What Is Adequate?

Operationalizing the Concept of Adequacy for New York

Thomas Downes

Department of Economics

Tufts University

February 2004

What Is Adequate?
Operationalizing the Concept of Adequacy for New York

Thomas Downes
Tufts University

Abstract

In New York, as in a number of other states in which recent state court decisions have invalidated the existing system of school financing, the court decision has focused on the failure of the existing system to provide each student with an adequate ("a sound, basic") education. While the legal strategies that have led to these decisions have been prompted, in part, by long-standing discussions in the research literature on how to best design intergovernmental aid systems, there is no consensus in the literature on how to operationalize the concept of adequacy. The court decisions do not give policy makers the luxury of waiting for consensus in the academic literature; they must settle on a method for structuring an aid system that insures that each student has access to an adequate education.

This paper will attempt to provide with some guidance on how to accomplish this difficult task. The starting point will be a working definition of adequacy. The paper will then provide a critical evaluation of methods that have been proposed for determining the spending districts in a state would need if they are to provide an adequate education. Particular attention will be paid to the lessons that can be learned from the efforts in other states and in New York to operationalize the concept of adequacy. The ultimate goal of this discussion is to suggest how existing methodologies could be combined to produce a defensible, replicable, and understandable method for determining adequacy.

The paper will then turn to several cross-cutting issues that must be addressed no matter which method is used to determine the adequate level of spending in each district. Among these issues will be the possibility of creating perverse incentives by compensating for factors in that can be influenced by districts and the need for avoiding a system of adequacy determination that appears to dictate to districts and schools specific input mixes. Particular attention will also be paid to the need for updating, to account for advances in research, for changes in standards, and for the availability of new or better data.

I. Introduction

In New York, as in a number of other states in which recent state court decisions have invalidated the existing system of school financing, the decision in the *Council for Fiscal Equity, Inc. v. State of New York* case has focused on the failure of the existing system to provide each student with an adequate ("a sound, basic") education. While the legal strategies that have led to these decisions have been prompted, in part, by long-standing discussions in the research literature on how to best design intergovernmental aid systems, there is no consensus in the literature on how to operationalize the concept of adequacy. The court decisions do not give policy makers the luxury of waiting for consensus in the academic literature; they must settle on a method for structuring an aid system that insures that each student has access to an adequate education.

This paper will attempt to provide some guidance on how to accomplish this difficult task. The starting point will be a working definition of adequacy. The paper will then move to a critical evaluation of methods that have been proposed for determining the spending districts in a state would need if they are to provide an adequate education. Particular attention will be paid to the lessons that can be learned from the efforts in New York to operationalize the concept of adequacy. In the New York case, these lessons will be crystallized through comparisons of the implications of three alternative methods that have been used to calculate estimates of the resources needed to provide each student in the state with access to an adequate education. The ultimate goals of this discussion are to attempt to identify sources of the differences in estimates of adequate spending and to suggest how existing methodologies could be combined to produce a defensible, replicable, and understandable method for determining adequacy.

The paper will then turn to several crosscutting issues that must be addressed no matter which method is used to determine the adequate level of spending in each district. Among these issues will be the possibility of creating perverse incentives by compensating for factors in that can be influenced by districts and the need for avoiding a system of adequacy determination that appears to dictate to districts and schools specific input mixes. Particular attention will also be paid to the need for updating, to account for advances in research, for changes in standards, and for the availability of new or better data.

II. What Level of Spending Is Adequate?

The starting point for any discussion of adequacy must be the simple reality that what is adequate in any single state must ultimately be determined by the voters in that state. In the literature, there is general acceptance of the argument that the standards that the students in a school district must attain determine the adequate level of funding for that district. The decision as to what level of education is adequate is a policy decision. The national acceptance of this premise is evident in the provisions of the *No Child Left Behind Act of 2001*, which assigns to the elected leaders in each state the responsibility for establishing state-specific standards in reading, mathematics, and science.

In most states, the standards that have been promulgated include both minimum resource requirements and minimum levels of academic performance that the students in each district

must attain. Since differences across districts in input costs and student needs translate into differential ability to meet those standards (Duncombe and Lukemeyer, 2002), most researchers agree that, if the type of comprehensive aid system described in Sonstelie (2004) or in Duncombe (2002) is to provide all students with access to an adequate education, that system must compensate for these differences. For example, all else equal, a high-poverty school district will not be able to attain the same performance levels as a low-poverty school district if those districts have the same resources. Thus, if all students in a state are to have access to an adequate education, higher poverty districts must receive more aid than lower poverty districts, all else equal.

Since accounting for the differential ability of districts to meet standards is a necessary part of the determination of the amount a district must spend to meet these standards, the process of calculating adequate levels of spending can naturally be decomposed into two steps no matter what methodology underlies the actual calculations (Duncombe, 2002). First, given the minimum resource and student performance requirements, determine the spending to attain these minimums in at least one benchmark district. Second, adjust cost of adequacy in the benchmark district to reflect different characteristics of other school districts. If only minimum resource requirements are in place, then the adjustments will require determining differences in prices that will have to be paid to purchase equal quality resources across the state. Since the majority of educational expenditures are for personnel, at a minimum these adjustments must account for variation in cost-of-living and working conditions. If the state's standard also includes minimum performance requirements, then there must be additional implicit or explicit adjustments for any factors that affect student performance, such as student poverty and the sparsity of the school district's population.

