A Cost-Benefit Analysis for Per-Student Expenditures and Academic Achievement

Dr. Sid T. Womack, professor of secondary education, Arkansas Tech University, Russellville, Arkansas

Dr. Kerry Roberts, associate professor of educational leadership, Stephen F. Austin University, Nacodoches, Texas

Dr. C. David Bell, professor and head of curriculum and instruction, Arkansas Tech University, Russellville, Arkansas

Mrs. Karen Womack, visiting lecturer, department of curriculum and instruction, Arkansas Tech University, Russellville, Arkansas

Published February 3, 2015

Abstract: Cost-benefit correlations have been subject to "selective sampling" in the media. Usually extremes of data from a very few high-funding and low-funding states are cited in the media to construct the case that there is no relationship between economic inputs and academic outputs. This study, using average per-pupil expenditures and ACT data from all 50 states, showed a 0.54 correlation. When data were systematically reduced from 50 states to 35 to eliminate the highest- and lowest- spending states, the cost-benefit relationship improved to 0.69, accounting for 47% of the variance in ACT scores. For the 2009-2010 school year, the 35 states in the most predictive range spent from $8712 (Arkansas) to $14,531 per student. Per-pupil expenditures outside that range were not significantly predictive of academic achievement as assessed by the ACT scores of graduating senior in spring 2010.

Particularly in economically austere times, the question sometimes arises "Are we getting our money's worth for our education dollar?" Comparisons of one state's economic inputs to its students' achievement outputs (achievement test scores) are often made with other states. This study employed national data to attempt to shed light on that question.

Such inquiries are not new. The Murphy Commission report in 1998 raised questions in Arkansas that there were few relationships of money spent to the level of academic achievement coming out of that state's high schools. Fox News in 2011 used carefully selected Census data to make the case that the tax burden for education was being transferred to local districts from the states, along with at least the implication that per-pupil costs were steadily increasing while achievement was stagnant. Conservative Action Alerts on April 15, 2012 was more direct in their assertion that "hundreds of billions spent on K-12 education gets only 'stagnant test scores.'" In a lengthy history lesson about the growing education bureaucracy, that report cited Neal McClusky of DownsizingGovernment.org about the increasing size of the education establishment while students allegedly were learning no more than they were two decades ago.

Arguments about non-correlation in the press seem to include a very few "cherry-picked" data with a few achievement test scores that seem to affirm the speaker's viewpoint. We selected national data from the Census and the Public Agenda website to investigate further.

Related Literature

This writing is not the first to show a relationship between costs per student and student achievement. Rasell and Mishel pled the case that the U. S. was not spending enough on education to be competitive with other countries (1990). Berliner (1993) reported that "academically more proficient teachers, who are more experienced, who are better educated, and

who work with smaller classes, are associated with students who demonstrate significantly higher achievement" (p. 636-637).

Berliner further found (1993) that

> An unusual set of data from Texas looks at the effects of teacher ability, teacher experience, class size, and professional certification on student performance in reading and mathematics. Data on millions of students in 900 districts were examined longitudinally from 1986 to 1990 Two rather simple findings emerged. First, teachers academic proficiency explains 20% to 25% of the variation across districts in students' average scores on academic achievement tests. The smarter the teachers, the smarter their pupils appeared to be, as demonstrated by results on standardized achievement tests administer to both groups. Second, teachers with more years of experience have students with higher test scores, lower dropout rates, and higher rates of taking the SAT. Experience counts for about 10% of the variation in student test scores across districts. The effects are such that an increase of 10% in the number of teachers within a district who have nine or more years experience is predicated to reduce dropout rates by about 4% and to increase the percentage of students taking the SAT by 3%. Dollars appear to be more likely to purchase bright and experienced professionals, who, in turn, are more likely to provide us with higher-achieving and better-motivated students (Berliner, 1993, p. 638)

Dollars affect class size. Class size is another research topic almost as hoary with age as the cost-benefit issue. Of class size, Berliner continues with

> The Texas data also show that, in grades 1 through 7, once class size exceeds 18 students, each student over that number is associated with a drop in district academic achievement, This drop is estimated to be very large–perhaps 35 percentile ranks on standardized tests–between a class size of, say, 25 and a class size of 18. Furthermore, the percentage of teachers with master's degrees accounted for 5% of the variation in student scores across districts in grades 1 through 7. So we learn from the Texas study and other data that support its conclusions that academically more proficient teachers, who are more experienced, who are better educated, and who work with smaller classes, are associated with students who demonstrate significantly higher achievement . . . . For those who point out that education costs have been rising faster than inflation, it is important to note that special education populations have been rising as well. It costs 2.3 times as much money to educate a child in special education as it does to educate a child in the regular education program. Most of the real increases in educational expenditures over the last 20 years have been the result of increased costs for transportation, health care, and special education. They have not been connected with regular instruction or teachers' salaries. (Berliner, p. 636-637).

