



AMERICAN INSTITUTES FOR RESEARCH®

Review of Kansas Education Cost Studies

Jesse Levin (AIR)

March 2, 2018

Submitted to: Gordon Self
Kansas Legislative Coordinating Board
300 SW 10th Ave., Ste. 370-W
Topeka, KS 66612-1504

Submitted by: American Institutes for Research

Dun and Bradstreet Number: 04-173-3197

Tax Identification Number (TIN) 25-0965219

Notice of Trademark: "American Institutes for Research" and "AIR" are registered trademarks. All other brand, product, or company names are trademarks or registered trademarks of their respective owners.

American Institutes for Research

1000 Thomas Jefferson Street NW, Washington, DC 20007-3835 | 202.403.5000 | TTY 877.334.3499 | www.air.org

Review of Kansas Education Cost Studies

Jesse Levin (AIR)

March 2, 2018



1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Copyright © 2016 American Institutes for Research. All rights reserved.

Table of Contents

1 – Introduction	1
2 – Costing-Out Study Objectives and Traditional Approaches.....	1
Objectives of a Costing-Out Study	1
Costing-Out Approaches.....	2
Cost Functions.....	3
Professional Judgment.....	4
Successful Schools.....	5
Evidence-Based	7
Summing Up the Different Approaches.....	7
3 – Review of Calculation of the Cost of a Suitable Education in Kansas in 2000-2001 Using Two Different Analytic Approaches (Augenblick, Myers, Silverstein & Barkis, 2002).....	9
Study Methodology	9
Professional Judgment Approach (Input-Oriented Approach).....	9
Successful Schools Approach (Outcome-Oriented Approach)	11
Key Results and Discussion	12
Key Results	12
Discussion.....	14
4 – Review of Elementary and Secondary Education in Kansas: Estimating the Costs of K- 12 Education Using Two Approaches (Kansas Legislative Post Audit Committee)	34
Study Methodology	34
Expenditure Analysis (Input-Oriented Approach)	34
Cost Function Approach (Outcome-Oriented Approach).....	35
Key Results and Discussion	36
Key Results	36
Discussion.....	40
References	46

1 – Introduction

The debate surrounding school finance in Kansas and specifically the question of how much is necessary to allow for the *suitable* provision for the financing of the state’s public education system has been at the forefront of policy discussion for years. Fueled by a series of court cases, most notably the series of cases known as *Montoy v. State* and more recently *Gannon v. Kansas* has resulted in various research efforts to better understand what constitutes a suitable education and how much would it cost to provide this to all students in the state. Two of these efforts are the following studies:

- 1) Calculation of the Cost of a Suitable Education in Kansas in 2000-2001 Using Two Different Analytic Approaches (Augenblick and Myers, Inc., 2002)
- 2) Elementary and Secondary Education in Kansas: Estimating the Costs of K-12 Education Using Two Approaches (Kansas Legislative Post Audit Division, 2006)

In addition, a new study is currently underway by the labor economist Dr. Lori Taylor. The purpose of this report is to provide a review of items 1) and 2), above, focusing on the methodology used in each and corresponding results to better understand the qualities of each and inform the current discussion surrounding the forthcoming remedy ordered by the Kansas State Supreme Court. A similar review of the study being developed by Dr. Taylor will be conducted after it has been finalized.

The report is organized as follows. Section 2 provides an overview of both the objectives of educational costing-out studies and the traditional methodological approaches used to perform cost studies. Sections 3 includes a review of the study performed by Augenblick and Myers, Inc. (A&M). Sections 4 provides a review of the study conducted by the Kansas Legislative Post Audit Division (LPA).

2 – Costing-Out Study Objectives and Traditional Approaches

Objectives of a Costing-Out Study

The need for costing-out studies is clear given the clauses found in virtually all state constitutions that dictate that the state has a responsibility to provide an education that is considered *adequate, sufficient* or some other term that represents a level that allows all students an opportunity to achieve the outcomes expected of the public education system (Baker & Green, 2014). If states are to follow through on this obligation, then it is necessary to understand both the amount of effort involved in terms the public funding required to offer educational sufficiency and how to appropriately distribute this funding. More formally stated, the main objectives of educational costing-out studies are to answer what have been referred to as the two fundamental questions of educational adequacy (Chambers & Levin, 2009):

- What does it cost to enable a public school system to provide all students with an adequate education?
- How can state school finance systems allocate their resources equitably, such that all students are afforded an adequate education regardless of their need or circumstance?

It important to note that these questions are neither simple to answer nor wholly independent from one another. First, we acknowledge that while the questions are conceptually separable,

adequacy and equity are inextricably linked in school finance.¹ While determining how much additional investment in education is *necessary* to provide an adequate educational opportunity, calculation of this bottom-line figure is not in and of itself *sufficient* to ensure every student realizes this opportunity. Only through the development of a mechanism capable of equitably allocating adequate levels of funding can true educational adequacy (i.e., providing the opportunity for all children to reach a desired level of outcomes irrespective of their circumstance or need) be achieved.

Second, we must realize that the concept of equity (upon which adequacy is determined) has evolved over time. Traditionally, the determination of adequacy was defined by the inputs provided to students with different needs and circumstances (Baker & Levin, 2014). From this input perspective, maintaining horizontal equity requires similar students to be treated in similar ways, while vertical equity requires students with differential needs to be treated in systematically different ways (Berne & Stiefel, 1984). The more recently adopted perspective is focused on equity of outcomes, where the goal is to provide all students with a similar opportunity to achieve some set of desired standards results.

Costing-Out Approaches

There have been great strides made over the past 20-plus years to better measure the cost of providing an adequate education (Rebell, 2006). Specifically, since the mid-1990s, numerous state legislatures, boards of education and advocacy groups have sought to derive empirical estimates of the “cost” of meeting specific state legislative and constitutional standards, including how those costs vary from one location to the next, and one child to the next (Baker, Taylor & Vedlitz, 2008).²

There have been four basic approaches traditionally applied to costing-out studies: Cost Functions, Professional Judgment, Successful Schools, and Evidence-Based. Despite there being four distinct methods, these can be conveniently classified into the following two categories:

- **Input-Oriented (Evidence-Based and Professional Judgment)** – Input-oriented analyses identify the various inputs – human resources/staffing, materials, supplies, equipment, and physical space – required to provide specific educational programs and services. Those programs and services may be identified as typically yielding desired educational outcomes for all student populations when applied in various settings.
- **Outcome-Oriented (Cost Functions and Successful Schools)** – Outcome-oriented analyses start with measured student outcomes, of institutions or specific programs and services. Outcome-oriented analyses can then explore either the aggregate spending on those programs and services yielding specific outcomes, or explore in greater depth the allocation of spending on specific inputs.

¹ For a discussion of the link between adequacy and equity in school finance, see the works by Chambers and Parrish (1982 and 1984) in Illinois and Alaska, which are amongst the earliest costing-out studies. The introductory chapters of these studies specifically address this link between adequacy and equity.

² While efforts to link such cost estimates to constitutional, statutory and regulatory standards were popularized in the era following the well-known education funding court case *Rose v. Council for Better Education*, empirical methods for estimating education costs, including costs of specific standards long pre-date this era.

The primary methodological distinction is whether one starts from an input perspective or with specific outcome measures. One approach works forward, toward actual or desired outcomes, starting with inputs, and the other works backwards from outcomes achieved. Ideally, both work in cyclical feedback with one another. Regardless, any measure of “cost” must consider the outcomes to be achieved through any given level of expenditure and resource allocation.

The following briefly describes each technique.

Cost Functions

The Cost Function (CF) approach uses statistical methods to estimate the relationship between educational costs, educational outcome(s), the price level of schooling inputs, and various measures of pupil need and scale of school or district operations. The approach has been credited for its use of real data on inputs, student needs, price levels, and outcomes to model educational production. The approach also offers a straightforward manner to derive the additional (marginal) costs of achieving education outcomes associated with cost factors such as specific pupil needs (i.e., poverty, special education, etc.), scale of district operations and other contextual factors (student density), as well as labor market conditions affecting the cost of attracting and retaining staff.

