

7

Science Violated: Spending Projections and the “Costing Out” of an Adequate Education

Eric A. Hanushek

THE RECENT MOVEMENT to hold schools accountable for student performance has highlighted a simple fact: Many students are not achieving at desired levels.¹ Moreover, it takes little additional evidence to realize that many schools are not on an obvious path toward eliminating the gaps. These simple facts have led people with widely varying reform perspectives to enter into the fray with plans and solutions. And a natural follow-on question is invariably “what will it cost?” To answer this important question, a series of very misleading methods for estimating the costs of an improved education have evolved, but the problems

1. This analysis benefited from helpful comments by Alfred Lindseth, Paul Peterson, Martin West, and Michael Wolkoff and from the research assistance of Brent Faville. An earlier version of this paper was presented at the conference on *Adequacy Lawsuits: Their Growing Impact on American Education*, Kennedy School of Government, Harvard University, October 13–14, 2005, and a companion analysis of costing out studies is found in Hanushek (forthcoming).

with these methods are generally unrecognized (or ignored) in the public and judicial debate.

“Costing out studies” should be interpreted as political documents, not as scientific studies. They are most frequently contracted for by parties interested in increasing spending for education (including teachers unions, state departments of education, and litigants), although they sometimes involve defensive reactions of parties trying to neutralize a rival costing out study that calls for large increases in spending. They are seldom used as analytical tools to aid in policy deliberations.

The greatest premium is placed on finding “a number,” because—regardless of its validity—a number for total “required” state spending can be used easily in a public relations campaign.² Discussion of the underlying basis for the number is usually relegated to the status of arcane academic debate, while the number itself frames the discussion. The debate about the basis for the number is not news, but the number is.

These studies *inherently* fail to provide usable information about the resources that would be required to meet a given student achievement level, at least when the resources are used efficiently and effectively. Instead, as described below, the studies merely provide spending projections that incorporate, and in general lock in, current inefficient uses of school funds.

But the other side is equally as important. Even if the specific method used in the spending projections is based on programs that have a proven track record of effectiveness—an infrequent occurrence in itself—there is no mechanism that will ensure the funds provided will be used in a way that is consistent with the effective programs. In fact, the final reports on spending projections invariably include a disclaimer that indicates one should

2. This explains why the Web sites for advocacy organizations give top billing to costing out studies. For example, see the ACCESS Project at <http://www.schoolfunding.info>.

not really expect the outcomes they consider because a variety of other forces are likely to dissipate any results. In other words, none of these studies suggests that the projected spending would actually have an effect on student achievement. To deflect criticism these studies frequently couch the analysis in terms of “opportunity” instead of outcome, but there is no scientific or objective way to define such an approach.

The warning of lack of results is perhaps the most accurate statement in a number of these studies. Little evidence supports the case that improvements have followed past court infusions of funds. This chapter concludes with additional data on such ineffectiveness, and Evers and Clopton (chapter 4) provide detailed case studies of the failure of large increases in funds to lead to noticeable improvements in student outcomes.

The fundamental issues surrounding the design and execution of these studies, described in this chapter, make these studies an inappropriate basis for judicial or legislative deliberations on school finance.

Approaches to Costing Out Adequacy

The pressures to solve the widely perceived problems with public schools have led courts and legislatures to look for a scientific determination of the amount of spending by schools that would be adequate to achieve the state standards. Indeed there has been no shortage of consultants who are prepared to provide an analytical answer to what level of spending is needed. This activity, dubbed “costing out studies,” has been conducted in more than thirty-three states, and the demand for such analyses has only increased.³ Courts are willing to write the specific numbers

3. A review of past costing out studies can be found in *Education Week* (2005). See also the ACCESS Project Web site, a project of the Campaign for Fiscal Equity (CFE), the plaintiffs in the New York City adequacy case, *Campaign*

from costing out studies into judgments,⁴ and legislatures come back repeatedly to these studies to guide their appropriations. Plaintiffs entering into lawsuits about school funding, recognizing the political power that can be generated by them, now tend to make a requirement for an official costing out study as the first remedy they seek.

Much of the allure of the existing study approaches derives from their commonsensical and logical approaches to analysis, all wrapped in a patina of science. These perceived traits benefit, however, from misconceptions about the underlying analyses. They do not meet the most basic standards of scientific inquiry.

A set of now-standard approaches has been developed to answer the question “how much would it cost to make all students achieve proficiency?” These approaches differ in important ways, but they share one common feature—none can provide a valid and reliable answer to this question. As a leading proponent of the use of these costing out studies concedes, “the aura of ‘scientific’ decision-making that is associated with these studies can be misleading. It is not, in fact, possible to definitively identify the precise amount of money that is needed for an adequate education. Although these studies use a variety of complex statistical and analytic techniques, all of them are premised on a number of critical judgments which strongly influence their ultimate outcomes” (Rebell 2006, 5).

There is little scholarly research on these analyses. A small number of firms have conducted contract work with organiza-

for *Fiscal Equity v. State of New York*, 100 N.Y.2d 893 (N.Y. 2003). CFE states that its primary mission is to “promote better education by conducting research, developing effective strategies for litigation and remedies (including cost studies), and providing tools for public engagement.” The count of earlier costing out studies comes from <http://www.schoolfunding.info/index.php3>, accessed on October 7, 2005.

4. See, for example, *Campaign for Fiscal Equity v. State of New York* and *Montoy v. State of Kansas*, No. 92032 (Kan. S.Ct. June 3, 2005).

tions in specific states. These analyses are, however, similar across states and across firms applying a common approach. It is also true that the common nomenclature for each type of study is itself misleading and does not accurately reflect the underlying approach to obtaining a cost estimate.

Perhaps the most commonly applied approach is the “professional judgment” method.⁵ With a few nuances, the underlying method involves convening a panel of educators—teachers, principals, superintendents, and other education personnel such as business officers—and asking them to develop an educational program that would meet certain specified outcome standards. Their efforts typically produce “model schools” defined through class size, guidance and support personnel, and other programs that might be necessary. The analysts running the process then provide elements missing from the model schools (e.g., central administration costs or computers and materials) and use externally derived cost factors (e.g., average teacher or principal salaries) to the model schools. Depending on the details of the panel activities, the panels may provide guidance on the extra resources for disadvantaged children, special education, and English language learners, or these extra resources may come from cost factors assumed by the consultants.

An alternative but similar approach directly substitutes the judgment of the analysts themselves for the judgment of the professional panels. This approach has been immodestly called the “state of the art” approach by the primary firm associated with it.⁶ At other times, building on the current mantra of educational

5. Examples of this (coupled with the leading groups applying the methodology) include Augenblick & Myers (2002), Augenblick, Myers, Silverstein, and Barkis (2002), Augenblick, Palaich, and Associates (2003), AIR/MAP (2004a), Picus, Odden, and Fermanich (2003), and Verstegen and Associates (2003).

6. See Odden, Fermanich, and Picus (2003).

policy, the consultants refer to it as the “evidence-based” method. The consultants sort through available research, select specific studies that relate to elements of a model school, and translate these studies into precise implications for resources in schools. It is advertised as applying research evidence to develop a set of model schools that are subsequently costed out in the same manner as the professional judgment model schools.

Neither of the previous methods makes any use of the actual spending and achievement experiences of districts in the specific state. The remaining two approaches rely on data from the schools and districts in a state.

The “successful schools” model begins by identifying a subset of the schools in a state that are effective at meeting educational goals. (Note that this is also conducted at the district rather than the school level.) The identification of successful schools may use differing methods but usually concentrates on the level of student achievement, possibly including identified input levels that relate to state policies and regulations and, infrequently, making adjustments to allow for the background of students.⁷ Spending on special programs—say, remedial education or special education—is stripped out of budgets in the successful schools to obtain a “base cost” figure for each district. Typically, then, the base costs for a portion of these schools—derived from excluding some number of schools in the tails of the distribution that are presumed to be outliers—are averaged to develop a level of spending that can feasibly yield effective performance. To get the full costs of the school, expenditures on special programs are then added back, based on the distribution of students with such special needs for each school.

The “cost function” approach, sometimes referred to as the

7. See, for example, Augenblick and Myers (1997), Myers and Silverstein (2005), and Standard & Poor’s School Evaluation Service (2004).

“econometric” approach, also uses the experiences of the state’s schools in spending and achievement to derive what different levels of achievement would cost according to the available observations on the current practices of schools.⁸ The exact methodology, while invariably involving a series of complicated statistical steps, differs in its application across states but has similarities to the successful schools analysis in attempting to characterize districts that are meeting desired achievement standards. Through statistical methods, the approach estimates how spending is affected by different student outcome levels and different student characteristics—which in turn can be used to derive the spending for different districts attempting to meet different performance levels. This approach may or may not attempt to distinguish between efficient and inefficient producers of outcomes, that is, between districts that spend more for some given level of achievement than others do.⁹

As explained below, each name is but a *nom de guerre*, used to market methods as serious scientific approaches to costing out adequacy. In reality, each method suffers from serious shortcomings, and each fails to provide a reliable or scientific way to estimate the needed expenditures for achieving prescribed levels of outcomes.

Why the Methods Don’t (Can’t) Work

Each of the approaches to determining the costs of an adequate education has some surface appeal, but their validity and reli-

8. Examples of this analysis include Duncombe, Lukemeyer, and Yinger (2003), Reschovsky and Imazeki (2003), and Gronberg, Jansen, Taylor, and Booker (2004).

9. Gronberg, Jansen, Taylor, and Booker (2004) explicitly analyzed the efficiency of districts, but this analysis was not well received in the courtroom; see the decision of Judge John Dietz in *West Orange-Cove Consolidated Independent School District et al. v. Neeley et al.*, No. GV100528 (Dist. Ct. Travis County, Texas, Nov. 30, 2004).

ability depend on their treatment of several important steps. The evidence about costing out studies is drawn from a selection of existing analyses. This selection was not drawn because the examples were particularly better or worse in application than others. Instead, they are used as convenient illustrations of the larger problems.

The theme of the discussion is that the identified problems with each approach are not ones of application that can be fixed by doing better. The problems are fundamental flaws that are not readily dealt with through fine-tuning one or the other of the approaches.

