Deficiencies & Misinterpretations in the Center for American Progress Method for Measuring and Comparing School District Return on Investment

Bruce Baker
Associate Professor
Department of Educational Theory, Policy and Administration
Rutgers, The State University of New Jersey
Bruce.baker@gse.rutgers.edu

The Center for American Progress released its new Return on Investment (ROI) Index for K-12 public school districts of greater than 250 students this week.[1] I had the opportunity to provide advice on this project early on and occasionally thereafter, and I do believe that at least some involved had and still have the best intentions in coming up with a useful way to represent the information at hand. I'll get back to the validity of the information at hand in a moment.

First, the policy implications and proposals, or even the general findings presented in the report are not supported by the analysis, however well or poorly done. The suggestion that billions of dollars might be saved nationally at no loss to performance outcomes is little more than a back-of-the-napkin extrapolation based on crudely estimated indicators. This conclusion is a huge, unwarranted stretch and quite simply arrogant, ignorant and irresponsible.

The analyses conducted by CAP provide no reasonable basis for the claim that any and all low "rate of return" districts could simply replicate the behaviors of high "rate of return" districts and achieve the same or better outcomes at lower cost, saving the nation billions. The limitations of these methods, when applied in their best and most rigorous complete possible form, no less in this woefully insufficient and incomplete form, simply do not allow for such extrapolation.

Further, given the crudeness of the models and cost and need adjustments used in the analysis, it is inappropriate to make too much, if anything, of supposed differences in the characteristics of districts with good versus bad rate of return indices. One can do some fun hunting and pecking through the maps, but that's about it!

[1] http://www.americanprogress.org/issues/2011/01/educational_productivity/index.html

For example, one major finding is that districts with good ROI's spend less on administration. However, much more rigorous studies using better data and more appropriate models exploring precisely the same question have found the opposite.[2] The CAP ROI methods and results of those methods are insufficient to draw any conclusion in this regard.

There is little basis in this analysis that states need to provide fairer funding – except that funding does appear to vary within states. But how funding varies is not explored, though it might be and should be, as I discuss below. A more appropriate use of similar methods for evaluating funding fairness and the need to improve funding fairness can be found in "*Is School Funding Fair?*" a recent national report.[3] And yes, states do need to provide fairer funding.

And there is no basis for suggesting that fairer funding would be accomplished by student-based funding (weighted student funding). In my own recent research, I find that adoption of student based funding does not, in-and-of-itself, lead to substantive improvements to equity overall or the targeting of resources to higher need districts, schools or children:

> "I find first that widely reported WSF success stories provide no more predictable funding with respect to student needs than other large urban districts in the same state."[4]

Adoption of weighted student funding is particularly problematic within tight overall budget constraints. The CAP ROI report does not even make the most basic attempts to link this recommendation back to underlying disparities in funding across districts within states or identify how the approach would resolve those disparities.

Perhaps the most problematic boldly proclaimed conclusion of the report was that districts serving higher proportions of low income children are simply less efficient. This conclusion was framed from the manipulative *call-to-action* perspective that low-income children are simply trapped in inefficient districts - the logical conclusion being that those children must be provided choices (charters and/or vouchers) and funding for those inefficient districts cut, because it's wasted anyway.

The authors fail entirely to consider the distinct possibility that their method is simply revealing to them a systematic bias. I would argue that exploring this bias for alternative explanations is a very basic responsibility of researchers conducting such analysis. Where findings show a particular strong pattern of this type, it is critically

[2] http://bush.tamu.edu/research/workingpapers/ltaylor/The_65_Percent_Solution.pdf
[3] www.schoolfundingfairness.org
[4] http://epaa.asu.edu/ojs/article/view/5

important to question that pattern and whether that pattern is a result of bias in the methods or models - important overlooked factors - or whether it is real.