III. Critique of Approaches to Estimating the Cost of Adequacy

This section presents a brief review of different methods that have been proposed for determining the cost of providing and adequate education. While the discussion will touch on all of the methods that have been suggested, the focus will be on the two methods that have become dominant in the research literature, the professional judgment and the statistical cost function methods. Given the current state of knowledge, none of the methods used alone provides definitive estimates of the cost of providing an adequate education. Improved estimates could be generated by using the cost function and professional judgment approaches in combination, since using the methods in combination addresses one or more of the flaws of each of these methods.

Geographic Cost Variation Approach

The simplest method for accounting for cost variation is to determine the resources needed in the benchmark district and then to use either a simple (consumer price index) or sophisticated (Chambers' (1998) geographic cost of education index) input price index to adjust for differences across the state in the cost of acquiring those resources. Rubinstein (2003) uses a variant of this approach to determine the amount each state would need to spend to provide an adequate education.

Since this methodology involves no explicit linking of resources to student performance

standards, the methodology is not appropriate for costing-out an output standard.[1] Existing evidence indicates that this approach may do a very poor job of accounting for critical variation in the cost of educating students, which is unsurprising since there is no explicit linking of resources to performance. Taylor and Keller (2003) show that, in Texas, input price indices result in very different cost adjustment recommendations than do cost-of-education indices; Duncombe and Lukemeyer (2002) find much the same thing in the New York context.

The results presented in Duncombe (2002) raise, however, another possibility; the choices made by researchers who construct input price indices may matter at least as much as does the decision to use an input price index rather than a cost-of-education index. Duncombe (2002) presents both a cost-of-education index and an overall teacher wage index, based on a hedonic wage model. Clearly evident are substantive differences between the cost-of-education index and the teacher wage index. But the implied differences in cost adjustments are far smaller than are the differences in cost adjustments implied by the cost-of-education and Chamber's (1998) geographic cost of education index.[2] Since the same basic methodology is used by Duncombe to generate his teacher wage index and by Chambers to calculate his geographic cost of education index, the differences in the implications of these two indices cannot be attributed to fundamental methodological differences. Instead, the differences must be due to differences in the manner in which the methodology is applied.

In estimating the hedonic equations that underlie his Geographic Cost of Education Index, Chambers explicitly excludes from the regression equation any "factors [that] are outside local control" (AIR and MAP, 2004, p. 13). Duncombe, on the other had, includes in his specifications a number of factors, like average district student performance and average class size, that can be influenced by school district choices. Duncombe then calculates his teacher cost index using "the average for the discretionary factors and, the actual district levels for the factors outside their control." (Duncombe, 2002, p. 8) These differences in the specifications of the hedonics are attributable to different judgments of how to best construct cost indices that are not susceptible to manipulation by school districts. There is broad agreement in the literature that any methodology used to estimate cost indices must distinguish costs that can be controlled by a recipient government from those that cannot. In general, aid should compensate only for costs that are beyond control of the recipient governments. Otherwise, recipient governments receiving aid do not have an incentive to reduce costs when it is possible to do so. (Downes and Pogue, 2002) A concrete example of perverse incentives that are frequently built into aid formulae are adjustments that provide more aid per capita to jurisdictions serving small populations. In these cases, aid compensates in part for the higher costs associated with small scale and reduces the incentive to consolidate or pursue other strategies that could take advantage of scale economies (Duncombe and Yinger, 2001).

---

[1]The professional judgment approach is best understood as an attempt to modify the geographic cost variation approach to accommodate output standards.

[2]This conclusion is based on calculated cost adjustments generated in William Duncombe's New York School Aid Simulation, available at http://www-cpr.maxwell.syr.edu/faculty/duncombe/special%20report/aid%20simulation%20on%20the%20web.xls. Throughout this paper the New York School Aid Simulation will be used to generate comparable distributions of cost-adjusted spending.

The approaches taken by both Chambers and Duncombe generate cost indices that do not create perverse incentives. The reality, however, is that these alternative approaches do not represent two methods that are both arguably correct. The hedonic methodology will only generate valid cost indices if the hedonic equation includes all of the characteristics that matter to consumers and suppliers of the good.[3] Omission from the hedonic equation of relevant characteristics results in biased estimates of the implicit prices of the included characteristics and in a hedonic specification that will be ill suited to characterize variation in salaries or in other input prices. Further, as Duncombe shows, it is feasible both to include all relevant characteristics in the hedonic equation and to generate teacher cost indices that are not susceptible to manipulation.