The Co-existence of Alternative Outcome Standards

The outcome standards that are considered should have a significant effect on the analysis of costs. For example, bringing all New York State students up to the level of having an elite diploma (a New York State Regents Diploma) is one of the loftiest goals of any state in the nation.¹⁰ This standard is clearly different from the constitutional requirement which, by the interpretation of the court of appeals, was a sound basic education—a standard explicitly below the Regents Diploma. Different outcome standards frequently coexist. In fact the existence of multiple standards has proliferated, since the states moved to more comprehensive accountability systems, and the federal government intervened in linking accountability to student performance under the No Child Left Behind Act of 2001 (NCLB). All estimation of performance and costs depends directly on the outcome standard that is applied.

10. New York State traditionally had two different diplomas with varying requirements. In 1996 the New York Regents decided that all students would have to qualify for a Regents Diploma (the previously optional high standard undertaken by roughly half of the students in New York State). This requirement has had a long phase-in period with altered testing requirements.

The choice of standards is a political decision, reflecting a variety of factors. Often the state department of education or the state board of education promulgates its standards, but they are not necessarily the views of the elected officials in the executive or legislative branches of the state. More important, these standards rarely bear any relation to constitutional standards, which are often phrased in broad generalities. Nor are they the same as the mandatory standards that might exist under state or federal accountability standards.

Clearly, decisions about the standards that should be applied are not within the purview of the hired researchers doing the costing out studies. But since many costing out studies are commissioned and paid for by parties with a position on what they would like the answer to be and with an understanding of the political import of the results, neither should the definition of outcome be left to the organization that contracts for the study to be done.

None of the extant methods for costing out adequacy avoids this issue. Each must explicitly or implicitly base estimation on a definition of outcomes, but this definition requires political judgments that are seldom introduced.

Take some examples. The New York City adequacy suit, after a full round of legal decisions, was remanded to the lower court to determine a final judgment on actions to deal with the constitutional failure of the extant system. The plaintiff in the case, the Campaign for Fiscal Equity, hired two consulting firms—the American Institutes for Research (AIR) and Management Analysis and Planning (MAP)—to cost out an adequate education in New York City under the New York State constitutional requirement for providing a “sound basic education.”¹¹ This group of

11. Details of the costing out exercises in the *CFE* case can be found in Hanushek (2005).

consultants chose, in consultation with their clients, to evaluate the costs of meeting the Regents Learning Standards that all children in New York should get a Regents Diploma. The Governor's commission, appointed to assess the appropriate State response to the court's decision, adopted a lower standard in its estimation of costs, conducted with Standard & Poor's School Evaluation Service (2004). The judicial referees, who were appointed by the court to advise it on the decision, simply ignored differences in the standards for cost estimation and were pleased by the consistency of the estimates—even though they were based on different outcome standards and should not have been the same by the logic of costing out (Hanushek 2005). The referees then went on in their report to recognize that the highest court had already said that the Regents Learning Standards were inappropriate, apparently oblivious of the fact that standards should affect any cost estimates.¹²

Take the studies commissioned in Kentucky. Three separate studies were conducted by two firms: Verstegen and Associates and Picus and Associates (who conducted parallel studies using a professional judgment and a "state of the art" approach). Picus and Associates (Odden, Fermanich, and Picus 2003) are generally willing to let their professional judgment panels define what the vague seven constitutional requirements of education laid down by the Kentucky Supreme Court mean as long as the requirements are fully met by 2014.¹³ Verstegen and Associates

12. John D. Feerick, E. Leo Milonas, and William C. Thompson, *Report and Recommendations of the Judicial Referees (CFE, Nov. 30, 2004)*.

13. The instructions given to the panelists about student outcomes to be achieved were: sufficient oral and written communication skills to enable students to function in a complex and rapidly changing civilization; sufficient knowledge of economic, social, and political systems to enable the student to make informed choices; sufficient understanding of governmental processes to enable the student to understand the issues that affect his or her community, state, and nation; sufficient self-knowledge and knowledge of his or her mental

(2003), on the other hand, call for these requirements along with an extensive set of input and process requirements included in the current Kentucky school regulations.

Or take Augenblick, Myers, Silverstein, and Barkis (2002) in Kansas. This analysis, which was later written into the judgment of the Kansas State Supreme Court, provides the following insight into the consultant's role in setting student outcome standards:

A&M worked with the LEPC [Legislative Education Planning Committee] to develop a more specific definition of a suitable education. We suggested using a combination of both input and output measures. For the input measures, it was decided that the current QPA [Quality Performance Accreditation] requirements would be used, along with some added language provided by the LEPC. This additional language included vocational education as a required course offering, and identified other programs and services that might be provided as part of a suitable education. Next we set the performance measures that would be used. Again, A&M worked with the LEPC. Together we determined which content areas and grade levels would be used. The math and reading tests are given in the same grade levels every year, the writing, science and social studies tests are given in alternating years. A&M felt that the reading and math tests, which are given every year, gave us the most flexibility in setting the output measures.

Perhaps more interestingly, the definition of adequacy is not always related to outcomes. In North Dakota, Augenblick, Palaich, and Associates (2003), the successor firm to Augenblick &

and physical wellness; sufficient grounding in the arts to enable each student to appreciate his or her cultural and historical heritage; sufficient training or preparation for advanced training in either academic or vocational fields so as to enable each child to choose and pursue life work intelligently; and sufficient levels of academic or vocational skills to enable public school students to compete favorably with their counterparts in surrounding states, in academics or in the job market.

Myers, noted that the state did not have explicit outcome standards but instead had input requirements. For their analysis, however, they layered on a set of outcomes that were related to state goals under No Child Left Behind. (Of course, if one were just interested in providing a well-defined set of inputs and did not have to worry about the relationship with student outcomes, it would be relatively easy to calculate the level of “adequate” funding using existing spending on the inputs.)

Duncombe, Lukemeyer, and Yinger (2004) analyze the effects of different goals on the estimated costs under alternative estimation approaches. They demonstrate that reasonable differences in the loftiness of the educational goal can lead to 25 percent differences in estimated costs in their own estimation approach and 50 percent differences across alternative estimation approaches, including the professional judgment approach.

The organizations commissioning different costing out studies appear to recognize the importance of the standard chosen, often arguing for the highest standard on record (e.g., the Regents Learning Standards in CFE’s instructions) or at least a full NCLB standard of 100 percent proficient. The exception is the successful schools approach, where the method *requires* that some schools meet the standard, that is, are successful. This requirement implies that the outcome standard chosen cannot be too far from current operations, and probably also explains why relatively few studies commissioned by special interest groups use the successful schools method (Baker, Taylor, and Vedlitz 2005).

The application of any standard, particularly in the professional judgment or the state-of-the-art approach, is usually left vague and up to the interpretation of the individual panel members or the consultants. This vagueness is entirely understandable, because it is far from obvious how the precise standard (or variations on it) could enter into the costing out approach. The

two approaches that build on observed outcomes in a state (the successful schools and cost function methods) have a different problem. They must have actual data on how close any school is coming to meeting the standard, and more important, the methods cannot feasibly consider more than one or two explicitly measured outcomes. These constraints often call for the consultants basing studies entirely on data availability and their own outcome choices.

But arbitrary choices of objectives yield arbitrary estimates of costs. The courts on the other hand seldom focus on the standard used by the consultant and instead tend to grasp the cost identified without apparent regard for the importance of the chosen objectives.

The appropriate outcome standard clearly differs by purpose, and a variety of people enter into setting the definition in varying circumstances. But in the judicial adequacy deliberations, it is simply inappropriate to divorce these definitions from the democratic policy process and to deed it over to consultants and interested parties.¹⁴

The Empirical Basis of the Cost-Performance Relationship

Costing out studies address questions of the relation between a desired outcome (“adequate education”) and the set of resources needed to reach that outcome. Put differently, the key to any such study is whether it accurately identifies how much achieve-

14. Surprisingly, not everybody would agree that outcome standards should be politically interpreted. Michael Rebell, a central figure in the New York City adequacy case and others, holds that the consultants *should* be the ones to determine the appropriate standards. In his words, “Education finance analysts should be held responsible for articulating and justifying the output measures used in their studies, and they should not be allowed to ‘pass the buck’ by stating that they are accepting vague or illogical output measures simply because they have emerged from the political process” (Rebell 2006, 53).

ment will change with added resources. Providing a reliable answer to this question has defied all past research, and none of the approaches to costing out an adequate education solve it.

The school systems in each state generate information about the relation between current spending and achievement, but this is seldom easy to interpret. Different school districts have different advantages in terms of the clientele they serve, and different districts make different choices about curriculum, programs, and personnel. These interact with spending decisions, often leading to little obvious relation between resources and achievement.

Decades of scientific research across a wide range of school experiences has focused on uncovering the contribution of schools to student outcomes. This substantial body of work shows, contrary to widely held popular beliefs, that there is not a consistent relation between school resources and student achievement (see Hanushek 2003). Such a finding of course presents a challenge to the consultants who attempt to describe the expanded resources needed to push student performance to the desired levels.

In the courtroom the plaintiffs seeking more resources have developed a variety of approaches to deal with this fundamental problem for their cases. One is simply to ignore the accumulated evidence, relying instead on common beliefs. Another is to set up a straw man by translating the research findings into the trivial question, “does money matter?” Some minimal level of resources is obviously necessary. Moreover, the research neither says that resources *never* matter nor that resources *could not matter*. It does, however, show that providing resources without changing other aspects of schools, such as the incentives for performance by teachers and administrators, is unlikely to boost student performance. The research evidence also fails to identify conditions or uses of money that translate resources into student

performance, making it impossible to specify a combination of resources and programs that will reliably boost achievement.

The challenge of squaring actual observations with costing out studies is best seen in a candid statement in Augenblick & Myers (2002), which is also repeated in most of their other studies:

The effort to develop these approaches stems from the fact that no existing research demonstrates a straightforward relationship between how much is spent to provide education services and performance, whether of student, school, or school district.

In the absence of such a simple relationship, and in light of the fact that some people believe that there is no clear association between spending and performance, four rational approaches have emerged as ways to determine a base cost level: (1) the professional judgment approach; (2) the successful school (district) approach; (3) the comprehensive school reform approach; and (4) the statistical approach.

In other words, the *advantage (!)* of the various methods is that they do not require any basis in the empirical reality of the specific state or, more generally, any state. The professional judgment panels or the state-of-the-art researchers in particular are completely free to declare anything without worry about being contradicted by the data.

