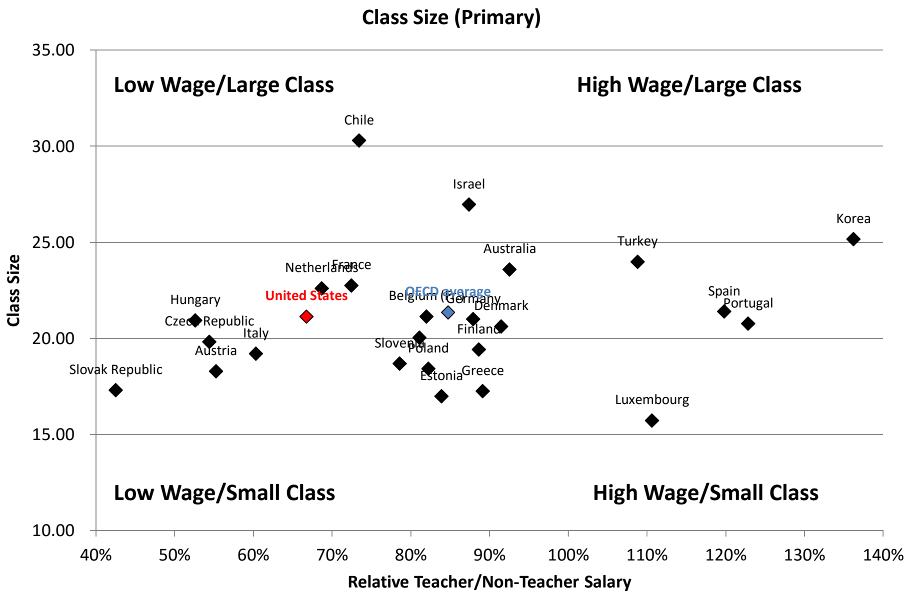
Salaries also play a potentially important role in improving the *equity* of student outcomes. While several studies show that higher salaries relative to labor market norms can draw higher-quality candidates into teaching, the evidence also indicates that relative teacher salaries across schools and districts may influence the distribution of teaching quality. For example, Ondrich, Pas and Yinger (2008) find that "teachers in districts with higher salaries relative to non-teaching salaries in the same county are less likely to leave teaching and that a teacher is less likely to change districts when he or she teaches in a district near the top of the teacher salary distribution in that county." Finally, a recent study by Britton and Proper (2015) on schools in England found that a 10 percent increase to the wage gap between teachers and non-teachers was associated with a 2% reduction in assessed outcomes.

Notably, while relative wages are a useful indicator for understanding education expenditures, it would be as fallacious to draw assertions from a simple relationship between nominal teacher salaries and PISA scores as it is to use simplistic spending/outcome comparisons.[16] Even within the personnel component of school budgeting, wages are only part of what drives costs. Personnel expenses in education are a function of staffing prices (wages) and staffing quantities, typically reflected in pupil to teacher ratios and class sizes. Thus, it is useful to explore the two together. In schools, districts, states or nations with relatively more total resources to spend, there exists greater flexibility to provide both competitive wages and smaller class sizes. As resources become scarce, tradeoffs become necessary. Because wages of non-teachers tend to be higher in wealthy countries, this creates upward pressure on education spending to maintain competitive compensation for teachers while maintaining palatable class sizes. Domestic research in the US indicates benefits of smaller class sizes in lower grades (Baker, 2016). Woessman and West (2006) find that cross-nationally, smaller class sizes seem more important where teacher wages are low.

Figure 14 shows the relationship (or, more accurately, the lack of a relationship) between relative teacher compensation and class sizes at the primary level. Nations to the left have relatively non-competitive teacher wages; those to the right have relatively competitive teacher wages, outpacing those of non-teachers (in Korea, for example, teacher wages are 30% higher than non-teacher wages). Teacher wages in the US are relatively non-competitive, both as a function of relatively low absolute wage and as a function of high (and more rapidly growing) non-teacher wage (Allegreto, Corcoran & Mishel, 2010). Meanwhile, class size, the vertical dimension in the figure, in US primary schools is relatively average. Class sizes in Chile, which also has relatively low teacher wages, are quite high. Korean primary class sizes are somewhat larger, but teacher wages much higher.