In what follows is discussion of results that will accentuate the relevance for the policy debate in New York of the differences in the Chambers and Duncombe approaches. These results will also show that, while the choice of a method for calculating the regional cost index matters, what appears to matter more is correctly implementing the chosen method.

Empirical Identification Approach

The empirical identification method builds on the seemingly plausible idea that the expenditures of districts already meeting a state's performance standard provide a clear indication of the amount that must be spent to provide an adequate education. Thus, to implement this method, researchers identify districts presently meeting the standard and then measure how much these districts are spending. To mitigate the influence of extraordinary cases, most practitioners cut the outlier districts (those that spend the least and the most). For the remaining benchmark districts, average (either mean or median) per pupil spending is calculated. Finally, to determine adequate spending in districts outside the benchmark group, benchmark per pupil spending is adjusted for differences in costs of doing business or higher student needs.

A variant of the empirical identification approach was used by the State Aid Work Group (2004) to generate the cost-adjusted expenditures that underlie the Regents Proposal on State Aid for 2004-05. The standard chosen was a performance level of at least 80 percent, where each district's performance level was calculated by taking a simple, unweighted average across seven statewide tests[4] of the percent of test-takers scoring at Level 3 or above.

While basing adequacy calculations on spending needed in districts that have already attained the standard has considerable intuitive appeal, the empirical identification approach has fallen into disfavor because of several problems that cannot be easily fixed. First, particularly if an ambitious standard is established, districts that have attained the standard are not representative of "typical" districts – this is a best practice approach. In other words, these districts are meeting standards when other districts are failing for a reason. That reason could be differences in the cost structures of the districts. But that reason could also be differences in the degree to which districts make efficient use of available resources or differences in other unobservable

[3]Chambers decision to omit certain characteristics may result from his conclusion that these characteristics are statistically endogenous. The appropriate solution to an endogeneity problem is not, however, to omit key characteristics from the hedonic equation. Instead, appropriate statistical methodologies should be used to correct for the endogeneity problem (Downes and Zabel, 2002).

[4]The tests in question were fourth grade English Language Arts, fourth grade Mathematics, high school Mathematics A, high school Global History, high school U.S History, high school English, and high school Earth Science.

characteristics of districts that make these successful districts atypical. Duncombe (2002) shows that, in the New York context, the extent to which successful districts are atypical (the "selection problem") increases as the standard increases. And, while researchers have implemented various ad hoc fixes to account for the fact that successful districts represent a select group of districts, researchers have not agreed on agreed on a standard method for accounting for selection.

Second, to adjust benchmark spending for cost differences, practitioners of this method often make a regional cost adjustment and use a set of pupil weights to reflect differing student needs. However, as the Texas evidence discussed above indicates, this is likely to be insufficient.

Whole School Design Method

Proponents of this method (e.g., Odden 1997, Odden and Busch 1998) have argued that experience from successful school reform efforts can be used to determine the expenditures needed to provide an adequate education. For example, Odden (1997) uses the expenditures for each school participating in the Modern Red Schoolhouse program to generate an estimate of the amount that would be needed to provide an adequate education. He chooses the Modern Red Schoolhouse program because of its documented success in improving student performance.

Again, the main flaw in this approach is basing an estimate of adequacy on a highly select sample. For example, if the only schools that had implemented the Modern Red Schoolhouse program were in an urban setting, the calculated implementation costs may be a poor indicator of the costs needed to implement the program in a rural setting. More importantly, accumulated evidence on whole school reform methods indicates that reforms that may be very successful in some settings may fail in others, in part because of unique features of those successful settings.

Professional Judgment/ Resource Cost Model (RCM) Approach

With this approach, researchers ask professional educators what resources they think are required to achieve an adequacy standard. Hailed as a "bottom up" approach to estimating the cost of adequacy, researchers design one or many prototypical classrooms, schools, and districts by asking professional educators ("experts") what resources are required for an organization to meet a specific standard. These "experts" provide great detail, provided from their best judgments based on their own experiences. This method focuses on the first step of Duncombe's two-step process: estimating the cost of adequacy in a prototype district, as defined by bundles of resources. Spending for the benchmark district (or districts) is then adjusted for input cost differences across districts, typically though the use of an input price index.

Guthrie and Rothstein (1999) believe that this approach best meets the needs of states attempting to ensure that all students are provided the opportunity to obtain an adequate education. They do not believe that some of the more statistical approaches are valid, partly because the outcome measurements used in the cost function approach, typically standardized test scores, only measure a small part of what defines an adequate education (math and reading), which is only a part of the full "basket of education goods and services."

Some practitioners of the professional judgment approach, including Guthrie and Rothstein (1999) and the AIR/MAP team that has prepared an adequacy study for New York, argue that research should be used to guide the deliberations of the panels. In the New York context, the

panels were provided with "an objective description of mainstream educational research as background for [their] deliberations." (AIR and MAP, p. 4) While, in the abstract, this approach sounds reasonable, the reality is that there is not consensus in the literature on the links between school and class size and student performance.