The professional judgment panels employ professional educators to develop programs and model schools, but there is never any indication that the members of these panels have any particular relevant expertise in terms of a knowledge of the extant research base, of an understanding of outcomes either inside or outside of their own locality, or of the effects of varying amounts of resources, especially when outside of their own experience. Indeed, no indication is generally given of the selection criteria for panelists. Were they chosen because they came from particularly innovative or high-quality districts? Were they cho-

sen because of previously expressed views on programs or resources? Or were they just the subset of a larger invited group representing those willing to attend a weekend session in exchange for added pay?

The consultants performing the study seldom know any of the education personnel in the state, so they obviously need to solicit nominations—frequently from the organization commissioning the study. But since these organizations generally have a direct interest in the outcomes of the study, it seems unlikely that they will produce a random selection of educators to serve on the professional judgment panels. The nature of the selection process ensures that the judgments of any panel cannot be replicated (a fundamental concern of any truly scientific inquiry).

But reality is worse than that. The educators recognize by the nature of the exercise that their input to the process may have an effect on their future well-being. This bias and conflict of interest is most apparent in the highly publicized court cases, such as that in New York City where the professional judgment panels were employed to suggest a remedy to an already decided liability (Hanushek 2005). Such a conflict is nonetheless also generally present in less publicized circumstances when educators are asked to develop a wish list of what they might like in their schools and districts. As noted in Massachusetts, “A review of the study (ex. 35, the professional judgment study by Dr. Verstegen) suggests that the resource needs identified represent to some extent a wish list of resources that teachers and administrators would like to have if they were creating an ideal school with no need to think about cost at all.”¹⁵

15. Exhibit 35 is the professional judgment study of Dr. Verstegen. The judgment goes on to note: “In this regard, as the defendants’ witness Dr. Robert Costrell pointed out, if Dr. Verstegen’s professional judgment model is applied to the comparison districts of Brookline, Concord/Concord-Carlisle, and Wellesley, it appears that none of the three is spending enough to provide an adequate

The lack of any empirical linkage described in Augenblick & Myers (2002) is precisely true for the professional judgment work and close to true for the state-of-the-art work. The empirical basis of the state-of-the-art analyses is a small number of selected research studies that relate to some schooling experiences, although not the experiences in the individual state. And most important, because these are highly selective studies from the research base, there is no reason to believe that they can be generalized or that they reflect the empirical reality *anywhere*.

The successful schools analysis uses information on a selected subset of the schools, based on the performance of their students. The identification and selection of the successful schools is obviously an important step. From a scientific perspective, simply taking high performing schools defined by the level of student test scores and other outcomes is inappropriate, because performance is affected by a host of nonschool factors including family background, peers, and prior schooling experiences. If these other factors are ignored, the interpretation of the observed spending-achievement relationships in the successful schools or successful districts is unclear, because there is no sense that the relation is causal or could be reproduced by simply altering the spending of a district. Nonetheless, virtually all existing successful schools studies rely on success defined just by the level of student achievement, not by the value added of schools.

The various cost function estimation approaches explicitly rely on the spending and achievement of the schools in a state,

education. Dr. Costrell could identify only five school districts in the Commonwealth that are spending at a level that would be considered appropriate according to the Verstegen model (see ex. 5449), and none of the five is included in Myers' seventy-five 'successful school' districts." This latter point reappears elsewhere, as noted in the analysis below. See trial record in *Hancock, et al v. Commissioner of Education, et al*, 882 N.E.2d 1134 (2005).

thus appearing to be closer to actual schooling experiences. But, the key to interpreting these remains whether or not they have adequately identified the *causal* relationship between student performance and spending.

A simple way to understand these cost function estimates is to begin with the closely related estimation contained in the extensive literature on educational production function. A wide range of past studies—as underscored by the quotation from Augenblick & Myers (2002) above—have looked for a relation between resources and achievement. This work involves estimating the statistical relation between achievement and a series of individual characteristics along with various measures of the resources available. This research has generally found little in the way of a consistent relationship between spending and student outcomes, and moreover almost all estimates that suggest such a resource-achievement relation often show a very small effect of resources on student outcomes (Hanushek 2003). If one were to take the estimates of the effect of resources from these, there would be the immediate implication that large amounts of resources were needed to obtain a noticeable achievement gain (again, because resources have little effect on achievement).

But now consider cost function estimates, which generally involve a statistical relation between spending as the dependent variable and achievement and characteristics of the student population as the explanatory variables.¹⁶ This analysis essentially moves spending from the right-hand side of the equation to the

16. Note that these estimates bear little relationship to classic cost functions in microeconomic theory that would use an underlying assumption of optimal firm behavior to translate the production function (achievement as related to various inputs) into a cost function that describes how cost relates to the prices of inputs. None of the work in education observes any variations in input prices (e.g., teacher wages, textbook costs, and the like). The empirical work in education described here relates spending to outputs and inputs such as the number or kind of teachers, the poverty rate, and so forth.

left, and achievement to the right.¹⁷ If the estimated effect of spending on achievement is small, this approach reverses it to indicate that it takes a lot of spending to obtain a little more achievement. But they have not necessarily identified the cost, or the expenditure needed, to obtain any outcome. They have only shown that the current pattern of spending is not very productive, exactly like the more extensive production function estimation.

This estimation is directly related to the production function estimation. It is given the new clothing of being a “cost function,” but it simply describes the existing spending patterns across districts with different achievement levels.¹⁸ The expenditure function does not indicate the minimum expenditure (or cost) of reaching any achievement level but instead identifies average spending behavior seen in districts.

No scientifically valid method is used to answer the question “how will achievement change for a given change in resources or spending?” even though the question is central to all the costing out approaches. This issue proves to be beyond the current

17. Some approaches to cost estimation are not done in this way but instead use various optimization methods to obtain the minimum cost of achieving some outcomes. They are nonetheless subject to the same interpretative issues about causation.

18. There are some serious statistical complications in this work. The econometric methodology places requirements on the modeling that are almost certainly violated in this estimation. The cost function estimation essentially assumes that districts first specify the outputs they will obtain and that this chosen achievement level and the characteristics of the student body determine the spending that would be required (i.e., achievement is exogenous in statistical parlance). This approach, while summarizing the average spending patterns of different districts, is inconsistent with the interpretation that the level of resources available to a district determines student outcomes.

The specific data and modeling are also very important. As Gronberg, Jansen, Taylor, and Booker (2004) state, “The measurement of efficiency in producing a set of outcomes is directly linked to the particular set of performance measures that are included in the cost model and the particular set of input measures.”

capacity of extant scientific investigations and is not overcome by the limited investigations of the costing out consultants.

The Treatment of Inefficiency

It seems clear, and the evidence supports the case, that not all school systems use their funds as effectively as others. This fact raises a serious problem if one studies spending to understand the cost of an adequate education. Should the starting point be the current spending, accepting whatever is being done, or should there be some attempt to deal with the inefficiency issue? And should there be allowance for the fact that some districts, when given extra funds, will not use them productively to increase student performance? Without accurately identifying current inefficiencies by schools and without specifying how added resources for a district will be used, the costing out methods lack any predictive value.

In fact, the natural definition of “cost” is the *minimum* spending needed for a given outcome. It is likely in the case of schools that some districts spend more to achieve a given outcome than others do. Inefficiency is simply spending more than the least that is required. It is apparent why cost must refer to just the minimum spending to obtain a level of achievement, because otherwise the value of cost would be completely arbitrary, depending on the whim of what a district wanted to spend.¹⁹

The problem is that it has proven difficult, if not impossible,

19. In education discussions, efficiency often has a bad name, in part because it is taken to mean least cost without regard to the outcome. The classic misstatement of efficiency in education is found in Callahan (1962), which like many subsequent considerations failed to hold outcomes constant but instead looked at pure minimization of spending. The spending of two schools that are producing very different amounts of learning does not, by itself, say anything about the efficiency of the two schools unless, of course, the high producer is also the low spender.

for researchers to identify the true costs of meeting any goal. In fact, only rarely do studies mention possible inefficiency in spending, let alone attempt to deal with it.²⁰

The divergence between observed spending and true costs has been almost entirely ignored or dismissed in past judicial proceedings. One line of judicial rulings (e.g., Wyoming and Montana) even elevates the distinction to dizzying heights by declaring that any differences in the financing of districts must be “cost based,” while meaning for practical purposes “spending based.” When the court retains jurisdiction and financing decisions are regularly revisited to verify the “cost basis,” districts are given a clear incentive to increase their spending, regardless of the efficacy or efficiency of the spending.

An example of the idea of how inefficiency is bizarrely dealt with can be readily found from the referees in the New York City case. The plaintiffs offered the estimates of AIR/MAP (2004a), while the State, using the much lower estimates of Standard & Poor’s School Evaluation Service (2004), had suggested that it was reasonable to concentrate on the spending patterns of the most efficient of the successful schools—those that did well in student performance with lower expenditure. They thus excluded the top half of the spending distribution by successful districts in their calculations. But when the referees attempted to reconcile the state’s recommendation of \$1.9 billion with the AIR/MAP estimates of more than \$5 billion, they insisted on adding in all the high-spending districts, even when such districts did not produce better academic outcomes. After all, the referees reasoned, “there was no evidence whatsoever indicating that the

20. An exception is Gronberg, Jansen, Taylor, and Booker (2004). The academic studies of cost functions have concentrated more on efficiency issues but have been subject to potentially severe specification issues that bias the results.

higher spending districts . . . were in fact inefficient.”²¹ In other words, spending more to achieve the same outcomes should not be construed as being inefficient. One might then ask “what would indicate inefficiency?”

The importance of this is immediately obvious. If spending must be enough to raise achievement regardless of how efficiently resources are used, the answer is likely to be a very large number.

The existing studies are clearly best described as spending studies and spending projections, and not as cost studies. Accurate language is not, of course, used because even sympathetic readers and judges would question simple reliance on spending without a demonstration that the spending was effective. And indeed plaintiffs have been very effective in avoiding the discussion of this issue.

The deeper conundrum is that the courts cannot simply rule that districts should spend money well, particularly when the districts have no past experience with using resources well. Thus, if courts are restricted just to dictating spending levels, they are confronted with having to decide how to treat the inefficiency that is built into the conclusions derived from empirical evidence for a state. Dealing with such issues is generally far beyond the expertise of the courts.