This single bias – failure to fully account for poverty related costs – is pervasive throughout the entire CAP ROI analysis. There is a systematic relationship between poverty and supposed inefficiency in many states in the analysis and in some versions of the ROI more than others. That bias exists in states where the state has provided additional resources to higher poverty districts, making them higher spending on average, and that bias exists even in states where spending per pupil is systematically lower in higher poverty districts. As a result, each map and scatterplot in the analysis must be viewed carefully with an understanding of the pervasive, uncontrolled bias against higher poverty districts. A bias that results largely from failure to fully account for cost variation.

Exploring the ROI

Here, I provide a walk-through of the ROI indicators, and major missing components of those indicators, when compared against research literature on the topic of cost adjustment.

Basic ROI

In any of the ROI's there are two sides to the analysis. There are the student outcome measures and the per pupil spending measures. Within the per pupil spending measures, there are cost adjustments for the "cost" of meeting student population needs and cost adjustments for addressing regional differences in competitive wages for school personnel. The Basic ROI uses an approach similar to that used by Education Week in Quality Counts as a basis for calculating "cost adjusted spending per pupil."

Weighted Pupil Count = Enrollment + .4*Free Lunch Count + .4*ELL Count + 1.1*IEP Count

After using the weighted pupil count to generate a student need adjustment, CAP uses the NCES Comparable Wage Index to adjust for regional variation in wages.[5] That is, they attempt to adjust for student needs, using a series of arbitrary weights, and for regional wage variation using a reasonable, macro-level adjustment.

The central problem with this approach is that it relies on setting rather arbitrary weights to account for the cost differences associated with poverty, ELL and special education. And in this case, CAP, like Ed Week, adopted weights nearer the lower bound

[5] Taylor, L. L., Glander, M. (2006). *Documentation for the NCES Comparable Wage Index Data File* (EFSC 2006-865). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

(if not well below that) claiming those low weights to be grounded in research literature, but that claim is a stretch at best and closer to a complete misrepresentation.

Adjusted ROI

For the adjusted ROI, CAP uses a regression equation which compares the actual spending of each district to the predicted spending of each district, given student population characteristics. Here's their equation:

ln(CWI adjusted ppe)= β0 + β1% free lunch + β2 % ELL+ β3 % Special Ed + ε

This method is reasonable for comparing how much districts spend relative to overall patterns of spending variation, but has little or nothing to do with adjusting for the costs of achieving comparable educational outcomes. That is, one can use a spending regression model to determine if a state, on average, spends more on high poverty than on low poverty districts. But a spending differential is not a cost factor. It's useful, and has meaning, but not the right meaning for this context. One would need to determine how much, more or less, needs to be spent in order to achieve comparable outcomes.

So, for example, one might apply this approach to determine that within a state, higher poverty districts spend less on average than lower poverty districts. But, the way the approach is used here, the negative or regressive poverty effect would become the cost adjustment. That is, it would be assumed that higher poverty districts have lower costs than lower poverty ones. This is wrong. They have lower spending, but they still most likely have higher costs of achieving constant educational outcomes. Including outcomes and holding outcomes constant is the key – AND MISSING – step toward using this approach to adjust for costs.

Further, the overly simplistic equation above completely ignores significant factors that do affect cost differences and/or spending differences across districts, such as economies of scale[6] and population sparsity as well as more fine grained differences in teacher wages needed to recruit or retain comparable teachers across districts of differing characteristics within the same labor market.[7]

Predicted Efficiency

Finally, there's the predicted efficiency regression equation, which attempts to generate a predicted achievement level based on a) cost adjusted per pupil spending, b) free lunch, ELL and special education shares. This one, like the others, doesn't attempt to

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

[7] Baker, B.D. (2005) The Emerging Shape of Educational Adequacy: From Theoretical Assumptions to Empirical Evidence. *Journal of Education Finance* 30 (3) 277-305

adjust for economies of scale or sparsity and suffers from numerous potential problems with figuring out how and why each district's actual performance differs from its predicted performance.

achievement = β0 + β1 ln(CWI adjusted ppe) + β2 % free lunch + β3 % ELL + β4 %Special Ed + ε

In this oversimplified production function approach, any individual district's actual outcomes could be much lower than predicted or much higher than predicted for any number of reasons, including the assumed *functional form*, or "shape" of the best fit line, as discussed below. It would appear from scanning through the findings that this particular indicator is most biased with respect to poverty.