Specifically, a comprehensive education cost function model considers spending as a function of a) measured outcomes, b) student population characteristics, c) setting characteristics (economies of scale, population sparsity), d) regional variation in input prices including competitive wages, and e) factors affecting spending that are not associated with outcomes (“efficiency” per se):

$$(1) \quad \text{Spending} = f(\text{Outcomes, Students, Context, Input Prices, Inefficiency})$$

Cost functions can be useful for exploring how otherwise similar schools or districts achieve different outcomes with the same level of spending, or the same outcomes with different levels of spending. That is, differences between districts in terms of their relative efficiency. While the approach can be used to identify the relative (in)efficiency of educational spending, researchers have come to learn that inefficiency found in an education cost function context isn’t exclusively a function of mismanagement and waste, and is often statistically explainable. Inefficient “spending” in a cost function is that portion of spending variation across schools or districts that is not associated with variation in the *observed* outcomes included in the model. That is, inefficiency might be that additional \$1 or \$1,000 spent that didn’t seem to affect the test scores included in the model. But that doesn’t mean it was “wasted.” It might, for example, have been spent to expand the school’s music or robotics program, which may be desirable to local constituents.

Factors that contribute to this type of measured “inefficiency” are also increasingly well understood. For one, local public school districts with greater fiscal capacity – greater ability to raise and spend more – are more likely to do so, and may spend more in ways that do not directly affect measured student outcomes. But that’s not to suggest that all additional spending is frivolous, especially where outcome measurement is limited to basic reading and math skills.

Common criticisms of the approach are that it relies on a limited set of outcome measures,³ the projections can be based on combinations of outcomes and student demographics that are outside of the sample from which the model was estimated, there is little to no transparency as to how resources are combined to generate educational outcomes (i.e., the model is “black box” relating inputs and outcomes to costs), and the technique is generally difficult to explain to non-researchers such as legislators and policy-makers (Chambers & Levin, 2006).

Professional Judgment

Professional Judgment (PJ) involves organizing panels of experienced expert educators to develop efficient resource specifications necessary to deliver a set of desired results or outcomes for students in a variety of hypothetical school settings, the cost of which may be affected by a host of characteristics (cost factors) associated with grade level, student needs (e.g., poverty, English learner and special education status, etc.), and contexts (e.g., enrollment size, urbanicity, etc.). The resource specifications are recorded into what is known as a Resource Cost Model (RCM), which explicitly organizes the resource data according to the specific activities and functions used to provide educational services to students. The RCM has its roots in the “ingredients” approach to cost analysis (Levin, 1983, 2017; and Levin & McEwan, 2001), which represents the gold standard in calculating educational costs through its modeling the structure and “ingredients” of services as they are actually or intended to be provided.⁴

The research team then uses the PJ resource specifications and RCM to calculate the costs of achieving the desired outcomes and to explore the patterns of variation associated with the various cost factors. Based on these patterns of variation, one can calculate the additional costs associated with the various cost factors. PJ has served as the central approach in many costing-out studies including one of the studies reviewed here and multiple studies conducted by the author of this report (Chambers et al., 2004a,b; Chambers, Levin & Delancey, 2007, and Chambers et al. 2008a,b).

Similar to CF and other approaches, PJ can also involve projecting costs beyond the existing sample of schools primarily because there are often few schools serving high need populations that are achieving at the standards used in these studies to define an adequate education (described in a goals statement that usually lists academic and sometimes other student outcomes the programs developed through the PJ process are intended to produce at a minimum cost). However, in contrast to CF, PJ offers much flexibility in terms of the breadth of outcomes that can be taken into account to define the adequacy objective, which may include a myriad of cognitive and non-cognitive dimensions.⁵ In addition, because

³ Virtually all studies using CF define educational adequacy based on average achievement scores or proficiency rates on one or a few standardized tests.

⁴ The approach is a systematic, well-tested procedure for identifying the comprehensive costs of implementing educational services and its use has not been limited to just costing-out studies such as those reviewed here. For example, it has also been used in recent studies for the U.S. Department of Education Institute of Educational Sciences investigating the cost-effectiveness of various interventions to promote high school completion, early literacy, and adolescent literacy, respectively (Levin et al., 2014, Hollands et al., 2013, and Somers et al., 2010).

⁵ Note that the educational goals statement used to define an adequate education in the New Mexico study conducted by Chambers et al. (2008a,b) included both cognitive (i.e., knowledge of content standards) and non-cognitive (i.e., development of personal qualities such as personal responsibility, civic participation, work ethic, etc.) elements. Given that research by Nobel laureate James Heckman and others suggests that, compared to cognitive skills, those of a non-cognitive nature (i.e., social skills, motivation, dependability, etc.) continue to develop over a much longer period of time and also generate large payoffs in the labor market (Heckman, 2008), it seems especially important that non-cognitive outcomes also be considered as educational goals in costing-out

PJ takes a bottom-up approach to costing out the resources, the process is very transparent to policy-makers and generally easy to explain.

The most common criticism of the PJ approach is that, while it relies on the practical experience of panels of educators who are closest to students and arguably the most knowledgeable about how to most effectively deliver educational services, the panels may not always specify the most efficient (minimally costly) combinations of resources necessary to achieve the desired student outcomes (Hanushek, 2006). In addition, because the PJ approach generates resource specifications and corresponding costs associated with hypothetical schools, as opposed to the CF approach which relies on data that directly relates resources to outcomes, the results are extremely difficult to validate empirically (i.e., one would have to implement the resource allocations). Later in this report, we detail research design components that have been used in costing-out investigations to address this concern (Chambers et al., 2004a,b; Chambers, Levin & Delancey, 2007, and Chambers et al., 2008a,b).

Successful Schools

Successful School – Traditional

The third method that has been commonly used to cost out educational adequacy is the Successful Schools (SS) approach introduced by Augenblick and colleagues (1993).⁶ The traditional SS approach attempts to identify the costs of adequacy by determining the average spending among districts that have been identified as successful in terms of academic achievement. While SS shares the transparency of the input-oriented professional judgment approach, like the output-oriented CF approach it relies on empirical observation to determine the costs of an adequate education. In addition to being simple to explain, depending on data availability the SS approach allows researchers to further investigate the types and quantities of resources being used at those schools/districts identified as successful and whether their organization of resources differ from schools that are not deemed successful.

On the surface, the SS methodology seems to be a logical costing-out approach to quantifying the cost of providing an adequate education. However, as it has been traditionally applied, it has a fatal fundamental flaw: specifically, it does not account for factors related to student needs or resource usage. Specifically, the successful districts identified may be those serving the most affluent student populations with lower needs and that operate in locales that are less costly (e.g., suburban areas) than their less successful counterparts. In turn, it can be argued that the approach provides little guidance in determining how much an adequate education would cost across the state, including for pupils in districts that are dissimilar to those deemed successful. Referring to the equation (1) used above to describe the CF approach, the application of SS can be thought of as a cost function that controls for nothing but outcomes as shown in equation (2):

$$(2) \quad \text{Spending} = f(\text{Outcomes}, \text{Students}, \text{Context}, \text{Input Prices}, \text{Inefficiency})$$

That is, the method is little more than a cost function a) without any controls for student characteristics, context or input price variation and b) without any, or with wholly insufficient controls for inefficiency.⁷

studies.

⁶ As many of these studies were performed at the district level, this might also be referred to as the Successful School *District* approach.

⁷ Notably, one could take average spending of schools or districts in various poverty categories, of various sizes, in various labor markets, etc. and also look within fiscal capacity ranges (to address indirect inefficiency predictors).

To this end, the SS approach as it has traditionally been applied has been discounted altogether as a rational costing-out approach (Baker & Levin, 2014).

Often the case is made that the SS approach is in fact appropriate to calculate a *base per-pupil cost* or the cost of providing an adequate education to students with no additional needs, however, this argument is easily dismissed as it suffers from the same issue mentioned above. That is, even the cost of providing an adequate education to students without additional needs (i.e., those who are identified as at risk, English learners or in need of special education services) may differ significantly across districts that face different levels of student needs or contextual challenges related to other cost factors such as scale of operations (size of enrollment), student density, or labor market conditions that make hiring and retaining staff more or less costly.