This approach, though thorough and detailed, is extraordinarily time-consuming and hard to coordinate. Further, in order to make the approach manageable, the panels cannot be asked to consider the full range of contexts in the state. As a result, for districts other than the benchmark districts the calculation of spending to meet the standards requires adjusting benchmark spending using a cost adjustment methodology. In New York, the principle cost adjustment used by the AIR/MAP team was a variant of Chambers' geographic cost of education index. The evidence from Texas suggests that only adjusting using a geographic cost of education index will not be sufficient.[5]

Further, participants in professional judgment panels are never asked to contemplate trade-offs, which means that participants tend to adopt a pie-in-the-sky view of the world. Sonstelie (2001) and Rose, Sonstelie, Reinhard and Heng (2003) suggest that this problem can be avoided if a two-step approach is used, with professionals first being asked to make resource allocation choices given a certain budget and then being asked to indicate which budget would enable them to satisfy certain standards.[6]
This professional judgment approach also suffers from two forms of human bias: self-serving behavior and habit. Some of the "experts" may have their own reasons for wanting funding to be at a certain level, and thus may act in ways that serve their own purposes but not necessarily the purpose of the project. While the remaining experts on the panel could serve as a discipline, the absence of a budget constraint substantially reduces the incentive for any participants to argue strenuously for compromises. The easiest strategy is to accede to the recommendations of the self-serving experts.

In addition, these "experts" are guesstimating – they are using their experience of current schools and districts and classrooms, which may or may not be ideal or a good base for understanding what is actually *needed.*

Several of the existing professional judgment studies highlight the extent to which the results the methodology generates are dependent on personal experience. For example, in a report that uses the professional judgment methodology to cost-out an adequate education in Nebraska, Augenblick and Myers (2003) include a table that summarizes, for elementary schools in a relatively large school district, the personnel recommendations generated by professional judgment panels in Nebraska, Kansas, Maryland, and Montana. What is striking about this table is the extent of variation in the recommendations, even though contexts were intended to be very similar. Because the process typically imposes constraints on class size, a point to which we will return below, the recommended number of teachers did not vary dramatically across the states.

[5] The preliminary report indicates that adjustments were also made for variation in students needs, but how these adjustments were made is unclear.

[6] In a personal communication, Jon Sonstelie indicated that he and Heather Rose have documented cases in which, given explicit outcome standards, the same set of professionals will make different resource recommendations when they do and do not face explicit budget constraints.

But the number of teacher aides per 1000 students ranged from 0 in Maryland to 25.7 in Nebraska. The number of librarians and media specialists ranged from 2 in Maryland to 5 in Kansas, and the guidance counselors ranged from 1.4 in Nebraska to 5 in Kansas. Total personnel recommended ranged from 80 in Maryland to 102.9 in Nebraska, with a recommendation of 89.7 in Montana.

In theory, these differences could reflect differences in state standards, but it seems unlikely that standards variation explains the variation in this case, since the standards used by Augenblick and Myers for Montana and Nebraska were very similar, as were the characteristics of the prototype schools. Further, in the Montana case, the output standards presented to the professional judgment panels were based on statewide performance results, not on district-specific performance information. As a result, professionals with experience in high-performing or low-performing schools would have little, if any, experience on which to draw to determine the resources needed for a prototypical school to achieve the performance gains specified in the standard.

Cost Function Approach

This method focuses on developing accurate adjustments for student needs and resource price differences. Researchers estimate (using econometric methods and actual data) cost functions relating data on actual spending in a district to student performance, resource prices, student needs and other relevant district characteristics. From the regression results, cost indices can be created for each district. These cost indices indicate how factors outside a district's control affect the spending required to reach a given student performance level.

There are three main problems with this approach. First, of all of the approaches, it is the least intuitive to non-economists and the least understood. Estimation of plausible cost functions requires the use of complex statistical techniques, which are typically difficult to explain to someone unfamiliar with multivariate statistics.[7] Second, the data quality must be extremely good, since the entire model is based on real historical data. Without accurate and complete data, it is difficult to obtain accurate results. Finally, the method has been called a "black box," as researchers do not dictate *how* funds should be spent, but rather simply *how much* should be spent.

Duncombe (2002) uses this approach in their study of New York State. In estimating the cost function, he uses as the measure of student performance each district's performance on the 4th and 8th grade math and English tests and the high school Regents Exams in math and English. After estimating the cost function, he calculates how much a district that has average everything would have to spend to reach a particular performance standard. Adequate spending in other districts is calculated by multiplying spending in the benchmark district by a cost index

[7] Guthrie and Rothstein (1999) argue that the cost function methodology fails the "ease of understanding" standard. No one can dispute the validity of their argument; what is at issue is the conclusion that this failure means that the other methodologies should be used in place of the cost function methodology, even if the cost function methodology is theoretically sound and is most likely to generate valid estimates of the spending levels needed to meet the standard. Taken to its extreme, this argument implies that, in choosing a method to determine adequate spending levels, one is better off choosing a method that is easy to understand but wrong rather than a method that is difficult to explain but that produces the right answers.

calculated for each district from the cost function estimates.