Part of what dollars for education do is to enable a reasonable class size for each student. The student in a class of 40 is clearly not going to get the amount of attention and help from a teacher that a student in a class of 25 will.

The Womack study of 2002 explored costs and benefits especially in Arkansas in an attempt to answer the concerns of the Murphy commission and others. The statewide per-pupil expenditures and ACT data show a correlation of 0.33, $p<0.023$, with 11 percent of the variance

in Arkansas student achievement being accounted for by the level of funding. A review of national expenditure and achievement data using SAT scores showed a correlation of 0.44 in the Womack study of 2002. The present study is mostly a replication or elaboration of that study.

Method

ACT scores for all 50 states were obtained via www.publicagenda.org (2010). Financial data on per-pupil revenues and expenditures were obtained on the Internet from the 2009 Annual Survey of Local Government Finances-School Systems. In one sense, this may have been a simplistic approach to setting up the correlation problem because the graduates of any particular year were influenced by the school finance practices of not just the preceding year, but rather the preceding 13 years. However, these data were chosen partly for simplicity's sake and partly in the realization that within states and districts, it is difficult to modify taxing and allocation practices very quickly. It was assumed that the financial influences of the nation's 2010 graduating classes had been reasonably consistent for the duration of their schooling. Table 1 contains the data used for this study:

Table 1

Mean per-pupil achievement, revenue, and expenditure data for the 50 U. S. states

State	Avg_ACT	Rev_per_student	PP_expenditure
Alabama	20.3	9636	8870
Alaska	21.1	16576	15552
Arizona	20	9882	7813
Arkansas	20.3	9976	8712
California	22.2	11588	9657
Colorado	20.6	10171	8718
Connecticut	23.7	17373	14531
Delaware	23	14335	12257
Florida	19.5	10098	8760

Georgia	20.7	10893	9650
Hawaii	21.8	14987	12399
Idaho	21.8	8141	7092
Illinois	20.7	12457	10835
Indiana	22.3	12360	9369
Iowa	22.2	11337	9707
Kansas	22	11939	9951
Kentucky	19.4	10010	8756
Louisiana	20.1	11967	10533
Maine	23.2	13666	12304
Maryland	22.3	15574	13449
Massachusetts	24	16270	14118
Michigan	19.7	11967	10483
Minnesota	22.9	12664	11098
Mississippi	18.8	8919	8075
Missouri	21.8	10456	9529
Montana	22	11266	10059
Nebraska	22.1	11796	10045
Nevada	21.5	10305	8422
New Hampshire	23.7	13725	11923
New Jersey	23.2	18874	16408
New Mexico	20.1	11266	9439
New York	23.3	20645	18126
North Carolina	21.9	10613	8587
North Dakota	21.5	11664	10151
Ohio	21.8	12811	10560
Oklahoma	20.7	9353	7885
Oregon	21.5	10862	9805
Pennsylvania	21.9	15023	12512
Rhode Island	22.8	15312	13707
South Carolina	20	10719	9277
South Dakota	21.8	9913	8507
Tennessee	19.6	8324	7897
Texas	20.8	10314	8250
Utah	21.8	7954	6356
Vermont	23.2	17108	15175
Virginia	22.3	12146	10930
Washington	23	11510	9550
West Virginia	20.7	10984	10367

Wisconsin	22.1	12435	11078
Wyoming	20	19238	14573
Means	21.55	12348	10636

Descriptive statistics were calculated from these data.