Successful School – Beating-the-Odds

As an alternative to SS, the Beating-the-Odds (BTO) approach takes a more sophisticated approach to identify successful schools. BTO uses statistical techniques to identify schools that are doing better than expected (“beating the odds”, if you will) given the needs of the students they serve and other contextual factors thought to affect educational costs.⁸ One can then collect data on relatively high-performing (beating-the-odds) schools to ascertain whether there are differences from relatively low-performing schools (i.e., those not beating-the-odds) in the types and quantities of resources used and how much is being spent. While the BTO methodology seems to provide a more defensible way to identify and cost out high performing schools, the typical application of this method also suffers from the common reliance on the limited set of outcomes that are at hand (average test scores or proficiency rates).

Moreover, it is important to understand that the BTO model as generally applied does not provide any definitive identification of schools that are operating *efficiently*. This is because the model only describes the relationship between a limited number of student outcomes (e.g., achievement in math and English language arts) and factors related to student needs and other contextual factors (scale of operation), but does not include direct measures of inputs or costs. A related method constitutes the first traditional costing-out approach presented above, cost functions, which account for cost factors (student needs), student outcomes and educational costs in the same model. Finally, while it may be tempting to identify individual schools that are deemed to be beating the odds and argue that all schools that are observationally identical should be able to operate in a similar fashion and necessarily achieve the same level of outcomes, this would be erroneous. The results only suggest that, *on average*, schools that are observationally similar to a given BTO school are expected to exhibit the same level of outcome. While on average schools that are observationally identical to a given BTO school will perform the same, there will be a spread of these schools that will perform better or worse than this average expectation.

But, by the time all of these cuts have been made, one has basically converged on estimating an actual cost function, but still missing critical components.

⁸ BTO analysis draws on what are referred to as adjusted performance measures in order to identify schools/districts that are considered extraordinarily successful given their characteristics. Examples of BTO analysis can be found in the studies by include Klitgaard and Hall (1972), Stiefel et al. (1999), and Perez et al. (2007).

Evidence-Based

The Evidence-Based (EB) approach was introduced by Odden et al. (2003a,b and 2006). This model draws upon the calculated costs of resource allocations found in literature on effective schooling practices as the foundation to estimate the cost of achieving adequacy in school funding. The notion of using the best available evidence on educational effectiveness has both intuitive and practical appeal. It is extremely transparent in terms of the types and quantities of resources used as the basis of costing out an adequate education. Moreover, the approach is quite simple to explain and is fairly easy to understand for policy-makers and stakeholders.

While there is much to be said for the concept of an EB approach to cost estimation, the manner in which this method has been implemented makes it rather suspect. The way in which EB uses the results of existing educational research has been highlighted as incorrect in terms of its summing the expected educational gains suggested from the various study interventions and their connection to the corresponding intervention resources and subsequent costs. The method is not only sensitive to the selection of literature chosen and the expected impact of implementing the combination of suggested resources (which come from widely different independent studies) on outcomes is unclear at best.⁹ However, this is not to say that the education literature upon which the EB approach depends is flawed in any way, only that the manner in which the EB approach has traditionally applied the results of the research to costing out an adequate education is deficient.¹⁰ Also, as noted by Taylor et al. (2005), users of this approach are limited to the outcomes contained in the effectiveness literature upon which the costing-out specifications are based, which may be quite different from those that are of direct interest to the client. Finally, the approach does not easily lend itself to measuring the additional (marginal) costs associated with providing adequate educational opportunity across students with diverse needs (i.e., poverty, English learner, special education, etc.) and hence offers little insight into how resources should be distributed to this end.¹¹

Summing Up the Different Approaches

Table 1 summarizes existing perspectives on education cost analysis as applied to measuring educational adequacy, organizing the methods into *input-oriented* and *outcome-oriented* methods, which are subsequently applied to hypothetical or actual spending and outcomes. The third column addresses the method by which information is commonly gathered, such as focus groups, or consultant synthesis of literature. The fourth column adds another dimension – the unit of analysis, which also includes the issue of *sampling density*. Most focus group activities can only practically address the needs of a handful of prototypical schools and student populations, whereas cost modeling, or even PJ applied to all actual schools and their data, involves all schools and districts, potentially over multiple years (to capture time dynamics of the system in addition to cross sectional variation).

All methods have strengths and weaknesses, but some weaknesses are critical flaws. Successful Schools is excluded from this table because it is not deemed a credible method of cost analysis. One might argue

⁹ Hanushek (2007) provides a critique of a recent adequacy study that makes use of the Evidence-Based approach, which emphasizes the unrealistic expected achievement gains implied by the study.

¹⁰ Indeed, the hybrid approach used in the comprehensive costing-out model described below explicitly provides expert briefs that draw upon the education research literature to provide information on the elements of successful schools to professional judgement panelists.

¹¹ That is, the Evidence-Based approach does little to formally address Question 2 put forth above.

similarly that a pure “evidence-based” approach, not integrated with context specific judgments is also moot, since it makes no attempt to estimate the costs of the state’s own outcome goals and further, because it fails to consider how needs vary across settings and children in the state specific context. The greatest shortcoming of a more robustly implemented PJ process is the tenuous, hypothetical link to outcomes. The greatest weakness of cost modeling is perhaps the quality and breadth of commonly available outcome measures and the potential influence of those quality and breadth concerns on model predictions.

Table 1 – Summary of Cost Analysis Methods in Education

General Method	Outcome/ Goal Basis	Information Gathering	Unit of Application	Strengths	Weaknesses
Input-Oriented [Professional Judgment and Evidence-Based]	Hypothetical	Focus Groups (Professional Judgment)	Prototypes (limited set)	Stakeholder involvement. Context sensitive.	Only hypothetical connection to outcomes. Addresses only limited conditions/settings.
	Hypothetical	Consultant Synthesis (Evidence Based)	Single model (transposed across settings)	Limited effort. Ability to use and apply boilerplate to any situation. Built on empirically validated strategies.	Aggregation of “strategies” to whole school is suspect. Transferability of “strategies” limited. Not context sensitive.
Outcome- Oriented [Cost Function]	Actual		All districts/schools over multiple years.	Base on statistical link between actual outcomes and actual spending. Evaluates distribution across all districts/schools.	Requires rich personnel, fiscal and outcome data. Potentially infeasible where outcome goal far exceeds any reality. Focus on limited measured outcomes. Limited insights into internal resource use/allocation underlying cost estimate.

Source: Baker & Levin (2014).

3 – Review of Calculation of the Cost of a Suitable Education in Kansas in 2000-2001 Using Two Different Analytic Approaches (Augenblick, Myers, Silverstein & Barkis, 2002)

Study Methodology

The 2002 study by Augenblick et al., makes use of two different costing-out methods, the input-based PJ approach and the outcome-based SS approach. We describe each of these briefly in turn.

Professional Judgment Approach (Input-Oriented Approach)

The first methodology used by the study is the PJ approach. There were four main tasks involved:

- 1) Defining a Suitable Education – This was done in consultation with the Legislative Education Planning Committee (LEPC) with the final definition including both input and outcome standards. The input standards were based upon the offered course, program and services included in the Kansas Quality Performance Act (QPA), while the performance standards were defined by districts that within a five-year period would meet specific percentage threshold standards of students scoring proficient or better (aka percent-above-cut-score) on six different grade level/subject specific criterion-referenced tests used for accountability purposes as shown in Table 2:¹²

Table 2 – Student Outcomes Used for Suitability Definition

Grade	Percent of Students Scoring Proficient or Higher	
	Math	Reading
4	65%	N/A
5	N/A	70%
7	60%	N/A
8	N/A	65%
10	55%	N/A
11	N/A	60%

Developing District and School Prototypes – The authors first developed 4 categories of districts that were distinguished by enrollment size. This was done by rank ordering the 304 districts in the state by enrollment and determining both raw district and pupil-weighted district quartiles, where the raw quartiles split the population into four groups with equal numbers of districts (76), while the pupil weighted split them into four groups with (roughly) equal enrollments (Table 3a).

Table 3b shows the final grouping used for the prototypes. Note, this grouping scheme made use of combinations of both quartile calculation schemes. Specifically, the raw quartile groups 1 and 2 for the Very Small and Small district categories, respectively, a combination of unweighted quartile 3 along with a portion of weighted quartile 1 and all of weighted quartile 2 for the

¹² Appendix B of the A&M study includes the formal definition of a suitable education used for the PJ approach.

Moderate district category, and all of weighted quartiles 3 and 4 for the Large district category. The authors provide no justification for the final designation of the district size categories.