[16]For a particularly egregious application, see: http://www.economist.com/news/international/21616978-higher-teacher-pay-and-smaller-classes-are-not-best-education-policies-new-school

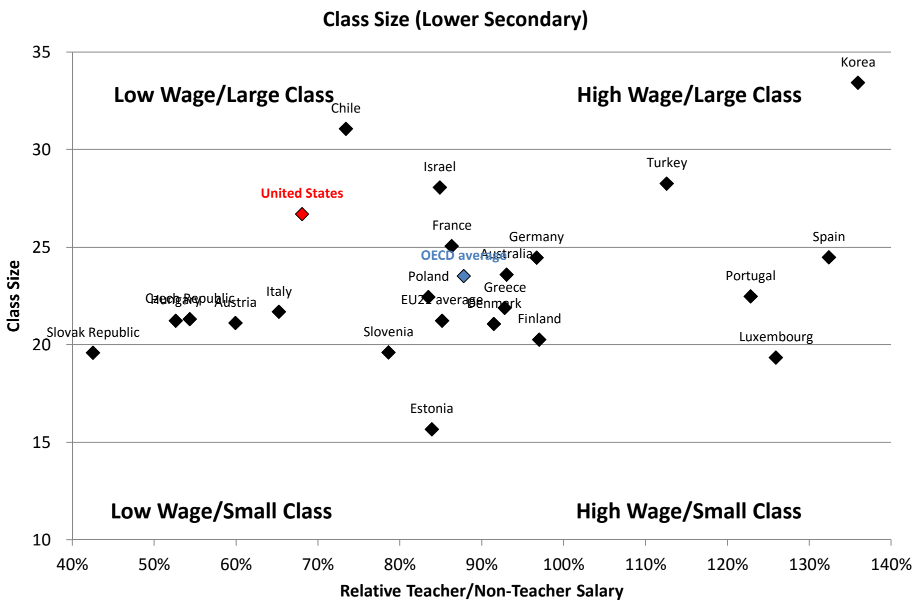
Figure 14. Class Size (Primary) and Relative Teacher Wages



OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
Table D2.1. Average class size, by type of institution and level of education (2012) &
Table D3.2. Teachers' salaries relative to earnings for full-time, full-year workers with tertiary education (2012)

Figure 15 shows the relationship between lower secondary class sizes and relative salaries. Again, US teacher wages are relatively non-competitive. In lower secondary grades, US class sizes are relatively large. The seemingly high per pupil spending figure of the US does not, therefore, translate into either competitive wages and small classes; instead, the US has relatively non-competitive wages and average to larger than average class sizes. These indicators provide a more accurate characterization of our investment in schooling, relative to other nations, than comparisons of nominal dollar inputs.

Figure 15. Class Size (Lower Secondary) and Relative Teacher Wages



OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
Table D2.1. Average class size, by type of institution and level of education (2012) &
Table D3.2. Teachers' salaries relative to earnings for full-time, full-year workers with tertiary education (2012)

The apparently high spending level of the US coupled with the apparently modest class sizes and low relative wages might lead one to assume that the US simply isn't getting resources into the classroom; that in fact, these findings do reveal inefficient spending. On the one hand, it may be the case that classroom teacher salary expenses are lower in the US for reasons discussed previously: that the US educational expenditure embeds far more than just classroom salary expenses, including pension and health benefits, transportation, food and extra-curricular activities. But it is also important to understand that the US has low relative wages for teachers in a very high GDP context, where the wages of non-teachers are high. In other words, given high GDP and non-teacher wages, higher spending may be needed in the US to achieve both competitive wages and reasonable class sizes.

Table 2 calculates the classroom salary per pupil, based on actual salary and class size data, for primary and lower secondary grades and evaluates classroom salaries per pupil as a share of per pupil spending. Table 2 shows that the US share of spending in classroom salaries per pupil is relatively low in the primary level and lowest for lower secondary, but not substantially out of line with other countries for which data were available.

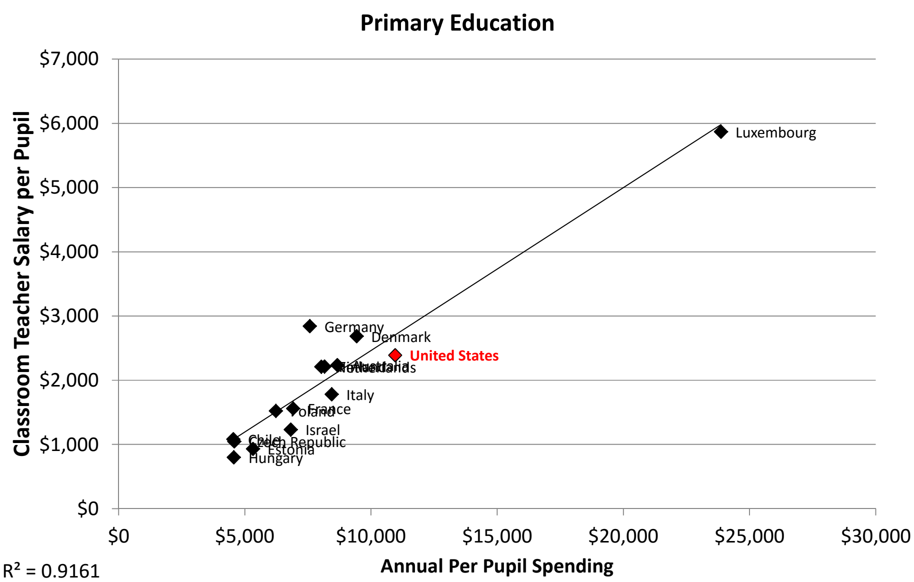
Table 2. Estimation of Classroom Shares of per Pupil Expenditure