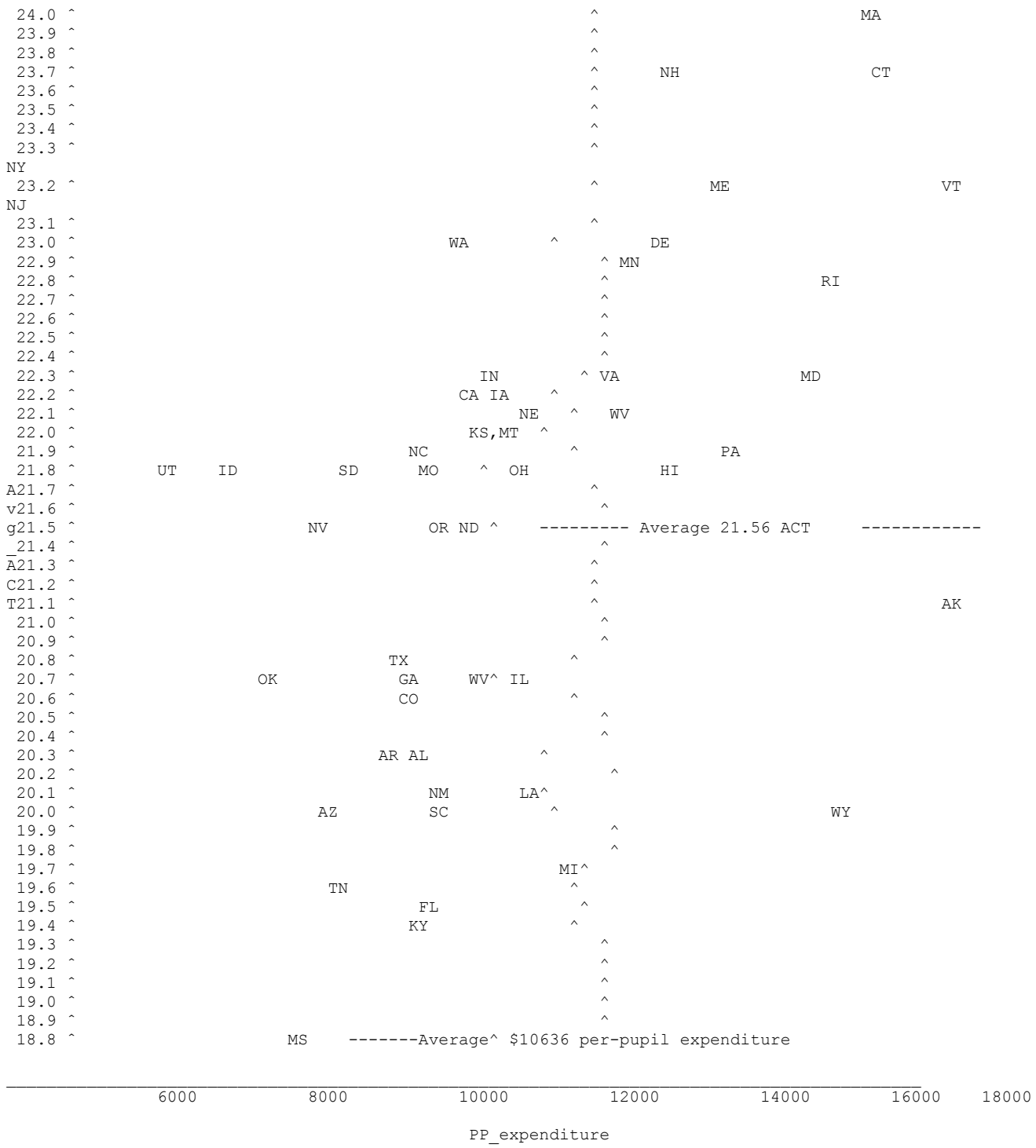
Findings

The relationship between per-pupil revenue and ACT composite scores was 0.55 ($p<0.0001$ for illustrative purposes although these are population parameters and need no probabilities). There was a significant correlation between revenues available per student and ACT scores as one outcome measure of achievement.

Would there be a relationship between money actually spent and ACT scores? The correlation was almost the same as with the revenue figures, 0.54 ($p<0.0001$). To explore these relationships visually, we plotted per-pupil expenditures against the state's average ACT scores.

Figure 1

Scatter plot of 2010 ACT scores and 2008-2009 per-pupil expenditures



The national average of elementary-secondary revenue for 2008-09 according to the Census Bureau was $12, 250; by our calculations using the same data, $12,348.04. The average for per-pupil expenditures in the same year was $10, 636.15. The standard deviations we calculated were $2945.55 and $2548.24, respectively.

We sought a way to determine which states got what they paid for. States that spent per-pupil amounts within plus or minus one standard deviation of the national average should, we reasoned, be able to expect ACT scores within plus or minus one standard deviation (1.29 ACT points) of the national average (21.56 by our calculations). By our definition, such states got what they paid for. We constructed Table 2 to see who got what they paid for.