Table 3a – Raw and Pupil-Weighted Quartiles of Enrollment Used to Define District Size Categories

	<i>District Size Quartiles</i>			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Raw Quartiles – Number of Districts (Enrollment Range)	76 (≤324)	76 (325-555)	76 (556-1,139)	76 (≥1,140)
Pupil-Weighted Quartiles – Number of Districts (Enrollment Range)	230 (≤1,140)	54 (1,150-3,599)	16 (3,600-16,499)	4 (≥16,500)

Table 3b – Final District Size Categories Used

	<i>District Size Category</i>			
	Very Small	Small	Moderate	Large
Enrollment Range	≤324	325-555	556-3,600	≥3,601

Table 4 – Final District and School Prototypes Used for Professional Judgment Panels

	<i>District Size Category</i>			
	Very Small	Small	Moderate	Large
Range in Enrollment	≤324	325-555	556-3,600	≥3,601
Average District Enrollment	200	430	1,300	11,200
<u>Average School Enrollment</u>				
Elementary	140	150	200	430
Middle	-	-	300	430
High School	60	130	400	1,150
<u>Average Numbers of Schools</u>				
Elementary	1	2	3	12
Middle	-	-	1	6
High School	1	1	1	3
<u>Average Incidences of Student Needs</u>				
Proportion of Students in Special Education	14%	14%	13%	14%
Proportion of Students Eligible for Free/Reduced Price Lunch	35%	35%	29%	36%
Proportion of Bilingual Students	2%	2%	3%	4%
<i>Note: Table adapted from study pages IV-2 and IV-3.</i>				

Within each district size category, the averages of district total enrollment, the numbers and enrollments of schools at the elementary, middle and high school levels, and incidences of students in special education, eligible for free/reduced price lunch, and identified as bilingual were calculated. Table 4 provides the final prototype definitions of districts and schools used in

the professional judgment panel work. It is important to note that the authors did not develop middle school prototypes for the Very Small and Small district size categories, as they claim that there were no stand-alone middle schools in these types of districts.

- 2) Selection of Panelists, Convening of Panels and Public Engagement – The authors consulted with the LEPC and the Kansas State Department of Education (KSDE) to select 25 individuals that made up four school-site professional judgment panels. One school-site panel was assigned to the Very Small and Small district size school prototypes, another school-site panel was assigned to the Large district size school prototypes, and two school-site panels were assigned to complete duplicate sets of the Moderate district school prototypes. A group of 15 panelists were chosen in a similar manner to serve on two district professional judgment panels charged with reviewing the work of the school-site panels and an expert panel of 6 panelists was chosen to review the work of the district professional judgment panel. The school-site panels convened for 1.5 days (December 4-5, 2001), during which time they deliberated and specified resources for the school prototypes. The district panels convened for 1.5 days (January 8-9, 2002) to review and amend the school prototype resources, as well as specify district-level resources to be added to the school-level prototypes. Finally, the expert panel met for 1 day (March 13, 2002) and made modifications to one of the two sets of prototypes for the schools and district under the Moderate size category.

The authors also conducted both a questionnaire and interviews lasting up to four hours with 10-person groups drawn from a pool of 59 participants included in a KSDE-provided list of 97 individuals that was made up of educators, school board members, education advisory group member, parents, and business community members. This engagement effort was done to get a better sense of public views on the Kansas school finance system concerning the funding foundation level, the current weights used to adjust funding for student needs (at-risk, bilingual and special education), scale of operations (district size), and programs such as vocational education. In addition, the data collection solicited input from respondents/participants on issues such as the appropriate provision of staff professional development. The meetings took place on November 13 and December 4, 2001, and on January 8, 2002.

- 3) Assigning Resource Prices, Calculating Costs and Developing Weights – The final step involved assigning unit prices for each type of resource and calculating the costs associated with each school prototype. Next, they added the corresponding costs of district-level resources, reported aggregate costs across the district size categories broken out by base spending versus additional spending necessary to support students with special needs, and determined base per-pupil funding and empirical weights for special education, at-risk, and bilingual students for each district size category prototype. The authors then used the information across the district size categories to generate schedules of base per-pupil funding and student need weights that varied with district enrollment size.

Successful Schools Approach (Outcome-Oriented Approach)

Implementation of the SS approach was far less involved than the PJ approach. The authors first determined districts that were successful in terms of their student outcomes. This was done by analyzing each district's percentage of students with scores that were proficient on the state's math and reading tests used for accountability purposes. To be deemed successful, a district had to be either meeting the percent thresholds mentioned earlier on five out of the six grade/subject specific tests or

be considered on track to meet these thresholds within five years. The determination of whether a district was considered being on track was made by looking at the changes in the percentage of students with proficient scores on each test from the 2000 to 2001 and comparing these year-over-year changes to the yearly progress that would have to be made to reach the test-specific thresholds within five years. According to this criterion, 86 of the statewide total of 304 districts were deemed successful in terms of their outcomes.

The authors next identified districts in terms of their compliance with the School District Finance and Quality Performance Act standards (QPA), which involved providing appropriate courses, programming and services. Only 1 of the 86 districts deemed successful according to the outcome criterion was found not to be meeting the QPA standards, leaving the final number of successful districts at 85.

Next, the authors isolated the basic expenditures of the districts, by excluding spending on services for special education, at-risk, and bilingual student populations, as well as expenditures on capital, food service, and transportation. Using these total spending figures, the authors calculated a pupil-weighted average base cost per pupil across the 85 districts.

Key Results and Discussion

Key Results

The key results from the PJ approach pertaining to suitable base and special needs per-pupil costs and corresponding weights are listed in Table 5. The base per-pupil cost resulting from the PJ approach ranged from \$5,811 for Large districts to \$8,581 for Very Small districts, with a pupil-weighted average across districts of \$6,362. This is about 40 percent larger than the pupil-weighted average base per-pupil cost calculated using the successful schools approach.

Additional special education per-pupil costs range from \$6,908 (Small) to \$12,090 (Large) with a pupil-weighted average of \$9,848 and corresponding special education weights ranging from 0.86 to 2.08. That is, the additional funding above and beyond the base cost that is necessary to support the cost of a special education student was between \$6,908 and \$12,090 across the district size categories or 0.86 to 2.08 times the base per-pupil cost for each of these categories. The at-risk per-pupil costs range from \$1,919 (Very Small) to \$3,392 (Moderate) with a pupil-weighted average of \$2,846 and corresponding weights ranging from 0.22 to 0.44. Bilingual per-pupil costs range from \$1,217 (Very Small) to \$5,993 (Large) with a pupil-weighted average of \$5,320 and corresponding weights equal to 0.14 and 1.03. Taking a ratio of the pupil-weighted average of the additional cost associated with each student need allows calculations of the weights associated with the pupil-weighted average costs are as follows: special education-1.55, at-risk-0.45, and bilingual 0.84.

The main result from the SS approach was a base per-pupil cost calculated at \$4,547. The SS per-pupil base figure (lower than the lowest PJ per-pupil base of \$5,811 generated for the Large district prototype) was combined with the weight figures generated using the PJ approach to develop cost schedules across the full district enrollment range. The cost schedules were then used to project the district-level and bottom-line adequacy costs, the latter of which was compared to current spending at the time. Using a current spending figure on comparable purposes (general school operations, which excludes capital, transportation, etc.) of \$2.837 billion, the authors conclude that total spending would need to increase by about \$236 million to \$3.073 billion (equal to a relative increase of 8.3 percent).