	Class Size [1]		Average Actual Salary [2]		Per Pupil Expenditure [3]		Classroom Salary per Pupil		Classroom Salary Percent	
	Primary education	Lower secondary education	Primary education	Lower secondary education	Primary education	Lower secondary education	Primary education	Lower secondary education	Primary education	Lower secondary education
Australia	23.58	23.58	$52,659	$52,928	$8,671	$10,689	$2,234	$2,245	26%	21%
Chile	30.28	31.05	$32,728	$32,728	$4,551	$4,494	$1,081	$1,054	24%	23%
Czech Republic	19.82	21.30	$20,743	$20,724	$4,587	$7,730	$1,047	$973	23%	13%
Denmark	20.62	21.06	$55,330	$55,330	$9,434	$10,971	$2,683	$2,627	28%	24%
Estonia	16.98	15.66	$15,803	$15,803	$5,328	$6,009	$930	$1,009	17%	17%
Finland	19.41	20.25	$42,910	$46,968	$8,159	$12,545	$2,211	$2,319	27%	18%
France	22.75	25.05	$35,432	$42,217	$6,917	$9,668	$1,558	$1,685	23%	17%
Germany	20.99	24.46	$59,598	$65,545	$7,579	$9,247	$2,839	$2,679	37%	29%
Hungary	20.94	21.23	$16,731	$16,731	$4,566	$4,709	$799	$788	18%	17%
Israel	26.95	28.06	$33,181	$32,228	$6,823		$1,231	$1,149	18%	
Italy	19.19	21.68	$34,162	$36,947	$8,448	$8,686	$1,780	$1,704	21%	20%
Luxembourg	15.73	19.32	$92,248	$104,991	$23,871	$16,125	$5,866	$5,433	25%	34%
Netherlands	22.60		$49,924	$59,469	$8,036	$12,031	$2,209		27%	
Poland	18.41	22.43	$27,986	$28,409	$6,233	$5,995	$1,520	$1,267	24%	21%
United States	21.13	26.69	$50,494	$51,487	$10,958	$12,338	$2,390	$1,929	22%	16%

OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
[1] Table D2.1. Average class size, by type of institution and level of education (2012) &
[2] Table D3.4. Average actual teachers' salaries (2012)
[3] Table B1.1a. Annual expenditure per student by educational institutions for all services (2011)

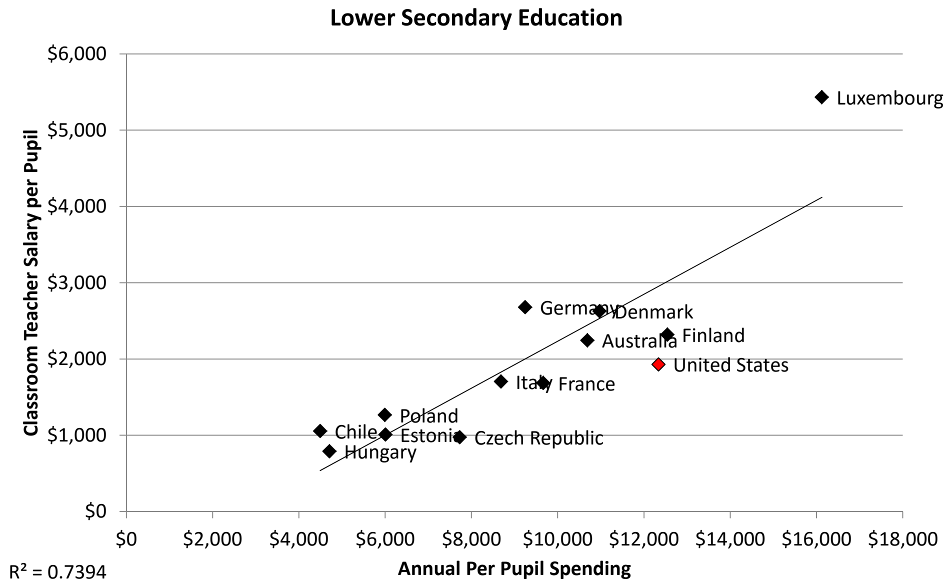
Figure 16 and Figure 17 provide some visual context for the values reported in Table 2. Notably, the relationship across countries between overall spending levels and classroom spending levels is strong and linear – indicating that countries spending more overall are spending more on teacher salaries per classroom. This includes the United States, which falls very near the trendline for Primary Education, but below the trendline (near Finland, however) for lower secondary education. Overall, the pattern is looser for lower secondary education (although the correlation is still strong), perhaps suggestive of more varied ranges of services and activities provided at this level.