Table 2

Depiction of which states got what they paid for, got less, got more, paid more/got more, or paid less than the national average and got less than average achievement

State	Got what paid for	Got less than paid for	Got more than paid	Paid more, got more	Paid less, got less
Alabama	X				
Alaska		X			
Arizona					X
Arkansas	X				
California	X				
Colorado	X				
Connecticut				X	
Delaware			X		
Florida		X			
Georgia	X				
Hawaii	X				
Idaho			X		
Illinois	X				
Indiana	X				

Iowa	X				
Kansas	X				
Kentucky		X			
Louisiana	X				
Maine			X		
Maryland		X			
Massachusetts				X	
Michigan		X			
Minnesota			X		
Mississippi					X
Missouri	X				
Montana	X				
Nebraska	X				
Nevada	X				
New Hampshire			X		
New Jersey				X	
New Mexico	X				
New York				X	
North Carolina	X				
North Dakota	X				
Ohio	X				
Oklahoma	X				
Oregon	X				
Pennsylvania	X				
Rhode Island		X (minimally)			
South Carolina	X				
South Dakota	X				
Tennessee					X
Texas	X				
Utah			X (considerably)		
Vermont				X	
Virginia	X				
Washington			X		
West Virginia	X				
Wisconsin	X				
Wyoming		X			
	28	7	7	5	3

Twenty-eight states **got what they paid for**: They paid per-pupil amounts within one standard deviation of the national average, and their students achieved within plus or minus one standard deviation of the national average, using ACT composite scores as the criterion. Some differences in per-pupil revenues and per-pupil expenditures are to be expected because dollars have more purchasing power in some parts of the country than others. Teachers need more money for housing or rent in metropolitan areas than rural ones. These and other things affect the salaries that are to be offered to educators.

Seven states **got less than they paid for**: They devoted revenue within +- one standard deviation of the national per-pupil amounts and spent within one standard deviation of the national average, but the achievement of their students was more than a standard deviation below the national average. Two of those, Alaska and Wyoming, might be expected to have higher per-pupil costs than the national average due to their geographic and meteorological characteristics. The reasons why the other states did not get everything they paid for will likely vary.

Seven other states **got more than they paid for**: They dedicated and spent revenues within a standard deviation of the national average, but their students scored more than the national average. Utah was a very striking case in point, wringing more student achievement out of every education dollar than any other state in the country. Whether Utah can continue to obtain these results on such meager amounts of money remains to be seen.

While it may be tempting for states or districts to use these findings to economize on education, consider the fate of the last category, spent less/got less. When taxpayers get too greedy with dollars for long periods of time, *spent less/got less* is where they end up.

Paid more, got more: Five states spent more than a standard deviation above the national average and had elevated achievement scores to show for it. To get an extra point or two above the average on the ACT, though, it cost them in revenues and per-pupil spending. It takes a sustained effort in funding to make a real difference in this area. The money dedicated to this effort needs to be carefully and thoughtfully spent. Money needs to be directed to goods, services, and facilities that should truly make a difference in student achievement.

Spent less, got less: Three states spent less than the national average and their students paid the price in academic achievement. Upon looking at which three those were, those appear to be states who have attempted this for long periods of time. The overall trend of all of these data is that in the long run, our children get what we pay for.

In both the "spent less, got less" schools and the "got more than they paid for" scenarios, there appears to be an underground economy of teachers spending money out of their own pockets on school supplies that is temporarily buoying education (Womack, 1990; 1992). Teachers can only be expected to give back to the school from their salaries and their free time for so long. At some point, they will vote with their feet. They will move to another state or leave teaching altogether. Ultimately the loss of talent to the education workplace via this scenario is much more expensive than dollars could ever approximate.

This depiction of the data has helped to explain why correlative studies of the past—including the 2002 Womack one—have had difficulty in showing statistically significant relationships between resources and academic outputs. Of particular interest to us was the number of states that obtained student achievement higher than that expected from monetary inputs. These instances make the case that money spent for education is money well spent, but

may confound the correlative cost-benefit graphs which could look so persuasive. The outliers—the extreme disparities between monetary inputs and cognitive outputs-- make the "test cases" that are paraded in the media to dispel the notion that citizens have a material responsibility toward our young.