In order to account for the possibility that some districts may not be getting as much as possible out of their available resources,[8] researchers have considered both direct and indirect methods for accounting for inefficiency. For example, Duncombe (2002) includes in his cost function a measure of inefficiency in resource use. Each district's cost index is calculated with that district's efficiency measure replaced with the average level of efficiency. Downes and Pogue (1994) also outline methods that could be used to generate cost indexes that were performance-based. They argue that the estimation methodology they use could, indirectly, account for some unobserved determinants of costs, unobserved outputs, and inefficiency.

Is There a Single Right Method?

Given the econometric advances of the last decade, the cost function approach is the most likely to give accurate estimates of the within-state variation in the spending needed to attain the state's chosen standard, if the data are available and of a high quality. The complexity of the method makes it likely, however, that policy makers will reject using the cost function methodology alone. For that reason, the best alternative may be to use the cost function and professional judgment methods in combination. The cost function methodology could be used to generate estimates of the amount of spending each district would need to provide an adequate education. Then, as has been suggested by Sonstelie (2001) and Rose, Sonstelie, Reinhard and Heng (2003), these estimates could then be used as constraint under which the professionals would need to operate. The professionals would be asked if, given the resources implied by the cost function methodology, achieving the state standards is feasible.

Using the professional judgment approach in a checking-the-estimates capacity can solve one of the flaws of the cost function approach by making more transparent the process by which adequate spending levels are determined. Combining the methodologies also can mitigate the effect of the inherent flaw in the professional judgment approach, that it relies on people so exclusively. Even experts are inherently biased and faulty, particularly when these experts are using their own experiences to recommend how to create the perfect school, having never seen the perfect school.

IV. How Do the Answers Compare? The New York Context

As the preceding discussion indicates, variants of the empirical identification, the professional judgment, and the cost function approach have all been implemented in New York. Comparing the results of these approaches is, however, difficult for three reasons. First, the outcome standards differ from approach to approach. Second, the AIR/MAP report provides adequate spending estimates by need to resource categories only.

Nevertheless, crude comparisons were possible. These crude comparisons shed some light on the differences across methodologies in the distribution of spending if each methodology in question was used to provide the estimates of the spending needed to attain the standards. In

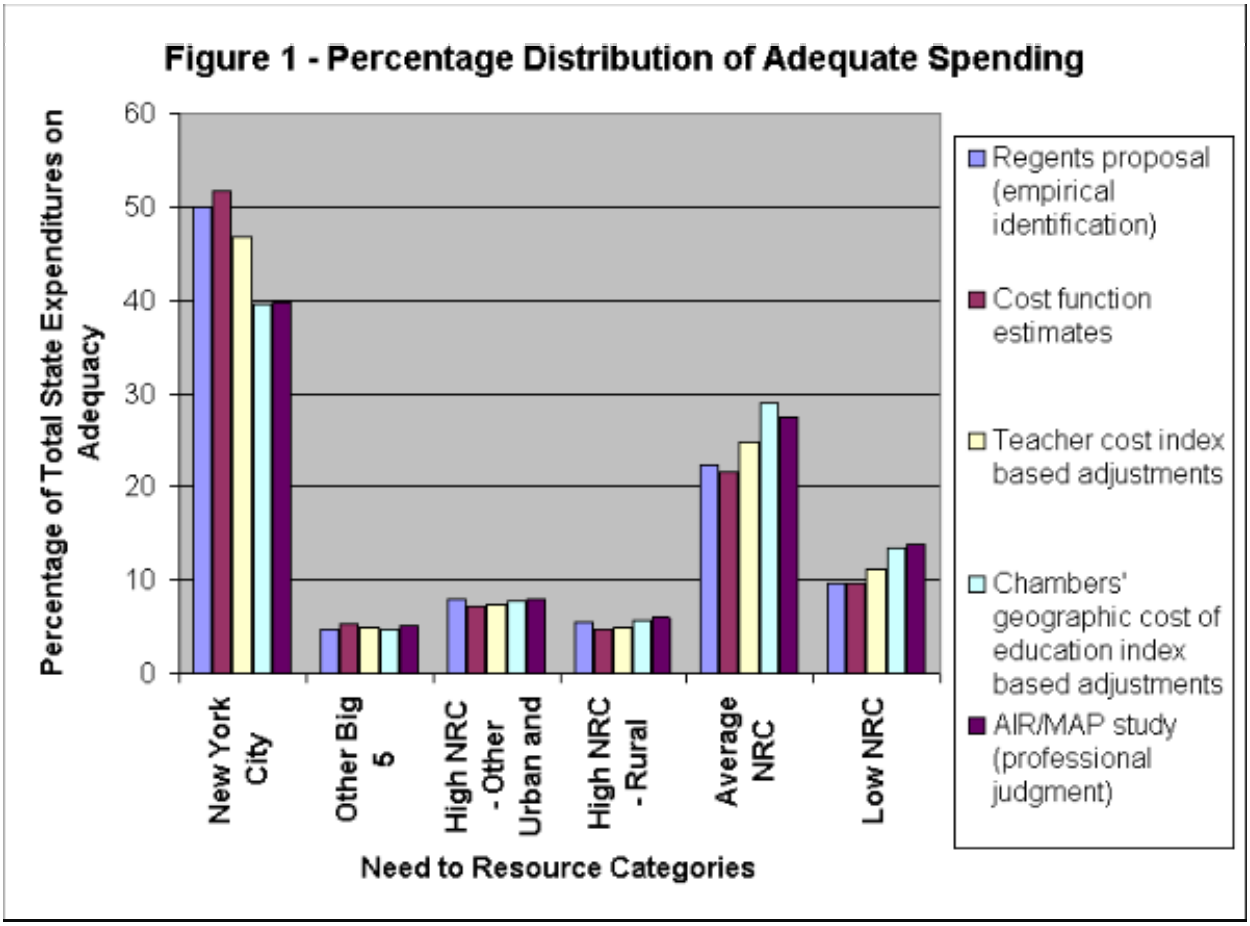
[8] Districts that are not getting as much as possible out of their available resources are said to be technically inefficient.