Table 5 – Suitable Base and Special Needs Per-Pupil Costs and Corresponding Weights from PJ Approach

	<u>District Size Category</u>				<u>Pupil-Weighted Averages</u>
	<u>Very Small</u>	<u>Small</u>	<u>Moderate</u>	<u>Large</u>	
<u>Total Base Cost from PJ Approach</u>					
<i>School Level</i>	\$6,692	\$5,786	\$5,499	\$4,724	
<i>District Level</i>	\$1,889	\$1,575	\$1,184	\$1,087	
<i>Total PJ Base Cost</i>	\$8,581	\$7,361	\$6,683	\$5,811	
<i>Pupil-Weighted Average Base from PJ</i>			\$6,362		
<i>Pupil-Weighted Average Base from Successful Schools</i>			\$4,547		
<i>Relative Difference Between PJ and Successful Schools Bases</i>			39.9%		
<u>Added Costs of Special Needs Students</u>					
	<u>Very Small</u>	<u>Small</u>	<u>Moderate</u>	<u>Large</u>	<u>Pupil-Weighted Averages</u>
<i>Special Education</i>	\$7,403	\$6,908	\$7,731	\$12,090	\$9,848
<i>At-Risk</i>	\$1,919	\$2,228	\$3,392	\$2,578	\$2,846
<i>Bilingual</i>	\$1,217	\$1,267	\$5,590	\$5,993	\$5,320
<u>Special Needs Weight Calculations</u>					
	<u>Very Small</u>	<u>Small</u>	<u>Moderate</u>	<u>Large</u>	<u>Pupil-Weighted Averages</u>
<i>Special Education</i>	0.86	0.94	1.16	2.08	1.55
<i>At-Risk</i>	0.22	0.30	0.51	0.44	0.45
<i>Bilingual</i>	0.14	0.17	0.84	1.03	0.84

Note: Derived from A&M study Table IV-10. Pupil-weighted averages of added costs of special needs students added by review author. 2000-01 statewide enrollments across size categories used to calculate pupil-weighted averages are as follows: Very Small (15,788), Small (32,872), Moderate (173,808) and Large (224,502). Pupil-weighted averages of special needs weight calculations based on ratios of pupil-weighted average special needs costs to pupil-weighted average PJ base per-pupil cost (e.g., pupil-weighted average special education weight of 1.55 equals \$9,848 / \$6,362).

They next offset estimated local and Federal revenues to calculate what the burden of the increase would be to the state, yielding a figure of \$284 million or 13.4 percent.

Discussion

My general impression of the A&M study is that it is a rather early effort implementing a PJ approach to costing-out educational suitability that includes some flaws in its design and implementation. In addition, I had some issues with how the study findings were translated into actionable funding policy. The following includes a critical discussion of the A&M study methodology and implementation focusing on the PJ approach and including how results may have been shaped by the data used and analytical choices made by the authors. As the study includes a rather dated implementation of the PJ approach, the text points out advancements used in more recent applications of the approach. The choice to focus on the PJ approach stems from a general lack of credibility in the SS approach as a valid costing-out methodology (Baker & Levin, 2014) and the larger share of the study findings that are made up of the PJ results (i.e., the SS approach was only used to calculate base per-pupil cost, while the PJ approach generated both base per-pupil cost and weight estimates).

Development of School Prototypes

A simple review of the district and school prototypes brings forth a major concern that almost certainly had significant influence on the key results presented above. Specifically, the review uncovered two issues that could not be ignored, but the effects of which are not clear.

First, it seems that the incidence of student needs used to define the district and school prototypes do not seem to be correct. Specifically, there is evidence that the average rates of students eligible for free or reduced price lunch (FRL) used to define the district and school prototype definitions that the PJ panelists based suitable education models do not comport with those calculated using data downloaded from the KSDE.¹³ The first panel of Table 6 shows the district average percentage of FRL reported in the A&M study (page IV-2) for each district size category, the same figures calculated for the purposes of this review, and the differences in incidence rates between the two sets of figures. While the differences for the Moderate and Large districts is quite small, we find that the FRL rates used in the study for Small districts was somewhat larger (by 2.4 percentage points) than the rate calculated for this review. Conversely, the average FRL rate used in the study for the Very Small district prototype was 4.4 percentage points smaller than what was calculated using KSDE data. To this end, it seems that in developing their models the panels were reacting to a key student need characteristic that was slightly too high for Small districts and too low for Very Small districts.

A second more fundamental problem that precipitated the investigation in this section is the fact that the authors used district averages to define student needs in both the district- and school-level prototypes. Ideally, the set of school prototypes used in the PJ approach should attempt to approximate the ranges of student need and school size naturally occurring in a state. It is this variation that will drive a more accurate calculation of how much more it costs to provide a suitable education to students with different types of needs and attending schools of different sizes. Because of this critical research design decision, the school prototypes are unfortunately quite limited in their ability to reflect the

¹³ School-level data on counts of students approved for free/reduced price lunch in Kansas for the 2000-01 school year were downloaded from the report generator on the KSDE website here: (http://datacentral.ksde.org/report_gen.aspx). These data were used to generate both district- and school-level pupil-weighted averages for each district category.

variation in pupil needs that actually existed across schools in the state. Specifically, the variation in student needs across the school prototypes used in the study only represents that found across the average districts within the four broad categories of district size. As seen in the prototype definitions listed in Table 4, above, while school size seems to follow district size, there is almost no variation in any of the average student needs incidences across the four district size categories. What is lamentable is the fact that the authors could have simply calculated school-level averages of the student needs variables across schools within each district size and by schooling level, which would have provided a more credible representation of needs across the state.¹⁴ Performing averages by schooling level is particularly important, given the well-known phenomenon whereby reported rates of students eligible for free/reduced price lunch for high schools are systematically lower than for their elementary and middle school counterparts.

Table 6 – Average District and School Incidences of Students Eligible for Free/Reduced Price Lunch Used in A&M Study and Calculated from KSDE Data

	<i>District Size Category</i>			
	<i>Very Small</i>	<i>Small</i>	<i>Moderate</i>	<i>Large</i>
<i>Averages Used in Study and Calculated from KSDE Data</i>				
<i>District Averages Used in Study for Both Districts and Schools</i>	35.0%	35.0%	29.0%	36.0%
<i>District Averages Calculated from KSDE Data</i>	39.4%	32.6%	28.7%	35.9%
<i>Difference in Study and KSDE Calculated Averages</i>	-4.4%	2.4%	0.3%	0.1%
<i>Schooling-Level Averages Calculated from KSDE Data</i>				
<i>Elementary</i>	44.6%	36.9%	33.7%	43.9%
<i>Middle</i>	40.1%	34.9%	28.8%	40.2%
<i>High</i>	33.6%	26.8%	21.5%	26.6%
<i>Differences Between District Averages Used in Study and Schooling-Level Averages Calculated from KSDE Data</i>				
<i>Elementary</i>	-9.6%	-1.9%	-4.7%	-7.9%
<i>Middle</i>	-5.1%	0.1%	0.2%	-4.2%
<i>High</i>	1.4%	8.2%	7.5%	9.4%

To check the degree to which the free/reduced price lunch rates used in the A&M study for both the district and school prototypes were different from the actual school-level averages that existed in Kansas in the 2000-01 school year the analysis was extended. The second panel in Table 6 shows the average FRL rates across schools at each schooling level within each of the four district size categories. The resulting average FRL rates show a consistent relationship across the district size categories at each schooling level; namely, schools in Very Small and Large districts tend to have the highest rates, while

¹⁴ Indeed, the authors were able to compute school-level averages of school size within each of the district size categories so it is curious that they did not do the same for the student needs characteristics. Perhaps the school-level student needs data were not available at the time.

those in Moderate sized districts tend have the lowest, and those in Small districts are somewhere in between. However, it should also be noted that within each schooling level the variation in average calculated FRL rates across the district size categories is much greater compared to those used in the school prototypes. The results also show a common pattern whereby FRL rates tend to be highest among elementary schools and lowest among high schools, with middle schools in between.

The third panel of the table contains the percentage point differences between the school-level FRL rates calculated from the KSDE data and those used for the school (and district) prototypes used in the PJ approach. The results are quite striking showing that the prototype FRL rates significantly over or underestimated student needs across the schooling levels and district size categories. Specifically, FRL rates at the elementary level were systematically underestimated by the school prototypes by 9.6 percentage points for Very Small districts, 7.9 percentage points for Large Districts, 4.7 percentage points for Moderate size districts, and 1.9 percentage points for Small districts. Conversely, the high school prototypes systematically overestimated the FRL rates for high schools by 1.4 to 9.4 percentage points. At the middle school level, the results are mixed. The school prototypes for Very Small and Large districts underestimated the average FRL rate by 5.1 and 4.2 percentage points, respectively.