Summary of what's missing or mis-specified

The table below summarizes the treatment of adjusted per pupil spending in each of the three ROI indices, with respect to what we know are the major cost factors that must be accounted for in any reasonable analysis of education spending data in relation to student outcomes.[8] Here, the basic conception of cost, and cost difference is "what are the differences in cost toward achieving comparable outcome objectives?" Cost cannot be estimated without an outcome objective.

First, I would argue that the selected weights in the Basic ROI are simply too low, especially in certain parts of the country.

Second, none of the models address economies of scale. CAP notes this, but in a section of the report most will never read. Instead, we'll all see the pretty maps that tell us that all of the rural districts in the upper Hudson Valley in NY State or in north central Pennsylvania are really, really inefficient.

Third, recall that the "adjusted ROI" model really doesn't control for cost at all, but rather for underlying spending variation, without respect for outcomes.

[8] Duncombe, W. and Yinger, J.M. (2008) Measurement of Cost Differentials. In H.F. Ladd & E. Fiske (eds) pp. 203-221. *Handbook of Research in Education Finance and Policy*. New York: Routledge.

Table 1

Cost Factor	Basic ROI	Adj. ROI	Predicted Efficiency	Research
Child Poverty	40% weight on Free Lunch Count (1.4)	No cost factor, only relative spending (no control for outcomes)	Variable in adj. performance model	100% weight on Free or Reduced Lunch or 150% weight on census poverty (2.0 to 2.5) Duncombe & Yinger (2005)
Limited English Proficiency	40% weight on ELL Counts (1.4)	No cost factor, only relative spending (no control for outcomes)	Variable in adj. performance model	100% weight on LEP/ELL share (2.0) Duncombe & Yinger (2005)
Disability a)High incidence (low severity) b)Low incidence (high severity)	110% weight on total disability count (2.1)	No cost factor, only relative spending (no control for outcomes)	Variable in adj. performance model	Average historical spending about 100% to 110% additional per child (2.1) Per pupil cost as much as 200% additional per child (3.1) Duncombe & Yinger (2005)
Regional Competitive Wage	YES (NCES ECWI)	YES (NCES ECWI)	YES (NCES ECWI)	NCES ECWI provides blunt geographic instrument for labor market level competitive wage Taylor and Glander (2006)
Local (Hedonic) Wage	NONE	NONE	NONE	Competitive wages vary as a function of local working conditions
Economies of Scale	Excluded districts <250 students	Excluded districts <250 students	Excluded districts <250 students	Perhaps largest overall cost factor Andrews, Duncombe & Yinger (2002), Baker (2005)
Sparsity	NONE	NONE	NONE	Related to scale

Regarding pupil need weights in particular, there exists at least some literature – the most rigorous and direct literature on the question – which suggests the need for much higher weights than those used by CAP. For example, Duncombe and Yinger note that in two versions of their models, in a study titled *How Much More Does a Disadvantaged Student Cost?*

> "Overall, this poverty weight ranges from 1.22 to 1.67 (x census poverty rate), the LEP weight ranges from 1.01 to 1.42, and the special education weight varies from 2.05 to 2.64."[9]

Across several models produced in this particular paper, one might come to a rounded weight on Census poverty of about 1.5 or weight on subsidized lunch rates of about 1.0 (100% above average cost, or 2x average, more than double the CAP weight), a weight on limited English proficient students around 1.0 and on special education students over 2.0 (slightly less than double the CAP weight).