addition, the comparisons offer some insight into the factors that lead to different distributions under different methodologies. As more results become available, particularly from the AIR/MAP project, the sources of differences will become more apparent.

William Duncombe's New York School Aid Simulation (available at http://www-cpr.maxwell.syr.edu/faculty/duncombe/special%20report/aid%20simulation%20on%20the%20web.xls) was used to generate comparable data that could help clarify why the distribution of spending implied by the AIR/MAP analysis differs so dramatically from that results from Duncombe's cost function estimates or from the Regents' School Aid Proposal. In order to come close to the standard used in the Regents' School Aid Proposal, a standard of 160 was used to generate all of the spending numbers produced by the New York School Aid Simulation. Matching the standard used in the AIR/MAP project was not possible, since no concrete standard was used in that project.[9]

Figure 1 gives the distributions of spending across the need to resource categories. The most striking result evident in this figure is that basing an aid formula on the AIR/MAP results would lead to a substantially lower fraction of spending in New York City than would either the Regents' Proposal or an aid formula based on cost function estimates. Until the final AIR/MAP report is available, it is not possible to determine to what extent the use of the professional judgment methodology contributes to the differences apparent in Figure 1. But, the results from the New York School Aid Simulation for the teacher cost index and the geographic cost of education index lend credence to the argument that these methods generate different distributions primarily because of differences in the cost adjustments used to transform benchmark district spending into each district's spending.

[9] In the "Statement of Desired Educational Outcomes" that was provided to the panels, the panelists were told, "Your job is to design an instructional program that will provide all students a full opportunity to meet the Regents Learning Standards, and to attain a Regents' diploma." (AIR and MAP, p. 23)



Similar distributions of spending are implied by the AIR/MAP study and by the simulation results with cost adjustments made using the geographic cost of education index. This similarity is not surprising, since the same methodology was used to generate both the AIR/MAP cost adjustments and the Chamber's geographic cost of education index. What is surprising is how different these distributions are from the distribution that would result if an alternative input price index, Duncombe's teacher cost index, is used. With the same standard using the same underlying data, the teacher cost index and the geographic cost of education index generate very different distributions. The above–discussed methodological flaws of the geographic cost of education index appear to have led the AIR/MAP analysts to produce flawed estimates of the amount each district would need to spend to attain state standards.

V. Crosscutting Issues in the Determination of Adequacy

Sonstelie (2004) makes clear that the amount each district would need to spend to provide an adequate education will be a critical component of any aid formula, independent of the method used to determine the adequate level of spending. As a result, any method used to determine the adequate level of spending must address certain design features that have been discussed in the intergovernmental aid literature. This section contains a brief discussion of these design considerations; more expansive discussion of design considerations can be found in Louis,

Jabine, and Gerstein (2003).

The Incentive Problem

As was noted above, there is general acceptance in the intergovernmental aid literature that aid should not compensate for factors that can be influenced by the behavior of the governments receiving aid. (Louis, Jabine, and Gerstein, 2003) For example, in the literature on aid to compensate for the costs associated with special education, much attention has been paid to the problematic incentives created by an aid formula that compensates for the number of students in a district who have been identified by the district as needing special education services. (Cullen, 2003) Many states have responded to the existence of these problematic incentives by modifying their aid formulas so that the aid a district receives does not depend directly on the number of students in the district who have been identified as needing special education services.

In theory, then, the calculation of the amount each district would need to spend to provide adequate services should be independent of the choices that district makes. In practice, breaking the link between aid and district choices may be very difficult to do, since many of the factors that account for cross-district variation in district costs are also affected by district choices. The canonical example is district size, which depends on decisions to pursue consolidation. But some observers have argued that even such cost factors as the fraction of students eligible for school lunch subsidies can be influenced by the aggressiveness of the identification procedures within the school district.