Additional analyses toward Best Fit:

Additional correlations and regressions were done to find the range of per-pupil funding that correlated highest with student achievement. The first iteration involved removing the five highest-funding states from the data and re-calculating. In removing the top-five per-pupil expenditure states and calculating with the remaining 45, the correlation rose from the previous 0.55 to 0.62. Removing the lowest five states in per-pupil expenditures and calculating with the data from the remaining 40 states yielded a correlation of 0.66. Further analysis using 35 states, with the lowest 10 PPE states and highest 5 PPE states removed, correlated at 0.69, accounting for 47% of the variance in ACT achievement. This was found to be the best model and to illustrate the most effective funding range for maximum student achievement. When the data set was further truncated to 30 states by eliminating the next five highest Per Pupil Expenditure states, r dropped to 0.61 from .69 and F dropped from 29.20 ($P<.0001$) to 16.17 ($P < .0003$). accounting for 37 percent of the variance.

The most effective range of Per Pupil funding, relative to ACT achievement, was between $8712 in Arkansas to $14531 for Connecticut in 2010 dollars. Per-pupil funding lower or higher than that range did not correlate well with student achievement as measured by the ACT in the graduating class of 2010. Using 35-state data, in the range between $8712 and $14531, each additional $166.26 on the average was associated with an increase in ACT score of

0.13 ACT points. If Arkansas, for instance, the lowest PPE state in the 35 state analysis, wished to raise its average ACT score from 20.3 to 22, it should increase its Per Pupil Expenditure from $8712 by $2174.17 to $10.886.17.

When further regression analysis was done with the lowest-10 PPE states, an insignificant r of 0.13 was found between Per Pupil Expenditures (F=0.1396). In 2010 dollars, there was no no relationship between economic inputs and academic outputs for under-$8712 PPE states. At the other end of the expenditure range, regression analysis of PPE data and ACT scores among the top five PPE states showed no ($p<.0.2239$, $F=2.3352$) relationship. Based upon these findings, it was believed that the extremes of the funding spectrum were what had weakened the correlation of data to 0.55 when data from all 50 states were included.

The correlation between Per Pupil Expenditures and ACT scores in 35 states may be explained in part due to the fact that costs of living vary somewhat in different parts of the country. A dollar in Arkansas likely buys more real estate than a dollar in Connecticut. A difficult-to-quantify additional variable may have to do with how education dollars are spent. In some communities, extra-curricular activities command dollars that would not have been spent on athletics, art, or music in other communities. The ACT test measures cognitive outcomes, not how well students catch footballs, paint in watercolors, or play scales on instruments. The more nearly an expenditure is aligned with core cognitive outcomes, the more predictive that expenditure correlates with improved ACT scores.

In order to make application of these findings, a little more analysis of the data could be useful. Regression analysis suggested that among all 50 states, it should have taken about $1096 in 2008-09 per-pupil expenditures to raise a state's ACT score by 1 point. A second calculation,

this time using an averaging process, indicated that what happened in 2008-09 was more like $1383.66 per pupil if the states with 23.0 ACT averages or above were excluded. If data from all 50 states were left in the analysis, an ACT point above the 18.64 ACT "floor" of the regression model cost $2263.46 each. Among the seven states with ACT averages above 23.0 in 2010, their average per-pupil cost in 2008- 2009 was $14,389.14. Per-pupil costs among the 43 below-23 ACT states were $9249.27. It took about $5,000 per child more per year to leave the "got what we paid for" club and join the "paid more, got more" one. As state-level policy makers consider the implications of this study, the law of diminishing returns should be considered. Do we want to have average schools, or excellent ones? The level of funding accounted for 30 percent of the variance in the regression model.

Conclusions, Discussions, and Recommendations.

The correlation between Per Pupil Expenditures and ACT scores in 35 states may be explained in part due to the fact that costs of living vary somewhat in different parts of the country. A dollar in Arkansas likely buys more real estate than a dollar in Connecticut. A difficult-to-quantify additional variable may have to do with how education dollars are spent. In some communities, extra-curricular activities command dollars that would not have been spent on athletics, art, or music in other communities. The ACT test measures cognitive outcomes, not how well students catch footballs, paint in watercolors, or play scales on instruments. The more nearly an expenditure is aligned with core cognitive outcomes, the more predictive that expenditure correlates with improved ACT scores.

One thing seemed mostly apparent in these data: If a state goes really low on funding for a very long period of time, what that state will purchase will be a below-average education for

their students. At a time when the U. S. as a whole is suffering in the international comparisons, finishing near the bottom in the U. S. may not be a good idea for the next generation.