[9] Duncombe, W., Yinger, J. (2005) How Much more Does a Disadvantaged Student Cost? *Economics of Education Review* 24 (5) 513-532 http://surface.syr.edu/cgi/viewcontent.cgi?article=1102&context=cpr

Other work by me, along with Lori Taylor and Arnold Vedlitz, done for the National Academies of Science, reviewing numerous studies also comes to higher average weights for children in poverty.[10]

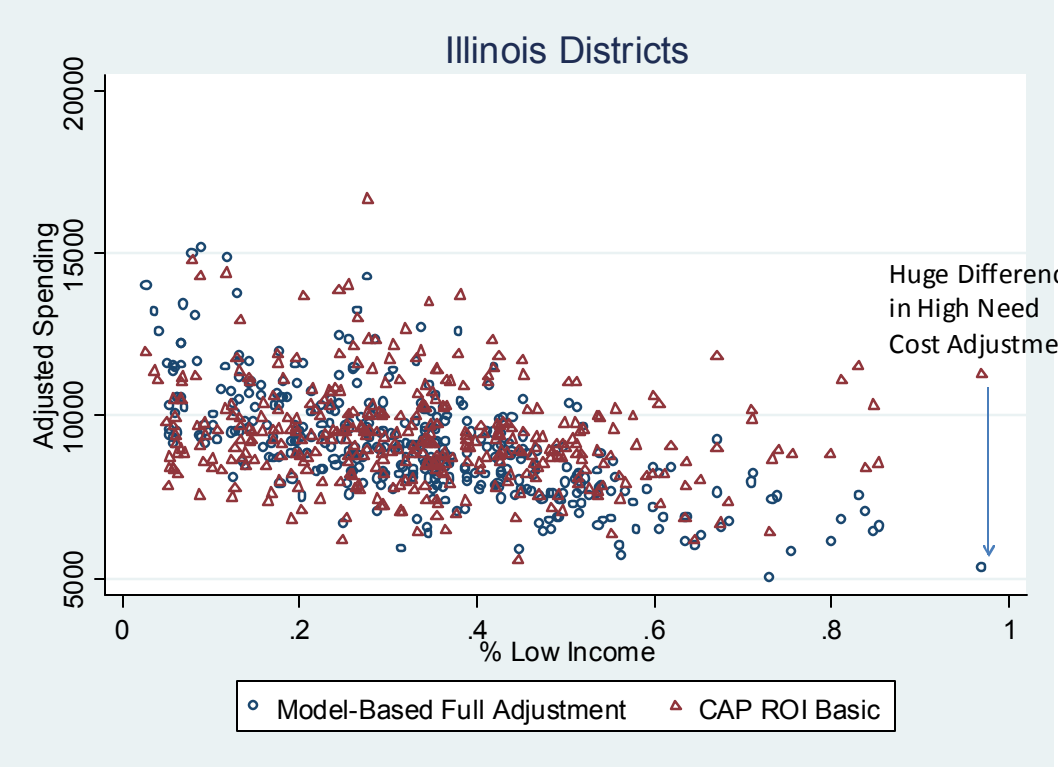
While one can quibble over the selection of "cost" weights from literature, the bigger concern is that the findings of the various ROIs indicate such a strong bias that any reasonable researcher would be obligated to explore further, and perhaps test out alternative research based weights as a way to reduce the bias. This process is a never ending battle and reducing the bias in one state can often make the distributions worse in another (because different patterns of poverty and distributions of ELL children lead to different appropriate weights in different settings – even within a state). If such problems persist, however, and a global method simply cannot be identified, the responsible thing to do is stop! Don't do it! Realize that it just doesn't work.