Figure 16. Total per Pupil & Estimated Classroom Salary Spending (Primary)



OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
Table D2.1. Average class size, by type of institution and level of education (2012) &
Table D3.2. Teachers' salaries relative to earnings for full-time, full-year workers with tertiary education (2012)
Table D3.4. Average actual teachers' salaries (2012)
Table B1.1a. Annual expenditure per student by educational institutions for all services (2011)

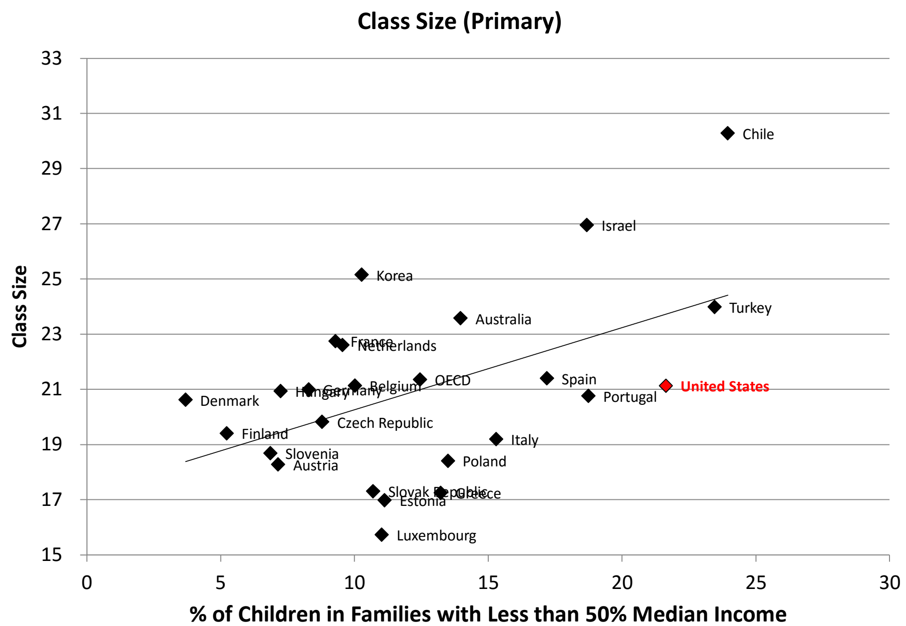
Figure 17. Total per Pupil & Estimated Classroom Salary Spending (Lower Secondary)



OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
Table D2.1. Average class size, by type of institution and level of education (2012) &
Table D3.2. Teachers' salaries relative to earnings for full-time, full-year workers with tertiary education (2012)
Table D3.4. Average actual teachers' salaries (2012)
Table B1.1a. Annual expenditure per student by educational institutions for all services (2011)