Fixing it. To apply these figures, let us take the case of Mississippi. In 2010, their average ACT score was 18.8, which is 2.75 ACT units below the national average of 21.56. If Mississippi could have had $1373.66 to multiply times the 2.75 ACT units needed to bring them to the national average, their 2008-09 expenditures per student would have been $10, 852.65. The national average that year was $10,636. Average costs to achieve average results? Imagine that!

It can be difficult for states to change categories from "paid less, got less" to "got what they paid for." Until the mid-1990s, Arkansas had had a long tradition of "paid less, got less." The *Lakeview* cases in the mid-1990s focused attention on the fact that the Arkansas funding mechanism was not operating within the Arkansas constitution which called for a "general, suitable, and efficient system of public schools." The Lakeview school district was able to demonstrate to the court that the system was not a suitable one. The legislature, buoyed by a change in public attitudes, raised the minimum mileages required of local school districts and appropriated other money, bringing Arkansas into the "got what they paid for" category. Arkansas' 2010 ACT average of 20.3 with a 2008-09 per-pupil revenue of $9976 and per-pupil expenditure of $8712 is on the prediction line, albeit a bit toward the lower left-hand side.

States that have recently redoubled their financial efforts may have to wait several years for the effects of those actions to take hold. States that have expropriated money originally intended for education may get away with it for several years until "the chickens come home to

roost." But in the longer term, everyone gets what they paid for.

Prudence in purchasing. For decades, there have been "cherry-picked" comparisons of per-pupil funding and achievement outcomes. Education has not usually fared well in such comparisons because the data were sometimes selected in intentionally biased ways to discourage funding for education by a press with an agenda. The national data in the present study were taken as a whole—with the occasional "flyers" (off the diagonal prediction line) included. They still showed a cost-benefit relationship. But how that that cost-benefit relationship be strengthened?

If academic achievement is what we are wanting to buy with our education dollars, then the expenditures need to be for academics. Perhaps donations or local fundraisers can help with extracurricular programs. Few events discourage millage increases more than seeing the coaches and band directors get raises while the classroom teachers get nothing. The ACT scores are being made in classrooms, not gridirons and band halls. (One of your writers is a former band director.) Educators need salaries that are high enough to enable them to live middle-class lifestyles. Few go into teaching expecting to be rich, but they don't go into education expecting to struggle financially from year to year, either.

Educators need the basic necessities of today's classrooms. Science labs need to be adequately equipped. Classrooms need the electronic equipment needed to bring the world to their students so the world can be studied.

Part of what is purchased with enhanced funding is reasonable class sizes. It seems likely that the only people who doubt the effects of class size upon achievement are those who have never taught as a teacher under contract.

References

Berliner, D. C. (1993). Mythology and the American system of education. *Phi Delta Kappan*, *74* (8), 632-640.

McLuskey, N. Hundreds of billions spent on K-12 education gets only 'stagnant test scores.' Retrieved June 14, 2012, at http://www.conservativeactionalerts.com/2012/04/report-hundreds-of-billions-spent-on-k-12-education-gets-only-stagnant-test-scores/

National Education Association, Research Division. (1998). *Rankings of the States 1997*: *A Data Resource for Examining State Public Education Programs*. Washington, D. C.: NEA

Organisation for Economic Cooperation and Development. (1992). *Education at a Glance*. Paris: Centre for Educational Research and Innovation.

Protheroe, N. (1998). The blob revisited. *The School Administrator*, 55 (8), 26-29.

Rasell, M .E., & Mishel, L. (1990). *Shortchanging Education: How U. S. Spending on Grades K-12 Lags Behind Other Industrialized Nations*. Washington, D. C.: Economic Policy Institute.

The Murphy Commission: An Arkansas Policy Foundation Initiative. (1998). A position paper prepared by the Murphy Commission Education Workgroup.

Womack, S. (1990). No extra charge. *American School Board Journal*, *177* (7), 34, 37.

Womack, S. (1991). Who says teachers only work from 8: 30 to 3: 30? *North Carolina Education, 21* (3), 16-17.

Womack, S. (1992). Demographics and attitudes of Arkansas educators. *The Clearing House*,

65 (3), 152-155.

Womack, S. (2002). The relationship between resources and academic achievement. Eugene, Oregon: *ERIC Clearinghouse on Educational Management.* (ERIC Document Reproduction Service, ED 456498).

http://www.publicagenda.org/charts/state-state-sat-and-act-scores

http://www.census.gov/newsroom/releases/pdf/cb11-94_table_11.pdf