Comparing the Basic ROI to Modeled Costs in Illinois

Here is an example of how much a corrected cost adjustment might matter, when compared with the Basic ROI, using data from the state of Illinois. The scatterplot below includes one set of dots (red triangles) which represent adjusted operating expenditures of Illinois school districts using the Basic ROI weights. The other set of dots (blue circles) uses a cost index derived from a more thorough statistical model of the costs of achieving statewide average outcomes for Illinois school districts. For the highest poverty districts, the adjusted spending figures drop by $4,000 to $5,000 per pupil when the more thorough cost adjustment method is used. This is substantial, and important, since the ROI is much more likely to identify these districts as inefficient and might be used by state policy makers to argue that cuts to these districts are appropriate (when they clearly are not).

[10] Baker, B.D., Taylor, L.L., Vedlitz, A. (2008) Adequacy Estimates and the Implications of Common Standards for the Cost of Instruction. *National Research Council*. http://www7.nationalacademies.org/CFE/Taylor%20Paper.pdf

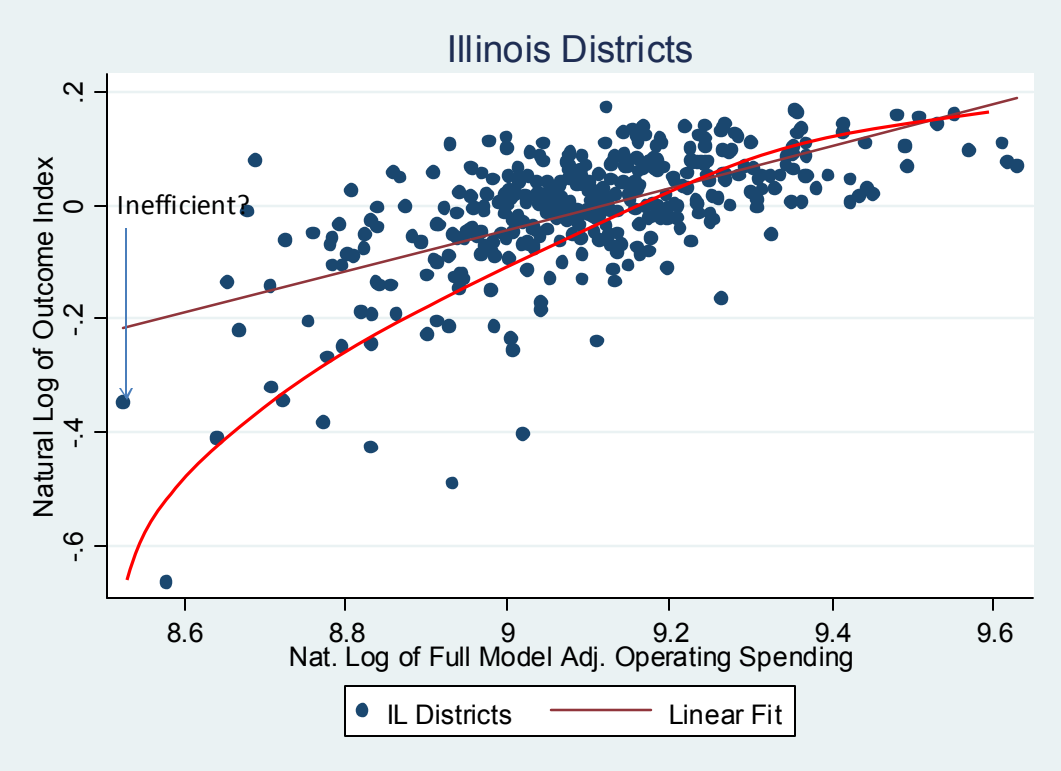
Figure 1



How you specify models to identify inefficient districts matters, a lot!

Here's one example of how this type of analysis, especially the Predicted Efficiency ROI, can produce deceiving results, simply based on the "shape" of the line fit to the scatter of districts. Below is a scatterplot of the cost adjusted spending per pupil for Illinois school districts (unified K-12 districts) in 2008, and the proficiency rates (natural log) for those districts. In this case, I'm actually using much more fully cost adjusted spending levels (same as above), accounting for regional and more local wage variation, accounting for desired outcome levels, for poverty, language proficiency, racial composition and economies of scale. As a result, the graph actually shows a reasonable relationship between cost adjusted operating expenditures per pupil and actual outcomes. Spending – when appropriately adjusted – is related to outcomes.