Unfortunately, publicly available data was not available on the other student needs characteristics defining the prototypes (incidences of special education and bilingual students) and therefore was not analyzed. However, one might hypothesize that given the significant correlation between the incidences of FRL and bilingual students that is often observed, a similar although less pronounced problem would also exist with the bilingual model components that were specified. Also, while the percentage differences may not seem like a lot, in relative terms they can be quite large. For example, the largest underestimates and overestimates found (for elementary schools in Very Small districts and high schools in Very Large districts) show that the values used for the prototypes were over one-quarter smaller and larger, respectively than they should have been.

In sum, it seems likely that the panelists likely would have specified more resources in the elementary school prototypes and fewer in the high school prototypes. However, looking at the differences between the school-level percent FRL used in the prototypes versus what is found from KSDE data across the three schooling levels for each district size category (i.e., down the columns of the last panel in Table 6), one could legitimately assume that overall the resources specified for Very Small and Large districts were too low, while those specified for Small and Moderate districts were too high. Unfortunately, while it would be hard to believe that this research design flaw could not have influenced the panelists' decisions, it is impossible to fully understand what overall impact this may have had on the final results. My thought here is that the school-level cost generated by the PJ approach is lower overall than it would have been if the school prototypes were defined with demographics that were true to the average needs specific to schooling levels within each district size category.

Translating Findings into Actionable Funding Recommendations

The authors made a good effort to translate the main results of both the PJ and SS approaches into funding recommendations that could be implemented. The first of these was to establish the base (foundation) per-pupil funding amount to which the various calculated weights for at-risk, bilingual, and special education were applied.

Base Per-Pupil Foundation

Exhibit 1 provides three cost schedules that show how suggested per-pupil base funding would be affected by district size. The solid-line schedule in blue represents the costs suggested by the A&M PJ approach (minimum of \$5,800), while the solid-line schedule in orange is that suggested by the SS approach (minimum of \$4,550). The third schedule in red (named “Raw PJ Base Cost” with a minimum of \$5,811) was developed by me directly from the data presented in Table 5, above. There is very little difference between the suggested PJ and raw PJ schedules.^{15, 16}

As can be seen, all three schedules produce the expected story that is consistent with economies of scale. That is, it is often found that the per-unit (per-pupil in this case) cost of production decreases as the scale of production gets larger. All three behave quite similarly, although the SS schedule is significantly lower at each enrollment level. The authors devote a discussion of why these differences might occur, stating that the districts identified for the SS approach might not meet all of the components that constitute a suitable education, which the prototype districts of the PJ approach by definition are assumed to meet. While the study is silent on any examples where this might be the case, one might be the fact that the SS districts were identified as successful if they met or were on track to meet test proficiency thresholds on five of the six tests, while the PJ panels were charged with developing models that would achieve the thresholds on all six tests.

However, the difference in the PJ and SS base per-pupil cost measures are most likely borne out of systematic differences in the characteristics of those districts deemed successful and other districts in the state, which the SS approach does not control for. It is precisely this issue that renders the SS approach useless for determine the costs of a suitable education (Baker & Levin, 2014). To this end, the suggested PJ base is preferable to that generated using the SS approach. Moreover, the scale adjustments seem appropriate. Indeed, the structure of the PJ prototypes were designed based upon differences in enrollment and therefore the approach seems to do a good job at distinguishing the differential costs associated with scale of operations.

An important decision is made by the authors was to use the lower SS base per-pupil cost as the driving the foundation level by which all districts were funded. The PJ base, or a scaled down version of the PJ base, would then be used as the limit on second tier funding (Local Option Budget or LOB).¹⁷ There are at least two things that are problematic with this decision. First, the choice to use the SS base per-pupil figure would seem to be endorsing an unreliable measure that seems to be an underestimate of the true base per-pupil cost (note that even the reported PJ base cost was deemed to be underestimated to some extent and the SS base is far lower than that). Second, using the PJ base per-pupil cost to set the LOB limit makes little sense in that these two things are meant to serve entirely different purposes. Specifically, a per-pupil funding base constitutes what must be spent on a student with no special needs in order to provide them with a suitable education. In contrast, the LOB is a limit of what can be spent

¹⁵ My thought is that the authors fit their suggested schedule to base per-pupil cost numbers that were rounded (e.g., using the minimum of \$5,800 rather than the raw \$5,811 produced by the PJ analysis).

¹⁶ In addition, I have taken the liberty of plotting smooth schedules (the dotted-lines) that do not have points of discontinuity.

¹⁷ The Local Option Budget (LOB) is a second-tier funding source by which districts are allowed to use local revenues to generate dollars above an adequate base of funding (one that would support a suitable education). At the time of the study, the amount of LOB funding a district could use was capped at 25 percent of the base.

above and beyond the base (i.e., intended to allow for districts to spend in excess of what is deemed adequate). In turn, it is unclear at best why you would use a base per-pupil cost figure to determine the LOB limit.

At-Risk Weight

Exhibit 2 includes a plot of the suggested schedule of the funding weight for at-risk students (in blue) and another that simply connects the raw weights calculated from the PJ prototypes for each district size category. In addition, I have included a function that best fits the raw data points. The suggested schedule was generated by the following equation:

$$(3) \quad \text{At-Risk Weight} = 0.60 - [(1,000/\text{Enrollment}) \times 0.08]$$

As is evident from the graphic, the intended poverty weight has a minimum of 0.20 and increases with district size, dramatically so at lower enrollment levels (from 200 to 800), and eventually levels off at 0.60. There are several concerns I have with this suggested weight schedule.

First, the positive relationship between district enrollment and the suggested PJ at-risk weight only partly follows the series produced by the raw PJ weights. The suggested PJ weight schedule is also consistently higher than the raw PJ weight series. The reader will also note that the raw PJ weight for the Large district size category (0.44) was lower than for the Moderate district size category (0.51), which seems illogical given the Moderate size prototypes had the lowest percentage of at-risk students of all the district size categories. Importantly, it may be that the pattern of the observed raw PJ weights are more of an artifact stemming from the organizational structure of the prototypes than the actual values of the at-risk percentages to which the panelists responded. Specifically, it does not seem that the prototypes provided sufficient variation in student needs to allow for accurate calculations of need-based weights. The only appreciable change in the at-risk percentage across the district size categories was for Moderate size districts, which was set at 29 percent and 35 or 36 percent for the other three district larger and smaller size categories.

In addition, the fact that only one panel addressed the prototypes in three of the four size categories (the Moderate district size prototypes were performed independently by two panels) is rather troubling (ideally there would be at least two panels developing models for each of the prototypes).¹⁸ Finally, the reader will note that the calculated at-risk weight for Moderate districts is not logical when taken in the context of those calculated for the other district size categories that had higher prototype FRL rates. For example, the Moderate at-risk weight associated with an FRL rate of 29 percent was 0.51, while the weights for Very Small and Small districts associated with an FRL rate of 35 percent were 0.22 and 0.30, respectively.

Second, I am concerned about the degree to which the suggested PJ at-risk weights increase with enrollment according to the schedule. While there are examples in both the research literature and state funding policy that the *concentration* of poverty has a significant impact on the outcomes of at-risk students,¹⁹ it is difficult to accurately determine how much additional funding might be necessary to provide an equitable suitable educational opportunity between at-risk students learning in

¹⁸ A more in-depth discussion of the importance of using multiple panels to perform the same exercises is included below (see section *Multiple Independent PJPs Performing Duplicate School/District Prototypes*).

¹⁹ See for example Reardon (2011).

environments with relatively higher and lower concentrations of poverty. Indeed, the Kansas costing-out study by the Legislative Post Audit Committee (LPA, 2005) described below provides results using a cost function costing-out approach that also suggests a significant relationship between the cost of providing a suitable education and incidence of student poverty in inner-urban districts.

In terms of an example of state funding policy, California's relatively new school finance system, the Local Control Funding Formula, provides an additional "concentration" grant funding adjustment (weight) in districts where the incidence of disadvantaged (at-risk) students (defined as the percentage of unduplicated counts of at-risk, English learners or foster youth) is above 55 percent. In these districts, funding is increased by 0.50 times the base per-pupil funding for each at-risk student accounted for in the excess incidence above 55 percent. To put the at-risk concentration weight in perspective, there is also an initial "supplemental" at-risk weight used where districts get an additional 0.20 times the base for all students that are deemed at risk. So, in California districts where the at-risk concentration weight is applicable, the effective additional funding for each at-risk student over the 55 percent incidence threshold is over three times as large as that for at-risk students under the threshold (3.5 times as large to be precise).²⁰ Exhibit 3 presents this discontinuous LCFF at-risk weight schedule that takes into account both the supplemental and concentration weights to show how the effective weight changes with increases with the incidence of at-risk students. The schedule shows an at-risk weight of 0.20 up until the incidence of at-risk incidence reaches 55 percent, after which the weight steadily climbs to 0.425. It is important to take notice that the ratio of the weight in the highest to lowest incidence districts is 2.125.