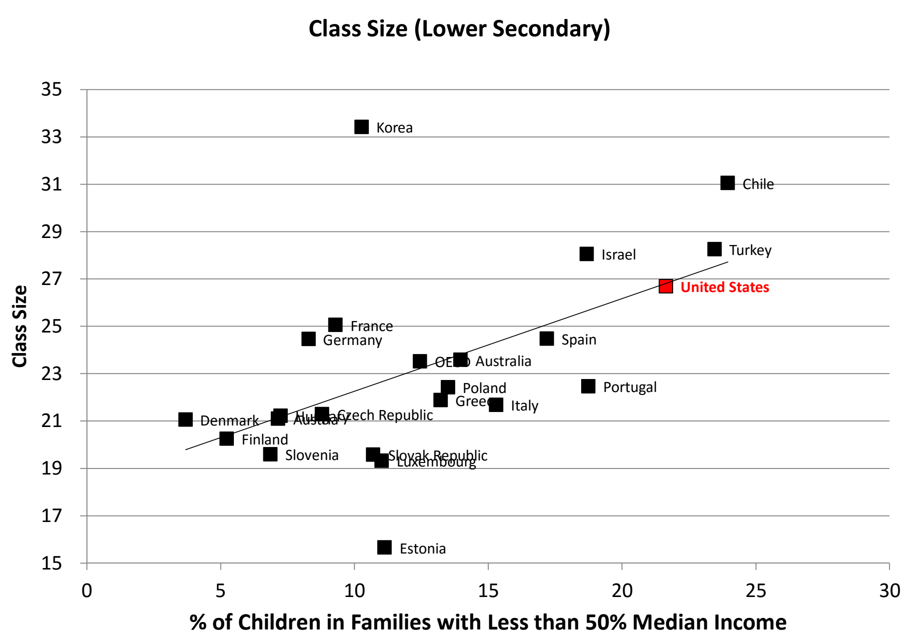
Figure 18 and Figure 19 show the relationship between relative poverty and class sizes across nations. Chile has particularly large class sizes and high relatively poverty. It also has lower than average relative wages and performs poorly on PISA (Figure 4). Israel and Turkey also have high poverty and relatively large class sizes but more competitive wages than Chile or the US. For lower secondary grades, the US joins these countries, having the combination of high poverty and large class sizes, and joins Chile in also having low relative wages. While still a limited view on resources and context, one would not likely expect superior international test scores from a nation with non-competitive compensation for its teachers, average to large classes, and high poverty. Perhaps the United States' seemingly mediocre PISA scores are, in fact, in line with expectations, given our inputs and context. Perhaps we are not the model of inefficiency, but rather, about average: in line with expectations, but nothing more and nothing less.

Figure 18. Class Size (Primary) and Relative Poverty



OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
Table D2.1. Average class size, by type of institution and level of education (2012) &
OECD Relative Poverty: Source: Provisional data from OECD Income distribution and poverty database (www.oecd.org/els/social/inequality).

Figure 19. Class Size (Lower Secondary) and Relative Poverty



OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en
Table D2.1. Average class size, by type of institution and level of education (2012) &
OECD Relative Poverty: Source: Provisional data from OECD Income distribution and poverty database (www.oecd.org/els/social/inequality).

6.0 Conclusions & Implications

What does this all mean? First and foremost, we can say with some confidence that existing expositions of US inefficiency based on OECD national spending data and PISA scores are so lacking in methodological rigor that they are of little if any value in public discourse or for informing national education policies. Second, it is unlikely that we could ever obtain data of sufficient precision, accuracy and comparability to meet the demands of more legitimate efficiency modeling for cross-national, intercontinental analyses. But that does not mean we can't learn anything at all from available data; as long as we deal with them cautiously, understand that we are viewing moment-in-time snapshots of limited measures, and realize the extent of what's missing from any such cross-national descriptive analysis, there are actually important insights to be gained from appropriate analyses of the international data.

Among other things, the OECD per pupil spending measure, as incomparable as it is, shows that the US may have higher per pupil spending than many nations, but falls right in line with expectations for nations of similar GDP per capita. The US is both a high spending and high GDP country, but some of that high education spending may be a function of the scope of services and expenses included under the education umbrella in the US. We also know that despite its seemingly high spending levels, the United States' teacher wages lag with respect to other professions, and the wage lag is not a result of providing relatively smaller class sizes. In fact, our primary class sizes are average and lower secondary class sizes large. Our wage lag is, to an extent, a function of high non-teaching wages (related to our high GDP per capita), necessarily making it more expensive to recruit and retain a high quality teacher labor force. To summarize: the US is faced with a combination of seemingly high education expense, but non-competitive compensation for its teachers, average to large classes, and high child poverty. Again, it's hard to conceive how such a combination would render the US comparable in raw test scores to low-poverty nations like Korea or Finland, or small, segregated, homogeneous enclaves like Singapore or Shanghai[17].

If there exists any possible example of classic over-allocation (as usually ascribed to the US), it might be found in Luxembourg, where poverty is very low, spending is high, class sizes are very small and teacher wages are very competitive. It would be difficult for PISA scores to line up with this full set of contextual and resource advantages. But this finding for Luxembourg comports with

[17] Shanghai in particular has several mitigating factors that make comparing its scores to other nations highly suspect; see: http://www.brookings.edu/research/papers/2013/10/09-pisa-china-problem-loveless

public finance literature regarding inefficiency: those who can afford to spend more are more likely to spend "inefficiently," at least so far as can be measured in outcomes that only focus on math and language arts test scores.