Figure 2



Even then, it's difficult to figure out what shape "best fit" line/curve should go through this scatter. If I throw a straight line in there, and compare each district against the straight line, those districts below the line at the left hand side of the picture are identified as really inefficient – getting much lower outcome than the trendline predicts. But, if I were to fit a curve instead (I've simply drawn this one, for illustrative purposes), I might find that some districts previously identified as below the line are now above the line. Are they inefficient, or efficient? Who really knows, in this type of anlaysis?

The most significant problem with the CAP production function analysis is that they came to a result that is so strongly biased on the basis of poverty and instead of questioning whether the model was simply biased – missing important factors related to poverty – they accepted as truth – as a major finding that higher poverty districts are less efficient. It is indeed possible that this is true, but the CAP analysis does not provide any compelling evidence to this effect.

Research literature on Relative Efficiency Analysis

Note that there exists a relatively large literature on measuring school and district performance and efficiency, and on whether or not we can, with any degree of precision, classify the relative efficiency of schools or districts. A handful of rigorous studies have shown relatively low accuracy at identifying which districts or schools are truly efficient, or inefficiency when using simulated data where true efficiency can be known and where common characteristics of real schooling data can be introduced, such as measurement error.[11] There are believers and there are skeptics, but even among the believers and the skeptics, all are applying much more rigorous methods and more refined models and more fully accounting for various cost factors than the present CAP analysis.

What, if anything, can we learn from those pretty maps and scatters?

Moving beyond this technical quibbling, is there anything we actually can learn from the interactive maps and scatterplots that CAP presents in its online tool? First, and most important, any exploration of the data has to be undertaken with the understanding that all 3 ROI's suffer from a severe bias toward labeling high poverty urban districts as inefficient and affluent suburban districts as highly efficient. But, with that in mind, one can find some interesting contrasts.

It would be useful for CAP to reframe and re-label their color schemes. Here's my perspective on their scatters and color coding. The assumption with the ROI is that there exists an expected relationship between adjusted spending and student outcomes. That's the diagonal line I represent in Figure 3 below. Districts in the lower left and upper right are essentially where they are supposed to be. There is nothing particularly inefficient about being in the lower left or upper right. The use of orange to represent the lower left makes it seem like the lower left is like the lower right, or bad, inefficient (worst!).

But, the lower left-hand districts in the scatterplot, in theory, are those that have too little funding and have low outcomes. Arguably, the lower left hand quadrant of the scatterplots is where one should go looking for school districts wishing to challenge their state over inequitable and inadequate funding. These districts aren't to blame. They are getting what's expected of them, given their meager resources. They are getting

[11] 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.