21. John D. Feerick, E. Leo Milonas, and William C. Thompson, *Report and Recommendations of the Judicial Referees* (CFE, Nov. 30, 2004). Much of the testimony and discussion with the referees revolved around what proportion of the high spending (or high and low spending districts) was appropriately left out of the calculations. The S&P calculations omitted the top 50 percent of the spending distribution for schools that had sufficiently high achievement to be successful, while the plaintiffs argued that this was not the general norm of those who did this type of work. Again, because it is not a scientific procedure, there is no objective way to decide among alternative cutoffs for inefficient schools. In contrast, the “efficient schools” according to the econometric approach will be many fewer—generally less than a dozen, depending on the specific analytical model.

Minimum Costs and Costing Out Approaches

Analyzing the minimum cost needed to achieve any given outcome—the putative job of the costing out consultants—requires that cost estimation be built on the joint consideration both of program effectiveness and of costs. Obtaining an estimate of the minimum costs to reach the achievement goal is seldom even a consideration in the costing out studies. Ignoring this ensures that the results are biased above the true costs of adequacy. Indeed such a bias is a design feature of most of the work.²²

The professional judgment panels are generally instructed at the beginning of the process not to consider where the revenues would come from or any restrictions on spending. In other words, dream big—unfettered by any sense of realism or trade-offs. (Indeed, one reason for taking adequacy cases to the courts is that the democratic appropriations process necessarily takes these matters into account—and the courts might be induced to avoid them). But those instructions to the panels apparently do not always work to the satisfaction of consultants and clients. As Augenblick, Palaich, and Associates (2003) state about the operation of the professional judgment panels in North Dakota, “Finally, we should say that the members of all of the panels behaved in a way that can best be described as parsimonious.

22. Rebell (2006, 59) wants to define ignoring efficient spending as a purposeful virtue of costing out studies, perhaps because he realizes that they miss the mark in this area: “Although efficiency and accountability are obviously major public policy concerns which should be vigorously pursued, it is questionable whether cost analysis *per se* is an appropriate venue for pursuing these concerns. After all, the basic purpose of costing-out analysis is to determine what level of resources, *using the best mix of current practices* [his emphasis], will meet stated achievement goals. The extent to which major changes in current practices might produce acceptable results for lower costs is not part of the stated mission of these studies, nor could it be without postulating a set of hypothetical variables that would be inconsistent with the methodological premises of professional judgment panels and successful school district studies.”

. . . We worked hard to push people to identify resources they thought were needed to help students meet state and federal standards in spite of their natural tendency to exclude items because local voters might not approve of them or schools could 'get by' without them." This process, more openly acknowledged in this case than in others, hardly suggests a quest for the minimum expenditure needed to achieve an outcome.

Similarly, AIR/MAP (2004a) used a two-stage panel process in analyzing the New York adequacy case where a superpanel was given the various inputs of the separate panels and could, input by input, aggregate across the panels. This process ensures that any trade-offs between programs and resources of the individual panels are lost, and the process essentially arrives at the maximum resource use sketched by the panels and not at the minimum resource use.

But the apparent irrelevance of focusing on minimum cost is nowhere as clear as in an oft-repeated discussion in the state-of-the-art analyses. Allan Odden, before he began consulting on costing out studies, wrote that educational policy should recognize that improved performance could be obtained by redirecting existing expenditures and did not have to rely on added expenditure. Such an answer does not square with the orientation of many organizations purchasing costing out studies, which are uninterested in an answer that current resources are sufficient. (If so, they would be unlikely to incur the expense of a costing out study). This incongruence of past perspectives and funders' objectives apparently leads to their standard disclaimer (Odden, Fermanich, and Picus 2003):

Odden (1997) identified the costs of seven school wide designs that were created by the New American Schools. In subsequent analyses he showed how via resource reallocation, they were affordable at schools spending at the average or median level of expenditure per pupil in the United States (Odden & Busch,

1998; Odden & Picus, 2000). His analysis, however, did not include adequate planning and preparation time for teachers and did not standardize costs across various designs, so his 1997 cost figures are underestimated.

The standardization across designs refers specifically to the fact that some whole school models require less expenditure than others. The state-of-the-art costing out studies proclaim that in such a case one should use the *maximum* expenditure level for any of the models.

This spirit of maximizing expenditure also comes through in their programmatic recommendations. The specific programs (repeatedly recommended across states) include ones that, according to their evidence, have widely varying effectiveness and costs. Yet, instead of recommending programs that yield high achievement per dollar invested, the consultants recommend doing everything. Some parts of their program, however, would purportedly produce ten times the achievement of others for each dollar spent.

The expenditure function approach with few exceptions simply traces out the past spending of districts. Thus, unless one can assume that all districts are spending money wisely—an assumption broadly contradicted by existing research—these estimates cannot be interpreted as tracing out the minimum costs.²³

Only the successful schools approach potentially considers such issues if high-spending districts are trimmed from the sample of successful districts that are used to calculate the cost estimate. But even here there is no uniformity, and the study might trim not only high-spending but also low-spending districts.

23. Other techniques found in the scholarly literature have been developed to consider cost minimization (see Hanushek [2002]). Even when considered, the problem is that it is generally impossible to describe how efficiency is achieved (see Gronberg, Jansen, Taylor, and Booker [2004]).

In any event, there is no way for a court or the legislature to determine how it could require other districts to behave like the successful low-cost districts. One cannot realistically specify that spending must be effective—because the existing research and knowledge base in districts is insufficient to support that. Moreover, the expenditure function analyses that consider efficiency and the successful schools analyses may be able to point to districts that are doing relatively well, but they cannot describe why they are doing well or how some other district might be able to replicate their performance.

Projecting Outcomes to an Adequate Level

All costing out studies are motivated by an argument that achievement falls short of desired levels and thus it is necessary to provide the resources needed to reach the state goals. The important question for assessing costing out studies is whether they can describe policies and resources that will reliably lead to the new, higher achievement levels. None can.

States have developed varying goals, but many of the goals have not been thoroughly tested in the sense that it is known how to reach them. Indeed, as mentioned previously, it is popular to link costing out studies to achieving the goals of No Child Left Behind, even if NCLB is generally not an obvious constitutional requirement. And no state has yet shown how it will reach the goal of having all students “proficient” in core subjects by 2014.

The professional judgment approaches assume that because the goal was given in general terms to the panel at its inception, the panelists have come up with a model school that will produce the desired results. None of the reports ever discusses this or evaluates that possibility. In fact, just the opposite. When the reports are produced, there is generally a disclaimer that indi-

cates there is little reason to expect that students will actually achieve at these levels. Take, for example, the statement in the New York City study (AIR/MAP 2004a):

It must be recognized that the success of schools also depends on other individuals and institutions to provide the health, intellectual stimulus, and family support upon which public school systems can build. Schools cannot and do not perform their role in a vacuum, and this is an important qualification of conclusions reached in any study of adequacy in education. Also, success of schools depends on effective allocation of resources and implementation of programs in school districts.

This “warning label” contrasts sharply with the extraordinary claim in the November 2002 AIR/MAP proposal that their study would answer the question, “What does it actually cost to provide the resources that each school needs to allow its students to meet the achievement levels specified in the Regents Learning Standards?”

Indeed, the programs and resources incorporated in the professional judgment model and its subsequent costing are predicated on just what is needed to overcome the problems in the warning label. Yet when the time comes to describe how to interpret the finished product, the consultants do not want to be judged on whether the resources actually affect outcomes.

The state-of-the-art approach relies on the consultants’ conclusions about the best evidence on the effectiveness of different policies. The more recent versions of the evidence-based model (e.g., Odden, Picus, and Goetz [2006]) quantify their assessments of effectiveness of components that they include in their model school. This new information thoroughly impeaches the evidence and vividly shows its selective and biased nature. It also shows why the consultants do not use their own evidence to make any projections of achievement.

A way of seeing the problems with their work is simply to

take their analysis at face value. They design a school around a series of programs that have surface plausibility: lowered class size, full day kindergarten, expanded summer school, more professional development for teachers, and the like. For each component, they report what they believe to be the best evidence about how much achievement would be improved with each. They then advocate doing all of the components.

Looking at their evidence, however, it is easy to see why these consultants never provide an explicit projection of how achievement would improve with their model schools. The programs they advocate would, by their own reporting of the evidence, lift the achievement of the average student to beyond that of today's best performing student.²⁴ With the history of program outcomes in the past, it is obvious that the consultants' programs—which are simply repackaging of existing programs—will not have any such results. The easiest interpretation of this summary of their work is that the evidence is not reliable. But it also shows that the research evidence cannot provide predictions of how these overall “evidence-based” models will alter achievement.

Again, however, the authors design an “ideal” school that relies on their notions of research findings. These schools are not necessarily found anyplace in the state (where the actual schools could choose to follow such a model if they wanted to do so). The provision of resources is never accompanied by a court or legislative directive that requires the resources be used

24. The technical basis for this conclusion comes from their assessment of the “effect sizes” or the standard deviations of improvement in achievement that are predicted. (An effect size of 1.0 means that achievement would improve by one standard deviation; an improvement of one standard deviation would move the average student to the 84th percentile.) Their model school is reported to have a total effect size of 3–6 standard deviations, a completely implausible outcome that would place the average beyond the 99.9 percentile of the prior distribution.

in the way identified by the consultants. (That would probably be an even greater disaster.) Thus, providing the resources is unlikely to lead to any of the changes the consultants like, giving no reason to believe that student outcomes would increase at all.

The successful schools approach is fully rooted in the current operations of a state's schools and considers only average expenditure for the relevant group of successful schools. Therefore, it gives no information about how changing the level of spending might affect achievement. It can at best say something about meeting the generally high goals of NCLB that tend to drive court arguments *only if* some subset of schools is achieving the full standards at the time. But that appears to be unlikely.

There is no way to extrapolate the successful schools results from the currently observed outcomes of schools to a new level that is outside the range of observations on outcomes. Specifically, assume for illustration that the set of schools identified as successful has 70 to 80 percent of students reaching proficiency (which is perhaps well within current standards); there is no way to extrapolate these results to a 95 percent proficient standard.