The implications of the A&M suggested at-risk weight schedule would be much more aggressive in terms of the funding equity that would ensue if it were enacted. Looking again at Exhibit 2, the smallest districts would receive additional funding for their at-risk students that would be one-third of that for the largest districts. This implies that it is only a third as costly to equally support the outcomes of at-risk students in the smallest districts than in the largest districts. Also, note that while there are no stark discontinuities or "jumps" in the schedule, the steep portion occurring between 200 and 800 students would provide an incentive for districts to increase their enrollment.

In the context of the A&M findings, to the extent that the concentration of at-risk students is related to district enrollment, there may be a call for some sort of upward graduated adjustment in the at-risk weight as district enrollment increases. However, a check of the looking at both the unweighted and pupil-weighted correlations between incidence of at-risk students and districts enrollment using 2000-01 data, I find that there is a negligible or weak correlation between these two variables.²¹

²⁰ Specifically, for at-risk students above the 55 percent threshold districts receive additional funding on the order of 0.70 of the base (this equals the 0.20 supplemental weight plus the 0.50 concentration at-risk weight), while at-risk students below this threshold only get the 0.20 supplemental weight.

²¹ Using KSDE data for 2000-01, I find that the pupil-weighted correlation between district-level percent at-risk and enrollment is 0.22. These were run within each of the district size categories with a mix of weakly negative and weakly positive correlations.

Exhibit 1 - Alternative Suggested Base Per-Pupil Suitable Costs by District Enrollment