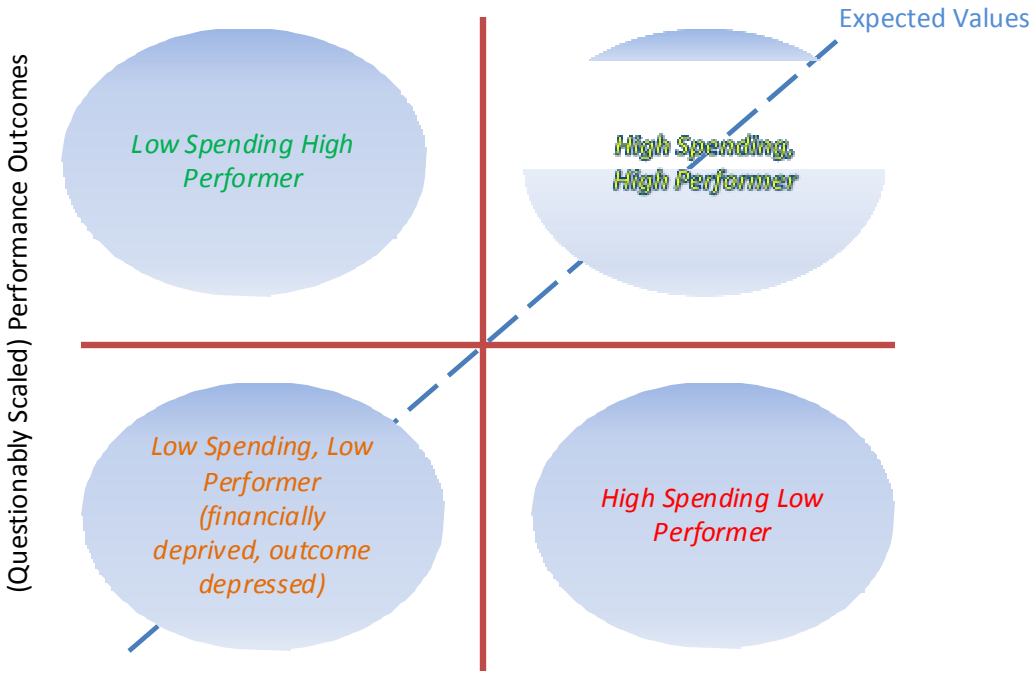
Robert Bifulco and Stewart Bretschneider (2001) Estimating School Efficiency: A comparison of methods using simulate data. *Economics of Education Review* 20

Ruggiero, J. (2007) A comparison of DEA and Stochastic Frontier Model using panel data. *International Transactions in Operational Research* 14 (2007) 259-266

shortchanged on funding and their kids are suffering the consequences – that is, if there really is any precision (which is a suspect assumption) to these models.

Actually, the presence in this quadrant of large numbers of high need districts, in many cases serving large student populations, provides a strong argument for major increases to funding in many states - the opposite of the major recommendation made by the supposedly progressive Center for American Progress.