A second extrapolation problem also occurs. When successful schools are identified just by proficiency levels on state tests, the schools identified as successful tend to have students from more advantaged families where the parents have provided considerable education to the students. The method concentrates on base spending for a typical successful school but then must indicate how much remedial spending would be needed to bring schools with students from more disadvantaged backgrounds up to the proficiency of the schools with better-prepared students. The appropriate way to do this is unclear, because again the situation is largely outside of the observations going into the successful schools analysis. The successful schools approach cannot provide any guidance to "unsuccessful" schools other than to

spend the same amount of money (which many already do with poor results).

The cost, or expenditure function, approach relates spending to student performance and student characteristics. Two factors are relevant. First, it interpolates the spending differences among very disparate districts. Thus, when there are large differences in the proportions of disadvantaged students as there are in New York State (the site of analyses by Duncombe, Lukemeyer, and Yinger [2004]), it relies strongly on the functional form of the underlying statistical relationship that connects the observations of districts. Second, and more important, it does not observe districts that achieve the levels of success that are considered in the evaluation of adequacy, leading to reliance on a simple linear extrapolation of the current observations of schools with no reason to believe that this will achieve the given ends. This problem is exactly analogous to the situation above with the successful schools analysis. The problems with extrapolation for success in schools with more disadvantaged students, identified for the successful schools approach, also hold in the cost function work.

The expenditure function analysis also does not identify programmatic ways of achieving outcomes. Instead, it assumes that just adding more of the resources observed (e.g., smaller classes or more experienced teachers) will lead to higher achievement. The version of expenditure functions that includes estimates of “efficient” districts is similar to the successful schools approach—districts doing relatively well are identified but poor performers are simply told that they could do better.

In summary, each approach lacks the information needed to project outcomes outside of those currently observed, but this is precisely what the costing out exercise demands. Again, however, this is not a problem with the execution of the analyses but instead is a fundamental roadblock to the analyses. There is ab-

solutely no reason to believe that the observed school operations provide a sufficient basis for projecting to outcomes outside of the current observations.

Incorporating Appropriate Input Prices

An integral part of all the approaches in arriving at their “cost” number is deciding on what prices to use for inputs, but the consultants not only lack appropriate data but also ignore the issues. In particular, using past salaries for teachers or past spending on administration and other inputs as the basis of calculations is inappropriate in almost all the circumstances of their projections. While some of the choices—particularly for modifying these inputs—sound reasonable, they introduce an arbitrariness that has significant effects on the resulting cost estimates.

If one wished to raise teacher quality, what would it cost? Clearly, the average salary, which is determined by the distribution of teachers of different experience levels and different amounts of academic training, cannot provide an answer to that question. What it would cost to improve teacher quality also depends markedly on whether one reproduces the current single salary schedule that does not recognize differences in quality or whether one contemplates a different pay and incentive scheme. It also depends on whether currently ineffective teachers can be replaced or whether it is necessary to wait until ineffective teachers decide to leave teaching. By considering just policies that involve adding resources to the current spending, the situation could in reality get worse. If all teachers, regardless of quality, are paid more, all teachers—including low-quality teachers—will have an incentive to remain teaching, and the ability to improve quality through replacement could become more difficult.

Such delineations of policy alternatives make it clear why the current typical behavior of a school district may not accurately indicate what improvements would cost if resources were used more effectively. It also underscores the difficulties of considering what can be done by only adjusting the funding of schools, and not considering other, more structural reforms.

The calculation of salaries is then a particularly interesting point of comparison across different studies. Sometimes the consultants simply use the average salaries for existing teachers (e.g., Odden, Fermanich, and Picus [2003]); other times they arbitrarily increase them by some amount (e.g., 10 percent in North Dakota in Augenblick, Palaich, and Associates [2003] and 18 percent in Arkansas in Odden, Picus, and Fermanich [2003]), vaguely arguing in terms of what other states spend; other times the bonus gets very high, such as the 54 percent advocated for New York City by Duncombe, Lukemeyer, and Yinger (2003), based on a regression comparison for New York districts in which New York City is a clear outlier in many of the dimensions of “uncontrollable” things such as density, poverty, and juvenile crime rates.

While the wide variance in teacher salaries has obvious and powerful effects on any cost estimates, none of these studies provides any evidence about the current quality of the teachers. Nor is there *any* research that relates teacher salary to quality, both in the ability to raise student achievement, and in the long-run supply of teachers of differing quality. So this becomes a whimsical adjustment based on the consultant’s vague sense of whether average salaries are high enough or not (for some unspecified quality level). And if they say they want to improve teacher quality, they simply increase the average salary by some arbitrary percentage.

Staying with existing structures and incentives—pervasive in all the methods—makes the reliance on average spending for the

components not specifically identified particularly dubious. For example, it is common to take existing central office and administrative expenditure as given and necessary. But there is no evidence that this is now the best way to organize schools or that it represents the minimum cost of providing a level of achievement.

The logic of developing estimates of the minimum cost of providing an adequate education calls for making decisions with an understanding of both the cost and the effectiveness of various inputs. The protocols of the costing out studies ensure that such decisions are never considered.

The Arbitrariness and Manipulability of Spending Estimates

While courts in various states have had differing responses to specific costing out studies, the general presumption is that all are trying to estimate basically the same thing—the resources required for adequacy. The results from past studies, however, indicate a clear arbitrariness in the results, which—because it is known to the parties supporting studies—permits strategic behavior and the manipulation of the results. Such circumstances compromise any claim to scientific underpinnings for the work.

The choice of approach for costing out is generally decided by the party requesting the work to be done. It appears that it might be a purposeful strategic choice, since many costing out studies are funded by parties with an interest in the outcome of the study (e.g., see Hanushek [2005]). For example, an analysis of differences across alternative analyses within the same state by the same researchers in four other states shows that the professional judgment method yielded estimates of “adequate” expenditure that were 30 percent above the successful schools

method (Baker, Taylor, and Vedlitz [2005]). This apparently has influenced the choice of method by clients.²⁵

A compilation of the estimated per-pupil expenditure for an adequate education across states and studies clearly indicates the arbitrariness of the estimates (Baker, Taylor, and Vedlitz [2005]). Even after adjusting for geographic cost differences across states and putting the estimates in real terms for 2004, the estimates differ by more than a factor of three. It is difficult to imagine what true underlying differences across states could drive such differences, since schools across the states look very similar, using similar curricula, approaches, and the like. But it is consistent with providing politically palatable estimates for the different state deliberations, because, for example, the citizens in many low-spending states would have difficulty accepting the current levels of spending in New York, let alone the post-judgment levels.

If the methods systematically produce very different results when addressing the same question, they obviously cannot be taken as giving a reliable and unbiased estimate of the resource requirements. Nor can they satisfy the most rudimentary criteria of scientific validity.

More Accurately Naming the Approaches

As with many concepts and ideas in school finance deliberations, the *nom de guerre* for each of the methods engenders

25. For example, Thomas Decker describes the choice of the professional judgment model for the costing out study to be commissioned by the North Dakota Department of Public Instruction: "The professional judgment approach we were aware would probably produce a higher cost estimate for achieving adequacy than successful schools." *Williston Public School District No. 1, et al v. State of North Dakota, et al*, Civil No. 03-C-507 (Dist. Ct., N.W. Jud. Cir. 2003 (Transcript of Deposition of Thomas G. Decker, August 17–18, 2005, 312).

confidence in the work, but it is a misplaced confidence. None of the names is accurate.

The professional judgment model relies on professional educators, but they generally lack expertise in designing programs to meet objectives outside of their experience. While they may have experience making trade-offs in current budgets, they do not have the research knowledge or personal experience to know how resources will change if they design a program for much higher student outcomes or of student body compositions that are outside their experience. But most important, they are asked to participate in a study where the outcomes of the study might directly affect their own pay, working conditions, and school situation, thus providing an incentive for them to distort whatever answers they might have. Thus, a much more accurate name of this approach is the *educators' wish list* model.

The state-of-the-art, or evidenced-based, model makes little effort to assess the accumulated evidence on different aspects of schooling. Instead, the highly selected evidence leads not to a scientifically grounded model but instead to the *consultants' choice* model. The results would vary dramatically if a different set of consultants, perhaps with a different focus, attempted to apply their understanding of the existing research base. In the end, the research base is simply too thin to have any consensus view about what an “evidence-based” school would look like (and, if that were not the case, it would be striking to find that none of the schools in the state already use the consultants' model).

The successful schools model begins with the identification of schools that are meeting some performance standard and then calculates the costs in an efficient subset of these successful schools. However, when the basis for judging school performance is student achievement, the resulting subset of schools conflates the various reasons why achievement may be high,

including family background and other peers in the schools. By relying on the observed performance for the “successful” set of schools, it has no way to project the results to a higher performance level. This approach is better labeled the *successful students* model, because it does not separate the success or failure of the school from other factors.

The cost function approach is designed to trace out the minimum costs for obtaining given outcomes. Unfortunately, this is true only if all school districts are operating efficiently—a situation that is known not to exist. The attempts of some to deal with inefficiencies have no general scientific foundation. These approaches capture the *expenditure function* for schools by identifying the average spending of districts with different achievement levels and student characteristics. They do not trace out the necessary cost of given performance levels, and thus cannot show the cost of an adequate education.

Evidence on the Results

The approaches to costing out produce an estimate of the resources needed to achieve an adequate education. For a variety of reasons, it is difficult to link these efforts to any results. First, courts and legislatures seldom faithfully enact the consultants’ dreams. Second, the consultants generally counsel not to take the results too seriously (see the AIR/MAP disclaimer above).²⁶

Augenblick, Palaich, and Associates (2003, II-3), go further in their analysis of North Dakota schools to discuss a lack of empirical validation of the professional judgment work. “The advantages of the approach [professional judgment] are that it reflects the views of actual service providers and its results are

26. This admonition is particularly strange in the state-of-the-art approach, however. They claim to have chosen the best methods based on research and evidence. If that is the case, shouldn’t it be mandated for all districts?

easy to understand; the disadvantages are that resource allocation tends to reflect current practice and there is *only an assumption, with little evidence, that the provision of money at the designated level will produce the anticipated outcomes*" (emphasis added).