Finally, it is equally important to understand the magnitude and heterogeneity of the US education system in the context of OECD comparisons, which mainly involve more centralized and much smaller education systems. Lower poverty, higher spending states that have been included in international comparisons, like Massachusetts and Connecticut, do quite well, while lower spending higher poverty states like Florida do not. This unsurprising finding, however, also tells us little about relative efficiency, and provides little policy guidance for how we might make Florida more like Massachusetts, other than by waving a wand and making it richer, more educated and perhaps several degrees colder.

References

Afonso, A., & St Aubyn, M. (2005). Non-parametric approaches to education and health efficiency in OECD countries. *Journal of Applied Economics*, *8*(2), 227-246.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Romain Wacziarg, R. (2002). *Discussion Paper Number 1959; Fractionalization.* Harvard Institute of Economic Research: Cambridge, MA.

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus? *Economics of Education Review*, *21*(3), 245-262.

Baker, B. D. (2016). Money Still Matters: An update. *Albert Shanker Institute.*

Baker, B. D. (2012). Revisiting the Age-Old Question: Does Money Matter in Education?. *Albert Shanker Institute.*

Baker, B. D. (2011). Exploring the sensitivity of education costs to racial composition of schools and race-neutral alternative measures: A cost function application to Missouri. *Peabody Journal of Education*, *86*(1), 58-83.

Baker, B. D. (2005). The emerging shape of educational adequacy: From theoretical assumptions to empirical evidence. *Journal of Education Finance*, 259-287.

Baker, B., & AIR, J. L. (2014). Educational Equity, Adequacy, and Equal Opportunity in the Commonwealth: An Evaluation of Pennsylvania's School Finance System.

Baker, B. D., Sciarra, D. G., & Farrie, D. (2015). Is school funding fair? A national report card (fourth edition). Education Law Center: Newark, NJ. http://www.schoolfundingfairness.org

Baker, B. D., Taylor, L., Levin, J., Chambers, J., & Blankenship, C. (2013). Adjusted Poverty Measures and the Distribution of Title I Aid: Does Title I Really Make the Rich States Richer? *Education*, *8*(3), 394-417.

Bessent, A. M., & Bessent, E. W. (1980). Determining the Comparative Efficiency of Schools through Data Envelopment Analysis, *Education Administration Quarterly*, *16*(2), 57-75.

Robert Bifulco & William Duncombe (2000) Evaluating School Performance: Are we ready for prime time? In

William Fowler (Ed) Developments in School Finance, 1999 – 2000. Washington, DC: National Center for

Education Statistics, Office of Educational Research and Improvement.

Robert Bifulco and Stewart Bretschneider (2001) Estimating School Efficiency: A comparison of methods using simulate data. *Economics of Education Review* 20

Borge, L. E., Falch, T., & Tovmo, P. (2008). Public sector efficiency: the roles of political and budgetary institutions, fiscal capacity, and democratic participation. Public Choice, 136(3-4), 475-495.

Britton, J., & Propper, C. (2016). Teacher pay and school productivity: Exploiting wage regulation. *Journal of Public Economics*, *133*, 75-89.

Carnoy, M., & Rothstein, R. (2013). What do international tests really show about US student performance. *Economic Policy Institute*, *28*.

T. A. Downes and D. N. Figlio, "Do Tax and Expenditure Limits Provide a Free Lunch? Evidence on the Link Between Limits and Public Sector Service Quality," *National Tax Journal* 52, no. 1 (1999): 113-128.

Coley, R. J., & Baker, B. (2013). *Poverty and education: Finding the way forward*. Educational Testing Service Center for Research on Human Capital and Education.

Duncombe, W., & Yinger, J. (2008). Measurement of cost differentials. *Handbook of research in education finance and policy*, 238-256.

Duncombe, W., Miner, J., & Ruggiero, J. (1997). Empirical Evaluation of Bureaucratic Models of Inefficiency, *Public Choice*, *93*(1), 1-18.

Duncombe, W., & Bifulco, R. (2002). *Evaluating School Performance: Are we ready for prime time?* In William J. Fowler, Jr. (Ed.), Developments in School Finance, 1999–2000, NCES 2002–316.Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Duncombe, W. and Yinger, J.M. (2007) A Comment on School District Level Production Functions Estimated Using Spending Data. Maxwell School of Public Affairs. Syracuse University.

Eom, T. H., & Rubenstein, R. (2006). Do State-Funded Property Tax Exemptions Increase Local Government Inefficiency? An Analysis of New York State's STAR Program. *Public Budgeting and Finance*, *Spring*, 66-87.