Figure 3

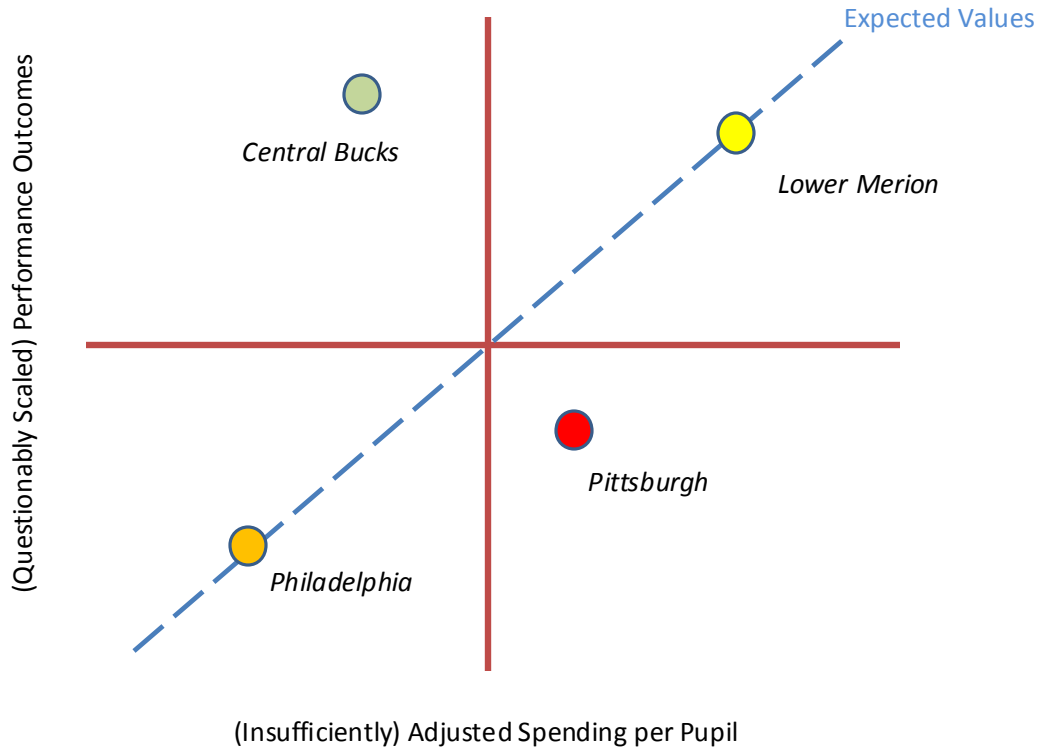


Examples from Pennsylvania

Historically, Pennsylvania has operated one of the least equitable, most regressive state school finance formulas in the nation (www.schoolfundingfairness.org). Philadelphia has been one of the least well-funded large poor urban core districts in the nation. Strangely, Pittsburgh has made out much better financially. Here's what happens when we identify the locations of a few Pennsylvania school districts in the CAP ROI interactive tool. I've recreated the locations of 4 districts. The location of Philadelphia actually makes some sense on the basic ROI - Low funding and low outcomes.

The implication of the orange shading (gradient between yellow and red) seems problematic. But if we ponder the meaning of the lower left quadrant it makes sense. Now, I'm not sure Pittsburgh is really overfunded and/or inefficient, as implied by being in the lower right quadrant – but at least relative to Philadelphia, it does make sense that Pittsburgh falls to the right of Philadelphia on the scatterplot. Lower Merion, an affluent high spending suburb of Philly, seems to be in the right place too. I'm not sure, however, what to make of any of the districts, including affluent suburban Central Bucks, which fall in the upper left. These are largely advantaged, suburban school districts that appear to spend somewhat less than the richest suburban districts.

Figure 4



Other districts falling in the lower-left quadrant include Reading, Allentown, Lancaster, Lebanon, York and Steelton-Highspire, all poorly funded high need school districts in the eastern half of the state. Other districts that fall in the lower right quadrant, implicated as having more funding but still lagging in outcomes, may equally belong in the lower left if more thorough cost adjustment is applied. But one can be quite confident that those presently in the lower left hand corner are significantly under-resourced with respect to their needs. If anything, the CAP ROI analysis provides a case for significantly scaling up funding for several school districts in Pennsylvania and across the nation

currently left behind under state school finance policies. This is a very different message from that conveyed by the Center for American Progress.

A Short List of Examples from New York and Connecticut

Similar to Pennsylvania, Connecticut and New York also have many high need districts of significant enrollment size that appear well into the lower left hand corner of their state's ROI Basic scatterplot. In Connecticut, those districts include the major city of Bridgeport, and smaller cities of Waterbury and New Britain.

In New York, another state with a systematically regressive school funding formula, the lower left hand quadrant includes not only New York City, but also most of the other major and minor cities throughout the state, such as Syracuse, Buffalo, Rochester, Schenectady, Yonkers, Poughkeepsie, Mt. Vernon, Newburgh, Utica, Albany and Middletown. Many of these are very high poverty, small cities with burgeoning non-English speaking populations (like New Britain, CT) and weakening tax bases.

Conclusions

The type of analysis attempted by CAP is an impossible task, especially across all states and dealing with vastly different student outcome data as well as widely varied cost structures. Only precise state by state analysis can yield more useful information of this type. A really important lesson one has to learn when working with data of this type is to realize when the original idea just doesn't work. I've been there a lot myself, even trying this very activity on more than one occasion. There comes a point where you have to drop it and move on. Sometimes you just can't make it do what you want it to. And sometimes what you want it to do is wrong to begin with. Releasing bad information can be very damaging, especially information of this type in the current political context.

But even more disconcerting, releasing bad data, acknowledging many of the relevant caveats, but then drawing bold and unsubstantiated conclusions that fuel the fire… that endorse slashing funds to high need districts and the children they serve – on a deeply flawed and biased empirical basis – is downright irresponsible.