While Augenblick, Palaich, and Associates (2003) did not look at the evidence, it is possible to do so in this case and in many other such costing out exercises. The authors use the professional judgment results to prescribe the spending for each of the K–12 districts in North Dakota in 2002. Two points are important. First, there is a wide variation in the calculated needs of districts. Second, sixteen districts were actually spending *more* in 2002 than the consultants (through their professional judgment panels) thought needed to achieve the full performance levels for 2014.

Because we have student performance information in North Dakota for 2002, we can see how performance is related to the fiscal deficits and surpluses that they calculate. It seems natural to think that districts with surplus expenditures are indeed performing above their achievement goals. It is also plausible to think that districts with smaller fiscal deficits are closer to achievement goals than those with larger fiscal deficits. (Note that the method and its application are designed to account for any different resource demands arising from the concentration of a disadvantaged population, school size, and the like—implying that the consideration of simple, bivariate relationships of deficits and performance are appropriate.)

A regression of reading or math proficiency percentages of North Dakota districts on the deficits indicates a statistically significant *positive* relationship. In other words, the larger the deficit, the higher is the student performance. Figures 7.1 and 7.2 plot calculated PJ (professional judgment) deficits against student achievement, immediately casting doubt on the value of the

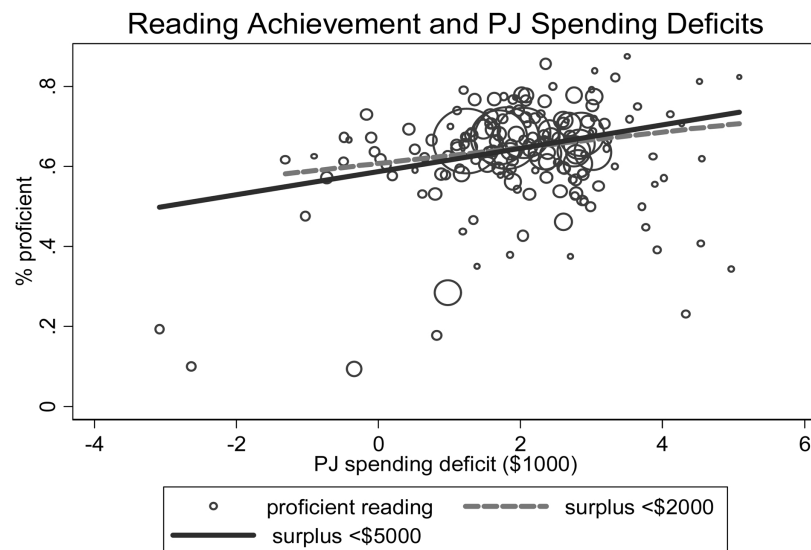


Figure 7.1 North Dakota School Districts' Professional Judgment Results (2002 Reading)

Note: Size of circles reflects student enrollment in each district.

professional judgment approach in this case. The solid line shows the regression of funding deficits on achievement.²⁷ Of course, because there are a few very large surpluses, the regression lines in the pictures could be distorted. But, the dashed line shows that a positive relationship between deficits and achievement still remains when all districts with surpluses greater than two thousand dollars are excluded from the calculations.²⁸

These are hypothetical exercises, however. It would be use-

27. By their method, the estimated needs should already account for differences in student backgrounds, and therefore the simple regression corresponds directly to their interpretation of the analysis. For this figure, the two school districts with surpluses greater than five thousand dollars per student are excluded. Including them would make the regression line even steeper.

28. Five districts out of the sixteen with identified surpluses have surpluses greater than two thousand dollars per student.

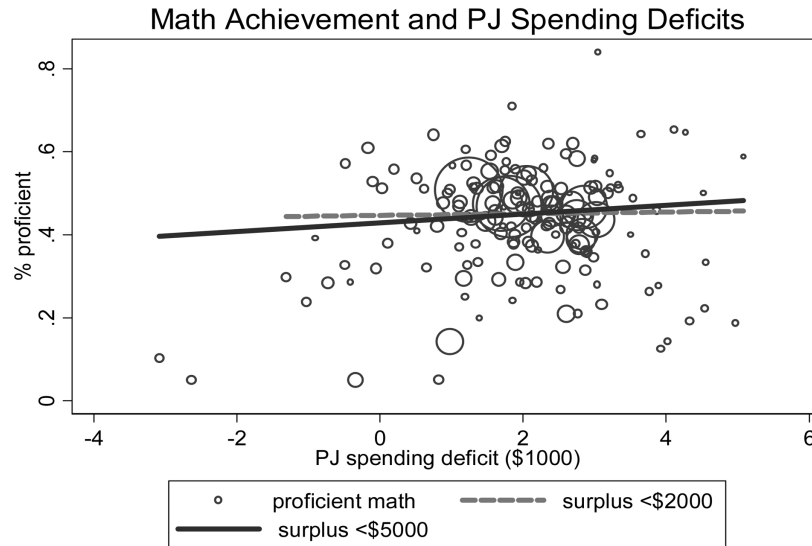


Figure 7.2 North Dakota School Districts' Professional Judgment Results (2002 Math)

Note: Size of circles reflects student enrollment in each district.

ful to see what happens when model results are introduced into actual decisions. This is difficult for a variety of reasons. First, while there is considerable current court activity, most of it has not fully worked through the courts and the legislatures and into the schools. Second, it is often difficult to obtain good comparisons to identify the effects of the court decisions.

Because Wyoming is tucked away out of sight of the East Coast media, few people outside of school finance insiders have followed the events of court decisions in Wyoming. But this example gives some insight into the effect of the adequacy decisions and court appropriations.

The Wyoming courts have considered the constitutionality of the school finance system since 1980. In *Campbell County School District v. State of Wyoming I* in 1995, the Wyoming

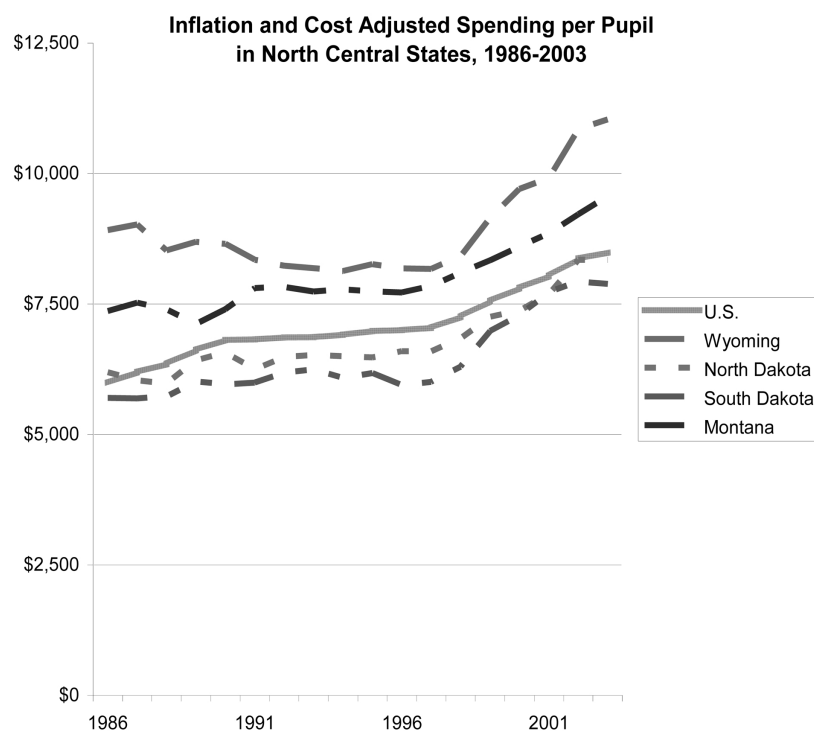


Figure 7.3 Spending History in North Central Comparison States

Supreme Court refined its schooling standard, as described in its subsequent 2001 decision:

This court made it clear it is the job of the legislature to “design the best educational system by identifying the ‘proper’ educational package each Wyoming student is entitled to have.” . . . Unlike the majority of states which emphasized additional funding, equalized funding, or basic education, Wyoming views its state constitution as mandating legislative action to provide a thorough and uniform education of a quality that is both visionary and unsurpassed. (*Campbell II*, 18)²⁹

29. *Campbell County School District v. State*, 907 P. 2d 1238 (Wyo. 1995); *Campbell County School District v. State*, 19 P.3d 518, 538 (Wyo. 2001).

This ruling presents a license for school districts to shop for virtually any program or idea that is arguably better than what they are currently using.

An element of this history that is important, however, is that the court has ruled that the school finance system must be “cost based” (which, as noted above, really means spending based). The legislature attacked this problem by asking MAP to develop a basic funding model, which it did based on an underlying professional judgment model. The basic model has been used in developing block grants to districts in order to meet differences in circumstances (disadvantaged populations, school size, etc.).

As might be imagined, this process of developing a visionary system—based on input criteria—leads to spending increases. Figure 7.3 compares Wyoming spending with that of a set of adjoining north central states (Montana, North Dakota, and South Dakota) and with the U.S. average.³⁰ The courts’ direct effect on spending is clear from this figure. Wyoming pulled away from the nation after the *Campbell I* decision in 1995. The

30. The other comparison states followed the normal democratic appropriations process and were not driven by court intervention in fiscal and policy decisions. Montana’s future may be very different, however. In Spring 2005, the Montana Supreme Court upheld a lower court decision that the state was in constitutional violation of its requirement to “provide a basic system of free quality public elementary and secondary schools.” *Columbia Falls Elem. School Dist. No. 6 et al v. the State of Montana*, No. 04-390 (Mont. S. Ct. Mar. 22, 2005). The District Court had identified the “major problems” in existing funding legislation as: “it provided no mechanism to deal with inflation; it did not base its numbers on costs such as teacher pay, meeting accreditation standards, fixed costs, or costs of special education; increases in allowable spending were not tied to costs of increased accreditation standards or content and performance standards; relevant data was already two years old when the bill was passed; and no study was undertaken to justify the disparity in ANB dollars [the ‘average number belonging’ entitlement] dispensed to high schools as compared to elementary schools. From these credible findings we must conclude that the Legislature did not endeavor to create a school funding system with quality in mind” (10). This reliance on input criteria could place Montana in a situation similar to Wyoming’s.