James D. Fearon, J.D. (2003). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth, 8*(2), 195-222.

D. N. Figlio, "Can Public Schools Buy Better-Qualified Teachers?," *Industrial and Labor Relations Review* 55 (2002): 686-699; D. N. Figlio "Teacher Salaries and Teacher Quality," *Economics Letters* 55 (1997): 267-271;

D. N. Figlio and K. Rueben, "Tax Limits and the Qualifications of New Teachers," *Journal of Public Economics* (April 2001): 49-71.

R. Ferguson, “Paying for Public Education: New Evidence on How and Why Money Matters,” *Harvard Journal on Legislation* 28, no. 2 (1991): 465-498.

Ganzeboom, H. B. (2010, May). A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002-2007. In Annual Conference of International Social Survey Programme, Lisbon (Vol. 1).

Glewwe, P. W., Hanushek, E. A., Humpage, S. D., & Ravina, R. (2011). *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010* (No. w17554). National Bureau of Economic Research.

Gronberg, T. J., Jansen, D. W., & Taylor, L. L. (2012). The relative efficiency of charter schools: A cost frontier approach. *Economics of Education Review*, *31*(2), 302-317.

Grosskopf, S., Hayes, K. J., & Taylor, L. L. (2014). Applied efficiency analysis in education. *Economics and Business Letters*, *3*(1), 19-26.

Grosskopf, S., Hayes, K. J., Taylor, L. L., & Weber, W. (2001). On the Determinants of School District Efficiency: Competition and Monitoring. *Journal of Urban Economics*, *49*, 453-478.

Haddad, H. (2014) A Finnish-ed Model for National Education. Berkeley Political Review, Jan. 12. http://bpr.berkeley.edu/2014/10/29/a-finnish-ed-model-for-national-education/

Heyneman, S. P. (2013). The international efficiency of American education: The bad and the not-so-bad news. *PISA, power, and policy: The emergence of global educational governance*, 279-302.

Jackson, K., Johnson, R.C., Persico, C. (2015) Money Does Matter After All. Education Next. http://educationnext.org/money-matter

S. Loeb and M. Page, “Examining the Link Between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-Pecuniary Variation,” *Review of Economics and Statistics* 82, no. 3 (2000): 393-408.

R. J. Murnane and R. Olsen, “The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching: Evidence from Michigan,” *Review of Economics and Statistics* 71, no. 2 (1989): 347-352.

OECD (2012) Does Money Buy Strong Performance in PISA? PISA in Focus. http://www.oecd.org/pisa/pisaproducts/pisainfocus/49685503.pdf

OECD (2010) Socio-economic indicators and the relationship with performance in reading, in *PISA 2009 Results: What Students Know and Can Do*, OECD Publishing, Paris.

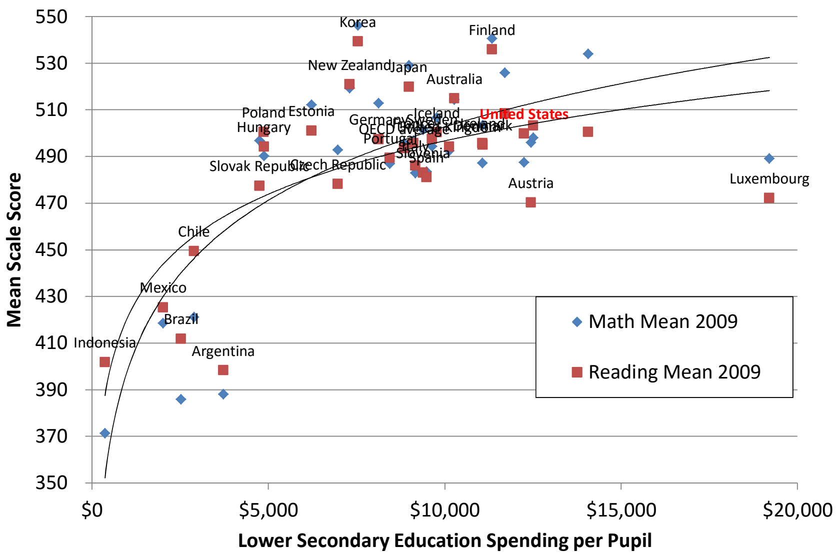
J. Ondrich, E. Pas and J. Yinger, "The Determinants of Teacher Attrition in Upstate New York," *Public Finance Review* 36, no. 1(2008): 112-144.

Rothstein, R. (2011). Fact-Challenged Policy. Policy Memorandum# 182. *Economic Policy Institute*.