Downes and Pogue (1994) and Duncombe (2002) have argued that one strategy for avoiding the incentive problem is to make cost adjustments dependent not on a district's actual values of these discretionary factors but on regional or statewide means. While this approach would solve the incentive problem, since district behavior can have little or no influence on regional or statewide means, this approach also leads to incomplete adjustment for cost variation.

A second alternative is to turn to the research literature for evidence on the extent to which district behavior does appear to be sensitive to these perverse incentives. The research by Lankford and Wyckoff (1996) on the extent to which special education assignments by districts in New York were sensitive to the incentives in the aid formula offers a nice example of the kind of research that can guide policy makers designing aid formulas. If the research evidence indicates that district behavior is not particularly sensitive to the type of incentives that would exist if aid formulas fully compensate for the impact on costs of certain factors, then the calculation of adequate spending should compensate fully for these factors. Hopefully, the trend towards designing aid formulas that compensate for discretionary cost factors will stimulate more research on the extent to which school districts respond to the incentives in these aid formulas.

Are the Results of the Adequacy Determination Process Normative?

If the professional judgment approach is used to determine the spending each district needs to provide an adequate education, then for the benchmark district one of the results of the process are the input levels the professionals feel are necessary to meet the state standards. While the process is not designed to dictate to districts how available resources should be spent, there is a real risk that policy makers inside each district will view as prescriptive the input levels that

result from the deliberations of the professional judgment panels. Given the nature of the process, a seemingly persuasive argument can be made that the input levels that the panel members feel are necessary for the benchmark district should be treated as minimum input levels for other districts in the same group as the benchmark district.

This argument ignores, however, the logic that underlies the foundation aid systems into which the adequacy recommendations are an input. The philosophy of the foundation aid system is that districts should be provided with the resources they need to provide an adequate education and then allowed to decide how to best utilize those resources. The input mix that may be appropriate for one district with one set of local conditions may not be right for another district that looks similar in some important dimensions but that differs in other, difficult to quantify dimensions. As a result, care must be taken to emphasize the fact none of the methodologies used to determine spending needed to meet state standards is designed to micro-manage districts.[10]

Updating the Adequacy Calculations

Almost method that could be used to determine adequate spending levels will be time-consuming and, potentially, costly. Thus, the natural inclination will be to use the results of the initial study as the basis of aid calculations for several years, with the only substantive year-to-year changes being inflation adjustments that are uniform across the state. Following this natural inclination could, however, be very bad policy. This conclusion does not mean, however, that the initial adequacy study will need to be replicated on a yearly basis.

Over time, new data will become available, implying changes in the relative status of districts. Simple inflation adjustments will not account for changes in the relative status of districts. Since, however, most methods use a process for accounting for differences in relative status that can easily incorporate new data, recalculating adequate spending amounts can easily be done no matter which method is used.

State standards are also likely to evolve. Depending on the method used to determine the spending levels needed to satisfy the standards, accounting for changes in these standards is more difficult. If, as is the case with the cost function methodology, the method used makes no explicit reference to the standards in determining spending needed in the benchmark district, then a change in standards can be handled as easily as can new data. If, as is the case with the professional judgment methodology, spending in the benchmark district is determined in reference to specific standards, then a change in the standards necessitates a new adequacy study.

Adequacy calculations should also be updated as the practice of determining adequate levels of spending improves. Major changes in acceptable practice call for newer adequacy studies that use methods consistent with acceptable practice. Less dramatic changes in practice may, however, require only minor changes in the formula used to calculate adequate spending levels. For example, if the State Aid Work Group's method (2004) is used to determine adequate spending levels, then, as new research is conducted on the links between spending, outcomes,

[10] If a relatively black box approach like the cost function methodology is used to determine adequate spending levels, it is less necessary to emphasize that the results of the process are not prescriptive.

and the fraction of a district's students who are at-risk, the results of that research should be used to update the weightings in the formula.

VI. Concluding Remarks

No single method of estimating the spending needed for adequacy has been universally accepted. This paper presents a critical evaluation of the methods for determining adequacy that have been proposed. The principal implication of this evaluation is that drawing on the strengths of several methods may prove to be the only way to craft a methodology that produces results that are understandable and defensible. In particular, by combining the cost function and the professional judgment methodologies, an approach could be crafted that is understandable to policy makers and is not subject to the types of biases inherent in a process that is based not on data but on the opinions of well-meaning but fallible experts.

As importantly, the results of this paper make clear that seemingly small methodological differences can translate into dramatic differences in spending distributions. Since undoing policy mistakes is so hard, it is imperative that policy makers get things close to right at the outset. And getting things right involves carefully thinking about the incentives that are implicit in aid formulas based on adequate spending estimates derived using any one of the methodologies. Of equal importance is developing a strategy for updating the estimates of the amounts districts would need to spend to provide an adequate education.

VII. References

American Institutes for Research and Management Analysis and Planning, Inc. "The New York Adequacy Study: 'Adequate' Education Cost in New York State." Unpublished report, February 2004.