Table 7.1 Rankings on 2005 NAEP Tests for North Central Comparison States

	<i>Math</i>		<i>Reading</i>		<i>Science^a</i>	
	<i>grade 4</i>	<i>grade 8</i>	<i>grade 4</i>	<i>grade 8</i>	<i>grade 4</i>	<i>grade 8</i>
<i>All students</i>						
Montana	17	6	10	6	5	1
North Dakota	8	5	8	2	4	4
South Dakota	14	4	18	8	—	—
Wyoming	7	17	14	10	7	9
<i>Free- or reduced-lunch students</i>						
Montana	9	5	7	5	5	1
North Dakota	3	2	3	2	2	3
South Dakota	7	1	10	3	—	—
Wyoming	1	7	1	4	4	5

^a Science rankings for 2000 based on thirty-nine states for all students or thirty-eight states for free- or reduced-lunch students.

other states' spending patterns have not been dictated by judicial actions but instead have resulted from the democratic appropriations process. These patterns are significantly below those of Wyoming and follow roughly the national pattern.

The interesting thing is to observe the outcomes of Wyoming's court-supervised spending and how they compare with those of other states. The four north central states shown in figure 7.3 are remarkably similar in demographics, implying that simple comparisons of student achievement are appropriate.³¹ Table 7.1 provides rankings on the National Assessment

31. The comparison states have similar demographics, although Wyoming has some advantages in income and education of adults. Montana and South Dakota have lower incomes and higher child poverty rates, while Wyoming has the highest income and the most high school graduates of the adult population. All states have more than 85 percent white populations with larger American Indian populations in Montana, North Dakota, and South Dakota and a larger Hispanic population in Wyoming.

of Educational Progress (NAEP) of the comparison states in 2005 for math and reading and in 2000 for science. The top panel gives comparisons for all students, while the bottom panel is restricted to students on free and reduced lunch. In fourth grade, Wyoming tends to do better than the comparison states in math and in both math and reading for low-income students. But in eighth grade, two things are important. First, Wyoming does worse across the board than the comparison states. Second, even though Wyoming consistently (and increasingly) spends more for schools, the rankings generally drop from fourth to eighth grade. In contrast, rankings in the other states generally improve from the fourth to eighth grades. Moreover, while comparisons over time are more difficult, Wyoming student performance relative to the nation declined from 1992 to 2005 in fourth grade reading and math and in eighth grade math.³²

Table 7.2 provides comparisons on the measures of school retention and college continuation. North Dakota and South Dakota, the two lowest-spending states, consistently outperform Montana, with Wyoming performing noticeably worst on each of these outcome measures.

Although some may interpret this record as saying that it is necessary to wait longer and to mandate even more spending, the Wyoming performance information to date gives little indication that this would be a productive path.

Interestingly, under the court mandate to periodically recalibrate the spending for schools, Odden et al. (2005) investigated the funding of Wyoming schools in 2005. They concluded that the current spending—already fifth highest in the nation in 2003—was some 17 percent shy of adequate.³³ They presum-

32. While all states participated the NAEP for math and reading in 2005, only a subset voluntarily participated in the earlier grades. Eighth grade reading was not assessed until 1998 but did improve between 1998 and 2005.

33. State rankings adjust for cost of living calculated from a wage index for

Table 7.2 School Attainment for North Central Comparison States
(State Rankings in Parenthesis)

	<i>Montana</i> (%)	<i>North Dakota</i> (%)	<i>South Dakota</i> (%)	<i>Wyoming</i> (%)
Ninth graders' chance for college by age 19	42.5 (12)	61.8(1)	48.1 (6)	40.4 (20)
College continuation rate of high school graduates	54.7 (30)	73.7(1)	60.9 (13)	55.1(29)
Percent of adults aged 18–24 with high school diploma	91.1 (10)	94.4 (2)	92.0 (5)	86.5 (31)
Percent 9th–12th graders who dropped out	4.2 (22)	2.2(1)	3.9 (14)	6.4 (42)

Source: U.S. Department of Education, Digest of Education Statistics, 2003. Washington, DC: National Center for Education Statistics, 2004; NCHEMS Information Center, <http://www.higheredinfo.org/dbrowser/index.php?submeasure=62&year=2002&level=nation&mode=data&state=0>.

ably believe that another substantial dose of funding would produce more than the last dose of funding, but never go so far in their lengthy report as actually projecting an improvement in student achievement. And, indeed, the Wyoming legislature in 2006 voted appropriations that exceed even the Odden et al. (2005) spending plans, moving Wyoming perhaps to the highest spending state in the nation. The legislature did not, however, specify that schools must put in place the “evidence based” programs, just that they get sufficient money that they could permit it by the consultants’ calculations.

The North Dakota and Wyoming data are not isolated in-

nonteaching college graduates. The new estimates, according to the Access account (<http://www.schoolfunding.info/news/policy/1-6-06WYcoststudy.php3> as accessed on March 10, 2006), would increase unadjusted spending from \$9,965 per pupil to \$11,635 per pupil. These calculations correspond to increasing total spending by \$142 million to approximately \$987 million, a 17 percent increase.

stances of lack of achievement gains from spending. Evers and Clopton (chapter 4) provide a series of case studies that involve significant spending increases—some from judicial actions and some from normal appropriations—that were unaccompanied by any gains in student outcomes. Their case studies cover large and very well-observed districts that nevertheless failed to use dramatically larger resources in effective ways.

Outcomes versus “Opportunity”

As previously noted, virtually none of the reports says that the calculated level of resources will yield the outcomes that the consultants are striving to obtain. When it comes time to write the reports—and to produce a document by which the consultants might be judged—the language generally changes to providing an “opportunity” to achieve the standard, not actually achieving the standard.

The motivation for the underlying costing out analyses is that children are not learning at a putative constitutional level (or an NCLB level or a state standards level), but the reports never say explicitly that the resources identified in the study are either necessary or sufficient to achieve these levels. Instead, they say that the resources will provide an opportunity to reach the standards.

This change of language means that the consultants are not predicting any level of achievement if the stated resources are provided. None of the reports states that the added resources will yield achievement that is any higher than currently observed. The reports provide no predictions about outcomes, and thus they are *completely unverifiable*. Said differently, there is no scientific basis for deciding among alternative “cost” estimates, because the data on student outcomes are not informative.

By implication of the report language, a wide range of spending could produce the same level of student outcomes. For example, why not project added spending of \$10 billion a year (as opposed to \$5.6 billion a year) for the New York City *CFE* case as the amount that would provide an opportunity to achieve some undisclosed higher student achievement? Why not \$1 billion a year?

The obfuscation about what is being calculated is easily seen in the AIR/MAP report for the CFE litigation in New York. Remember that the report is entitled “The New York Adequacy Study: Determining the Cost of Providing All Children in New York an Adequate Education.” Since an adequate education is defined in terms of student outcomes, one might think that this implies that students provided with the specified resources would achieve the adequate outcomes (in this case, achieve the Regents Learning Standards). Moreover, the report is laced with language suggesting that the AIR/MAP consortium is considering actual student achievement and not some more ethereal concept:

- In describing the purpose of the report, the AIR/MAP team states, “To remedy this injustice, Justice DeGrasse ordered a number of reforms. As a first, ‘threshold task,’ he charged the state with assessing ‘the actual costs of providing a sound basic education in districts around the State’” (AIR/MAP 2004b, 6).
- Subsequently, when describing what was the objective, the report states, “32 organizations from throughout the state came together to initiate a one year, cutting-edge costing-out study—supported by grants from several major national foundations—that will determine the actual amount of funding needed in each school district to provide an adequate

education to all students throughout the state” (AIR/MAP 2004b, 6).

- Finally, in instructing the professional judgment panels, it states, “Specifically, your task is to design adequate instructional and support programs for students in Kindergarten through twelfth grade that you are confident will meet the expectations specified in Exhibit 1 for the student populations described in the assumptions listed below” (AIR/MAP 2004b, 64). Exhibit 1 then discusses both the NCLB student outcome standards and what is necessary to reach the Regents Learning Standards.

The language is qualified, however, whenever a reader might infer that some explicit outcome is being considered in the analysis. For example, when the goals related to Regents Learning Standards are mentioned in the AIR/MAP report, they are prefaced with “an opportunity to achieve.”³⁴ Nonetheless, there is little doubt that the reader is intended to interpret this as the actual student outcomes to be expected from providing the added resources.

This situation is not specific to the AIR/MAP report but pervades all the methods and all the available reports. The possible exception is some of the successful student or expenditure projection studies, where the authors might suggest that a given school could achieve a given level of performance *if it could figure out why some other school achieved that level and if it could reproduce it in another setting*. Yet no guidance on either the source of achievement or the way to reproduce it is ever given.

34. The judicial referees who declared that \$5.63 billion a year was the right number consistently use the “opportunity” language. Perhaps knowing that this spending is unlikely to produce actual achievement of the kind they believe represents a sound basic education, the referees also call for regular costing out studies on a four-year cycle.

The translation of the objective into an undefined opportunity is particularly problematic when, like Wyoming, the finance system is supposed to be cost (i.e., spending) based. *Cost* necessarily refers to what is needed to purchase some good or service. But if the good or service to be purchased is undefined, and if it could mean a broad range of different things, the cost must logically also be undefined—because there is no way to link an observable outcome to an expenditure.

Instead of attempting to parse the very careful language of all the costing out studies, however, consider the opposite perspective. If the costing out studies do not provide a clear view of the outcome that would be expected, they become just the whim of the consultant—even when based on a method that has previously been applied or has a “scientific” air to it. There is no way to judge among alternative spending projections based on any evidence that will become available about outcomes, thus putting each projection in the category of personal opinion and not science. There is no obvious reason for giving deference to the personal opinion of consultants hired by interested parties in the debates.

This work also does not help the political and legislative debate on school finance. The studies are designed to give a spending number. They do not indicate how achievement is likely to be different from the current level if such an amount is spent. Neither do they suggest how achievement (or even opportunity) would differ if a state spent 25 percent more or 25 percent less than the consultants’ personal opinions about what should be spent.

Conclusions

The traditional focus of courts on equity in school finance, defined simply as funding for schools, has given way to one on

outcomes and adequacy. And this has moved the courts into areas in which they are completely unprepared. Specifically, if one wants to improve outcomes or change the distribution of outcomes, how can the court do it? After all, even if the courts want to do so, they cannot simply mandate a given level of student achievement. Instead they must define any judgments in terms of instruments that will lead to their desired outcomes but that can be monitored by the court. This necessity returns the decision making to a focus on money and resources.