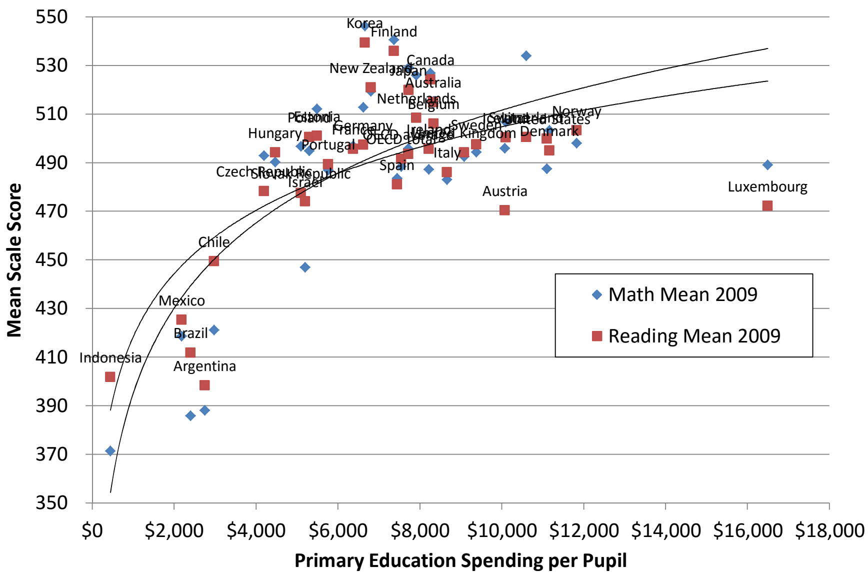
Taylor, L. L. (2006). Comparable wages, inflation, and school finance equity. *Education*, *1*(3), 349-371.

Wößmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, *50*(3), 695-736.

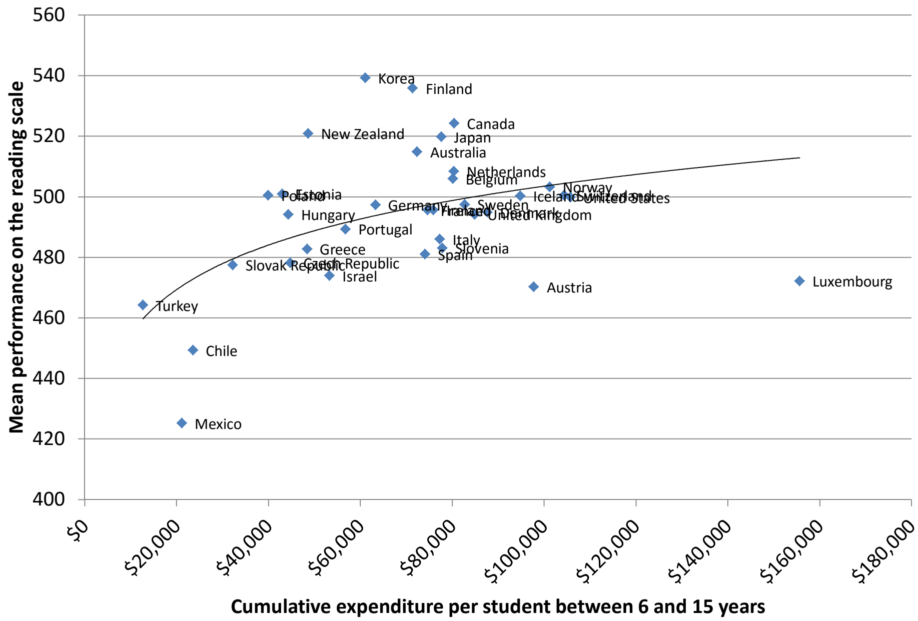
Appendix A. Alternative Unconditional, Nominal Production Curves



OECD (2012), Education at a Glance 2012: OECD Indicators , OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en
Indicator B1.1a: http://dx.doi.org/10.1787/888932665867
OECD (2010), PISA 2009 Results: What Students Know and Can Do – Student
Performance in Reading, Mathematics and Science (Volume I) http://dx.doi.org/10.1787/9789264091450-en



OECD (2012), Education at a Glance 2012: OECD Indicators , OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en
Indicator B1.1a: http://dx.doi.org/10.1787/888932665867
OECD (2010), PISA 2009 Results: What Students Know and Can Do – Student
Performance in Reading, Mathematics and Science (Volume I) http://dx.doi.org/10.1787/9789264091450-en



OECD (2010), Table I.2.20. Socio-economic indicators and the relationship with performance in reading, in *PISA 2009 Results: What Students Know and Can Do*, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789264091450-table76-en