Augenblick and Myers, Inc. "Calculation of the Cost of an Adequate Education in Montana in 2001-2002 Using the Professional Judgment Approach." Unpublished report, Denver, 2002.

Augenblick and Myers, Inc. "Calculation of the Cost of an Adequate Education in Nebraska in 2002-2003 Using the Professional Judgment Approach." Unpublished report, Denver, 2003.

Chambers, Jay G. "Geographic Variations in the Public Schools' Costs." National Center for Education Statistics Working Paper No. 98-04 (Washington, DC: U.S. Department of Education, National Center for Education Statistics, February 1998).

Cullen, Julie Berry. "The Impact of Fiscal Incentives on Student Disability Rates." Journal of Public Economics 87 (August 2003): 1557-89.

Downes, Thomas A. and Pogue, Thomas F. "Adjusting School Aid Formulas for the Higher Costs of Educating Disadvantaged Students," National Tax Journal 47 (March 1994): 83-102.

Downes, Thomas A. and Pogue, Thomas F. "How Best to Hand Out Money: Issues in the Design and Structure of Intergovernmental Aid Formulas," Journal of Official Statistics 18 (December

2002): 329-352.

Downes, Thomas A. and Zabel, Jeffrey. "The Impact of School Quality on House Prices: Chicago 1987-1991," Journal of Urban Economics 52 (July 2002): 1-25.

Duncombe, William. "Estimating the Cost of an Adequate Education in New York." Center for Policy Research Working Paper Number 44, Maxwell School of Citizenship and Public Affairs, Syracuse University, February 2002.

Duncombe, William D., and Lukemeyer, Anna. "Estimating the Cost of Educational Adequacy: A Comparison of Approaches." Mimeograph, Syracuse University, March 2002.

Duncombe, William D. and Yinger, John. "Does School District Consolidation Cut Costs?" Center for Policy Research Working Paper Number 33, Maxwell School of Citizenship and Public Affairs, Syracuse University, January 2001.

Duncombe, William D. and Yinger, John. "Performance Standards and Educational Cost Indexes: You Can't Have One without the Other." In Ladd, Helen F; Chalk, Rosemary; Hansen, Janet S, editors, Equity and Adequacy in Education Finance: Issues and Perspectives (Washington, D.C.: National Academy Press, 1999), pp. 260-97.

Guthrie, James W. "Twenty-First Century Education Finance: Equity, Adequacy, and the Emerging Challenge of Linking Resources to Performance." Mimeograph, Vanderbilt University, 2001.

Guthrie, James W. and Rothstein, Richard. "Enabling 'Adequacy' to Achieve Reality: Translating Adequacy into State School Finance Arrangements." In Ladd, Helen F; Chalk, Rosemary; Hansen, Janet S, editors, Equity and Adequacy in Education Finance: Issues and Perspectives (Washington, DC: National Academy Press, 1999), pp. 209-259.

Lankford, Hamilton and Wyckoff, James H. "The Allocation of Resources to Special Education and Regular Instruction," Ladd, Helen F., editor, Making Schools Accountable: Performance-Based Approaches to School Reform (Washington, DC: Brookings Institution, 1996), pp. 221-57.

Louis, Thomas A.; Jabine, Thomas B.; and Gerstein, Marisa A., editors, Statistical Issues in Allocating Funds by Formula (Washington, DC: National Academies Press, 2003).

Odden, Allan R. "The Finance Side of Implementing New American Schools." Report for the New American Schools, Alexandria, VA, 1997.

Odden, Allan R. and Busch, Caroline. Financing Schools for High Performance. (San Francisco: Jossey-Bass, 1998)

Rose, Heather; Sonstelie, Jon; Reinhard, Ray; and Heng, Sharmaine. High Expectations, Modest Means: The Challenge Facing California's Public Schools (San Francisco: Public Policy

Institute of California, 2003).

Rubinstein, Ross. "National Evidence on Racial Disparities in School Finance Adequacy." In William J. Fowler, Jr., editor, Developments in School Finance: 2001-02 (Washington, D.C.: National Center for Education Statistics, 2003), pp. 91-109.

Sonstelie, Jon. "Toward Cost and Quality Models for California's Public Schools." In Sonstelie, Jon and Richardson, Peter, editors, School Finance and California's Master Plan for Education (San Francisco: Public Policy Institute of California, 2001), pp. 103-123.

Sonstelie, Jon. "Fiscal Federalism, State Academic Standards, and the Adequacy of Public School Resources." Mimeograph, University of California, Santa Barbara, February 2004.

State Aid Work Group, New York State Education Department, "Regents Proposal on State Aid for School Districts for 2004-2005." Unpublished report, Albany, January 2004.

Taylor, Lori L. and Keller, Harrison. "Competing Perspectives on the Cost of Education."In William J. Fowler, Jr., editor, Developments in School Finance: 2001-02 (Washington, D.C.: National Center for Education Statistics, 2003), pp. 111-126.