But how much money translates into the desired schooling outcomes? For this, the courts have come to rely on outside consultants (frequently hired by interested parties) to provide the answers.

These consultants, and the people who hire them, suggest that the subsequent “costing out” exercises provide a scientific answer to the disarmingly simple question, “how much does it cost to provide an adequate education?” Nothing could be farther from the truth. The methods that have been developed are not just inaccurate. They are generally unscientific. They do not provide reliable and unbiased estimates of the necessary costs. In a variety of cases, they cannot be replicated by others. And they obfuscate the fact that they are unlikely to provide a path to the desired outcome results.

As Augenblick, Myers, Silverstein, and Barkis (2002) eloquently state in their study, which was the basis of the Kansas judgment, “None of these approaches are immune to manipulation; that is, each is subject to tinkering on the part of users that might change results. In addition, it is not known at this point whether they would produce similar results if used under the same circumstances (in the same state, at the same time, with similar data).” This possibility gives considerable latitude to the courts to pick whatever number they want. Judge Bullock in his Kansas decision speaks favorably of the Augenblick & My-

ers cost estimates (with the above caution included), while justifying his choice in part by noting that a parallel ruling in Montana opined:

The testimony of Dr. Lawrence Picus of the University of Southern California (who also testified for Defendants in the instant action) was found to lack credibility in that, while testifying for the defense in Kansas and Massachusetts he had opined those systems were equitable and thus constitutional, but in Montana (while testifying for the plaintiffs) he opined Montana's funding was inadequate and violative of constitutional requirements—both opinions being based astonishingly on undisputed numbers showing Montana's system more equitable in virtually every measurement than either Kansas or Massachusetts. In other words, Dr. Picus “danced with the girls that brought him.”³⁵

Costing out studies are political documents, almost always purchased by clients with an agenda. When there are no accepted scientific standards for their conduct, when there are few empirical restraints, when they cannot be replicated by others, when the outcomes of any changes cannot be verified based on observed data, and when there is no requirement for consistency across applications, it should come as little surprise that the estimates please the party who has purchased them.

The history of the use of costing out studies in the New York City case highlights the political nature of such studies. During the original trial, the defense sought to introduce a professional judgment analysis of the costs of an adequate education in New York. It concluded that the school district's existing \$10-billion-plus budget was sufficient to meet the constitutional requirements for a sound basic education. The plaintiffs successfully argued that the approach had not been shown to be generally

35. *Montoy v. State*, Case No. 99-C-1738 (Dist. Ct. Shawnee County, Kan., Dec. 2, 2003).

scientifically accepted and that it was inadmissible hearsay, leading to rejection of the study conducted by the MAP consulting firm. The plaintiffs then hired the same firm, MAP, along with another consulting firm to cost out an adequate New York City education, although this time based on the plaintiffs' specifications of what was adequate. The judicial referees received the plaintiffs' report and passed it back to the judge with none of the qualms that had led the judge originally to exclude such testimony or analysis.

Courts need guidance if they are to enter into the adequacy arena, because they have no relevant expertise in the funding, institutions, and incentives of schools. They are generally eager to have somebody tell them the answer, so they are willing to jump on "the number" even while recognizing that it might have problems.

The message here is that the existing costing out methods do not and cannot support such judicial decision making. There is also the distinct possibility that pursuing such a policy will actually worsen rather than help students and their achievement.

The methods provide spending projections, based crucially on existing educational approaches, existing incentive structures, and existing hiring and retention policies for teachers. Essentially, each calls for doing more of the same—reducing pupil-teacher ratios, paying existing teachers more, retaining the same administrative structure and expense. These are just the things that districts have been doing for the past three decades.

On the other side, *none* of the existing costing out studies claims that providing the resources they call for will have *any* effect on achievement. They very carefully skirt a statement that would tie them to results, couching explicit spending figures in the vague and undefined language of "opportunity." And for good reason. Past experience provides plentiful evidence of in-

stances where funding was increased with no fundamental change and where student performance did not change. The consultants know well that even if we take a large leap of faith and believe that the programs they describe will be effective, nobody enforces the adoption of these programs when resources are added. But if the “required” spending for an adequate education is not related to an expectation about student outcome, what is the meaning of the spending that is called for? We know that it is possible to get no results while spending even more. Couldn’t we also get no results by spending less?

There is a pernicious result, however. It is not just that money is wasted by investing in ways that have no payoff. Following the recommended spending projections reinforces and solidifies the existing structure of schools that has not produced, almost certainly to the detriment of students. They offer only a blind and unsupported hope of bringing about the kinds of improvements that they purport to cost out.

References

- AIR/MAP. 2002. *A Proposal for Determining Adequate Resources for New York Public Schools*. American Institutes for Research and Management Analysis and Planning (November 27).
- AIR/MAP. 2004a. *The New York Adequacy Study: Determining the Cost of Providing All Children in New York an Adequate Education, Volume 1—Final Report*. American Institutes for Research and Management Analysis and Planning (March).
- . 2004b. *The New York Adequacy Study: Determining the Cost of Providing All Children in New York an Adequate Education, Volume 2—Technical Appendices*. American Institutes for Research and Management Analysis and Planning (March).
- Augenblick & Myers, Inc. 2002. *Calculation of the Cost of an Adequate Education in Indiana in 2001-2002 Using the Professional Judgment Approach*. Prepared for Indiana State Teachers Association. Denver, CO: Augenblick & Myers, Inc. (September).

- Augenblick, John, and John Myers. 1997. *Recommendations for a Base Figure and Pupil-Weighted Adjustments to the Base Figure for Use in a New School Finance System in Ohio*. Denver, CO: Augenblick & Myers, Inc. (July).
- Augenblick, John, John Myers, Justin Silverstein, and Anne Barkis. 2002. *Calculation of the Cost of a Suitable Education in Kansas in 2000-2001 Using Two Different Analytical Approaches*. Prepared for Legislative Coordinating Council. Denver, CO: Augenblick & Myers, Inc. (March).
- Augenblick, Palaich, and Associates, Inc. 2003. *Calculation of the Cost of an Adequate Education in North Dakota in 2002-2003 Using the Professional Judgement Approach*. Bismarck, ND: North Dakota Department of Public Instruction (September).
- Baker, Bruce D., Lori L. Taylor, and Arnold Vedlitz. 2005. "Measuring Educational Adequacy in Public Schools." Bush School Working Paper #580, Texas A&M University (September).
- Callahan, Raymond E. 1962. *Education and the Cult of Efficiency*. Chicago: University of Chicago Press.
- Duncombe, William, Anna Lukemeyer, and John Yinger. 2003. "Financing an Adequate Education: A Case Study of New York." In *Developments in School Finance: 2001-2002*, edited by William J. Fowler, Jr., 127-154. Washington, DC: National Center for Education Statistics.
- . 2004. "Education Finance Reform in New York: Calculating the Cost of a 'Sound Basic Education' in New York City." CPR Policy Brief No. 28/2004, Center for Policy Research of Syracuse University, New York (March).
- Education Week*. 2005. "Quality Counts 2005: No Small Change—Targeting Money toward Student Performance." *Education Week* (Washington, DC). January 6.
- Gronberg, Timothy J., Dennis W. Jansen, Lori L. Taylor, and Kevin Booker. 2004. *School Outcomes and School Costs: The Cost Function Approach*. Texas A&M University.
- Hanushek, Eric A. 1999. "The Evidence on Class Size." In *Earning and Learning: How Schools Matter*, edited by Susan E. Mayer

- and Paul E. Peterson, 131–168. Washington, DC: Brookings Institution.
- . 2002. “Publicly Provided Education.” In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein, 2045–2141. Amsterdam: Elsevier.
- . 2003. “The Failure of Input-Based Schooling Policies.” *Economic Journal* 113, no. 485 (February): F64–F98.
- . 2005. “Pseudo-Science and a Sound Basic Education: Voodoo Statistics in New York.” *Education Next* 5, no. 4 (Fall): 67–73.
- . Forthcoming. “The Alchemy of ‘Costing Out’ an Adequate Education.” In *School Money Trials: The Legal Pursuit of Educational Adequacy*, edited by Martin R. West and Paul E. Peterson. Washington, DC: Brookings Institution.
- Myers, John L., and Justin Silverstein. 2005. *Successful School Districts Study for North Dakota*. Augenblick, Palaich, and Associates, Inc. (August 15).
- Odden, Allan, Mark Fermanich, and Lawrence O. Picus. 2003. *A State-of-the-Art Approach to School Finance Adequacy in Kentucky*. Lawrence O. Picus and Associates (February).
- Odden, Allan, Lawrence O. Picus, and Mark Fermanich. 2003. *An Evidence-Based Approach to School Finance Adequacy in Arkansas*. Lawrence O. Picus and Associates (September).
- Odden, Allan, Lawrence O. Picus, and Michael Goetz. 2006. *An Evidence-Based Approach to School Finance Adequacy in Washington*. Lawrence O. Picus and Associates (June 30).
- Odden, Allan, Lawrence O. Picus, Michael Goetz, Mark Fermanich, Richard C. Seder, William Glenn, and Robert Nelli. 2005. *An Evidence-Based Approach to Recalibrating Wyoming’s Block Grant School Funding Formula*. Lawrence O. Picus and Associates (November 30).
- Picus, Lawrence O., Allan Odden, and Mark Fermanich. 2003. *A Professional Judgment Approach to School Finance Adequacy in Kentucky*. Lawrence O. Picus and Associates (May).
- Rebell, Michael A. 2006. “Professional Rigor, Public Engagement, and Judicial Review: A Proposal for Enhancing the Validity of

Science Violated

311

Education Adequacy Studies.” Paper presented at the *Annual Meeting of the American Education Finance Association*. Denver, CO.

Reschovsky, Andrew, and Jennifer Imazeki. 2003. “Let No Child Be Left Behind: Determining the Cost of Improving Student Performance.” *Public Finance Review* 31 (May): 263–290.

Standard & Poor’s School Evaluation Service. 2004. *Resource Adequacy Study for the New York State Commission on Education Reform* (March).

Verstegen and Associates. 2003. *Calculation of the Cost of an Adequate Education in Kentucky: Prepared for The Council for Better Education* (February).