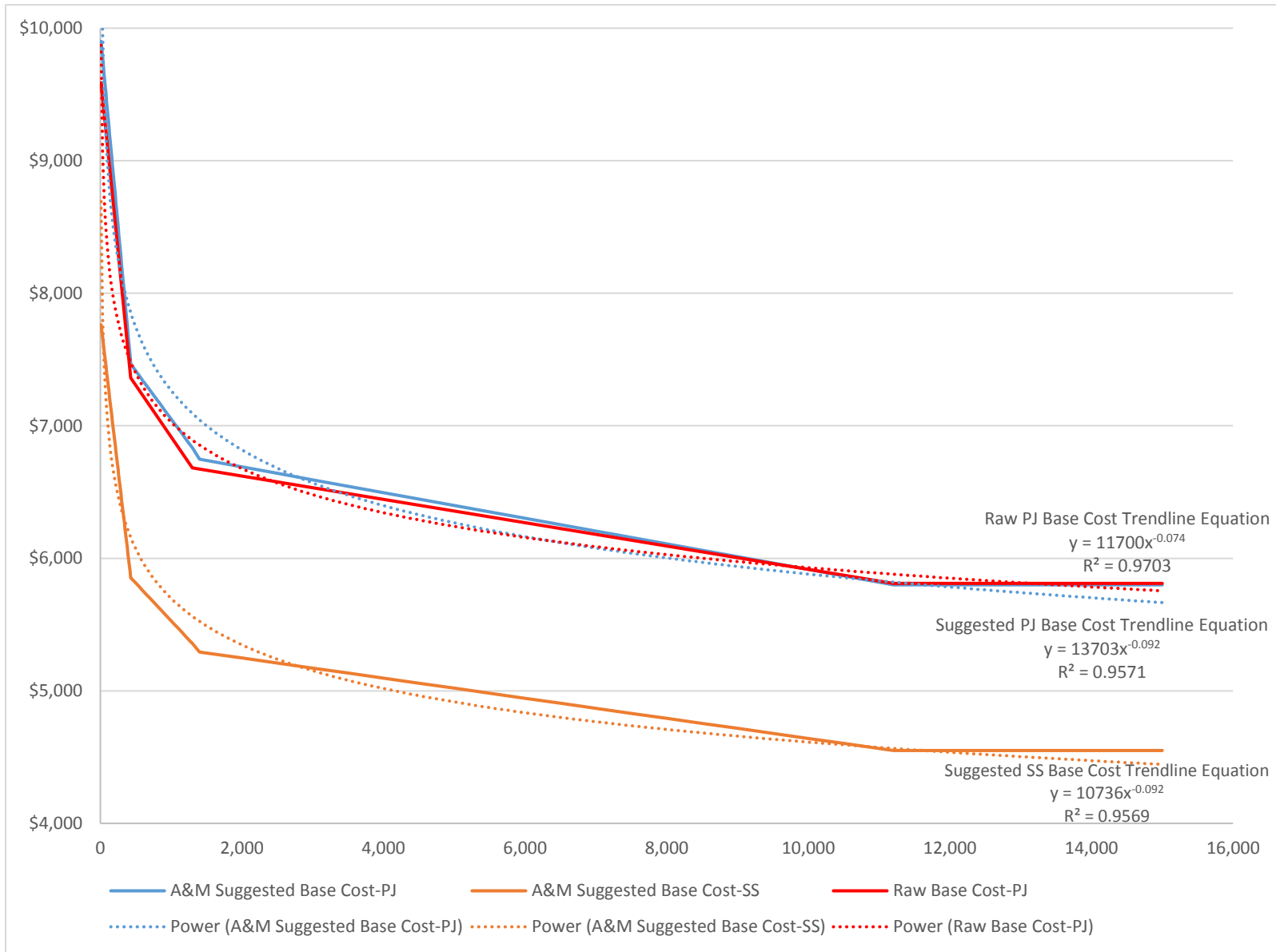


Exhibit 2 - A&M Suggested At-Risk Weights by District Enrollment

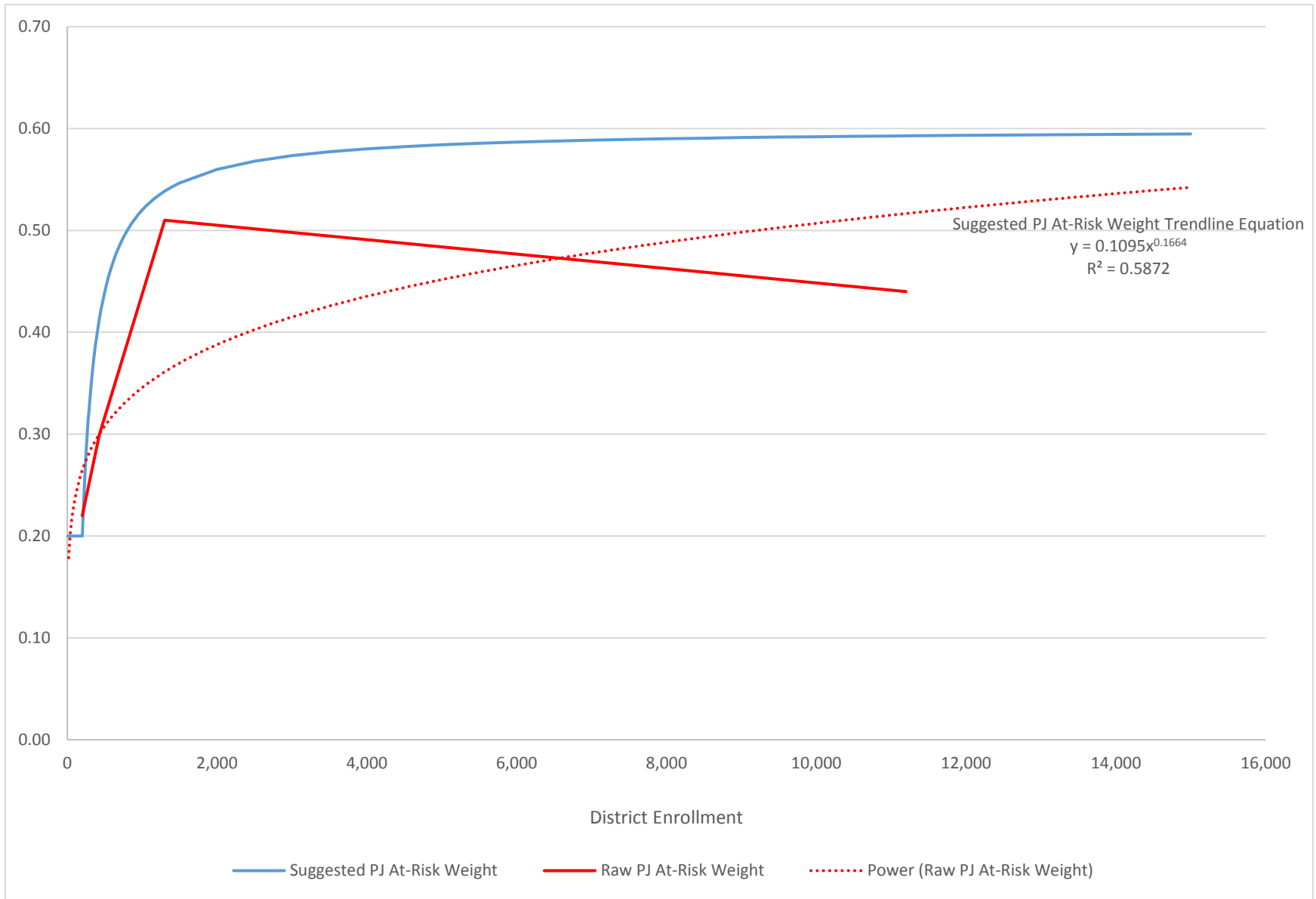
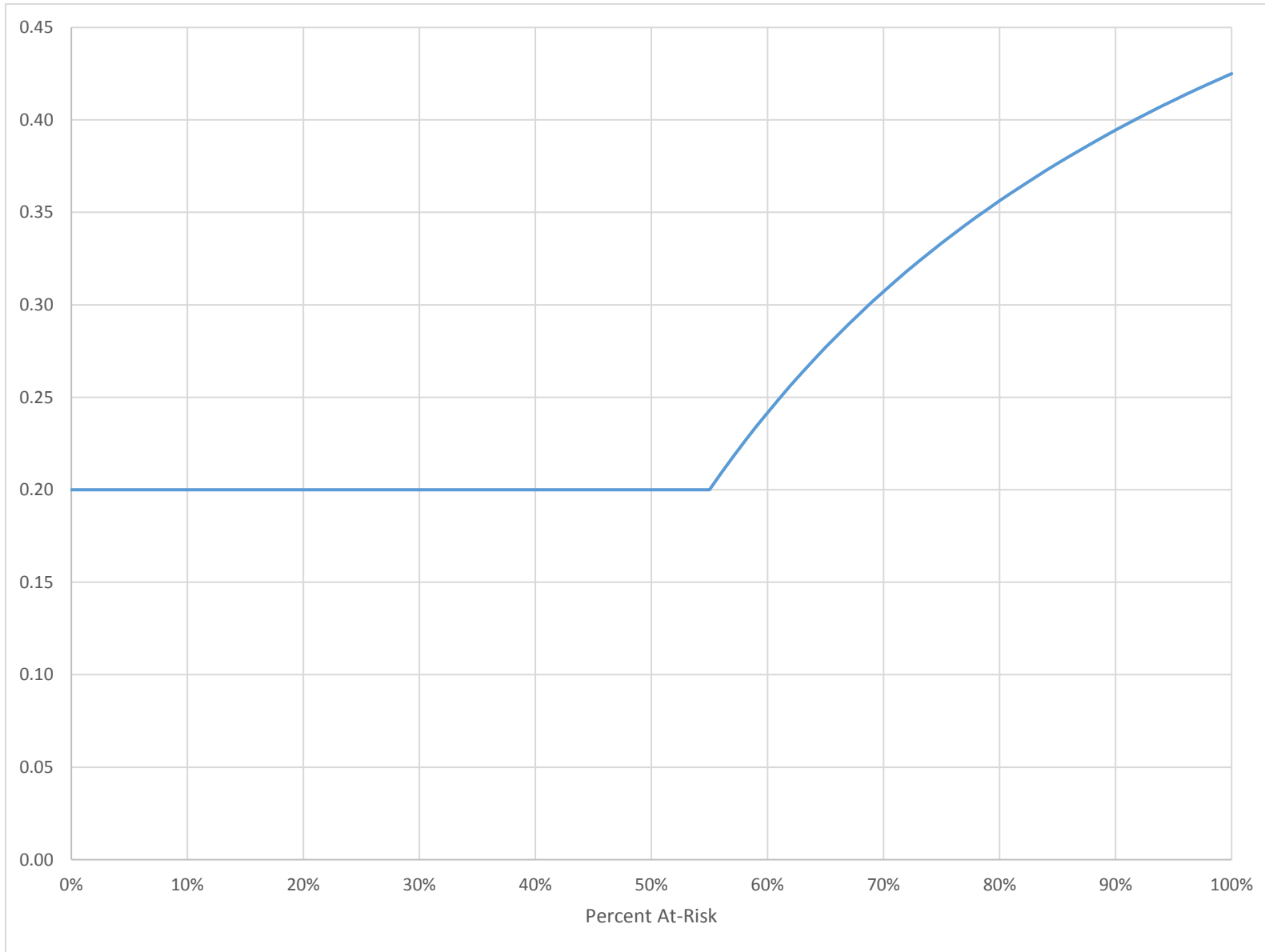


Exhibit 3 – At-Risk Weight Schedule from California Local Control Funding Formula (LCFF)



Given the large relative difference between the suggested PJ at-risk weight in the largest versus smallest districts, perhaps a better solution would be to suggest a standard at-risk weight to be used across all district enrollment sizes. One obvious choice would be to go with the pupil-weighted average of the weights calculated for each district size prototype. My calculations show this would be 0.45, which is admittedly rather conservative compared to other costing-out studies, including the range of at-risk weights computed in the LPA cost function approach.²²

Bilingual Weight

The suggested schedule for the bilingual weight is presented in Exhibit 4. I have similar concerns about the A&M suggested bilingual weight schedule for reasons mentioned above in the discussion of the at-risk weight schedule. The resulting increasing weights across district size are most likely due to the lack of variation in the incidences of bilingual student used across the prototypes specific to schooling levels and district size categories, as well as a lack of multiple panels completing duplicate prototypes. Indeed, similar to the case of the at-risk weights, there may be concentration effects at play (often the incidences of at-risk and bilingual are at least moderately correlated). However, it is difficult to understand why the additional cost of providing a suitable education to a bilingual student would be so much higher in large districts. The equity effects resulting from implementing the suggested bilingual weight schedule would be pronounced, with the relative difference in additional per-pupil funding for bilingual students between the largest and smallest districts measuring over 600 percent. A more logical way to apply the prototype bilingual weights might be to implement their pupil-weighted average equal to 0.84, which is not outside of the range of English learner weights generated by PJ studies (0.39 to 2.0) as reported in the literature review on this very subject by Castellanos-Jimenez and Topper (2012).

Special Education Weight

The authors basically did not make use of the special education weight for the Large district size category because it was considered too high (2.08). Instead, they noted that the other weights were more reasonable (0.86, 0.94 and 1.16 for the Very Small, Small, and Moderate prototypes, respectively), and developed a schedule (Exhibit 5) that starts at a weight of 0.90 for the smallest district sizes and increases with district enrollment as follows:

$$(4) \quad \text{Special Education Weight} = 0.90 + (\text{Enrollment} \times 0.00002)$$

One should notice that the A&M suggested schedule (blue line) is much flatter than the raw schedule (orange line). The 0.90 is a well-established, but outdated, figure calculated in a 2002 report of the Special Education Expenditure Project (Chambers, Parrish & Harr, 2002). However, this is not a weight based on an adequacy cost study, but rather one describing how much was being spent on the average special education student across the county relative to the average student with no special needs *without explicitly taking into account any specific definition of educational suitability*. To this end, the 0.90 weight might be seen as an underestimate of what it would cost to provide a suitable education for the average special education student.

²² See the compiled list of estimated poverty weights from costing-out studies performed from 1997 to 2007 in Baker, Taylor & Vedlitz (2008) which range from 0.58 to 0.92 for those using the PJ approach.

Exhibit 4 – A&M Suggested Bilingual Weights by District Enrollment

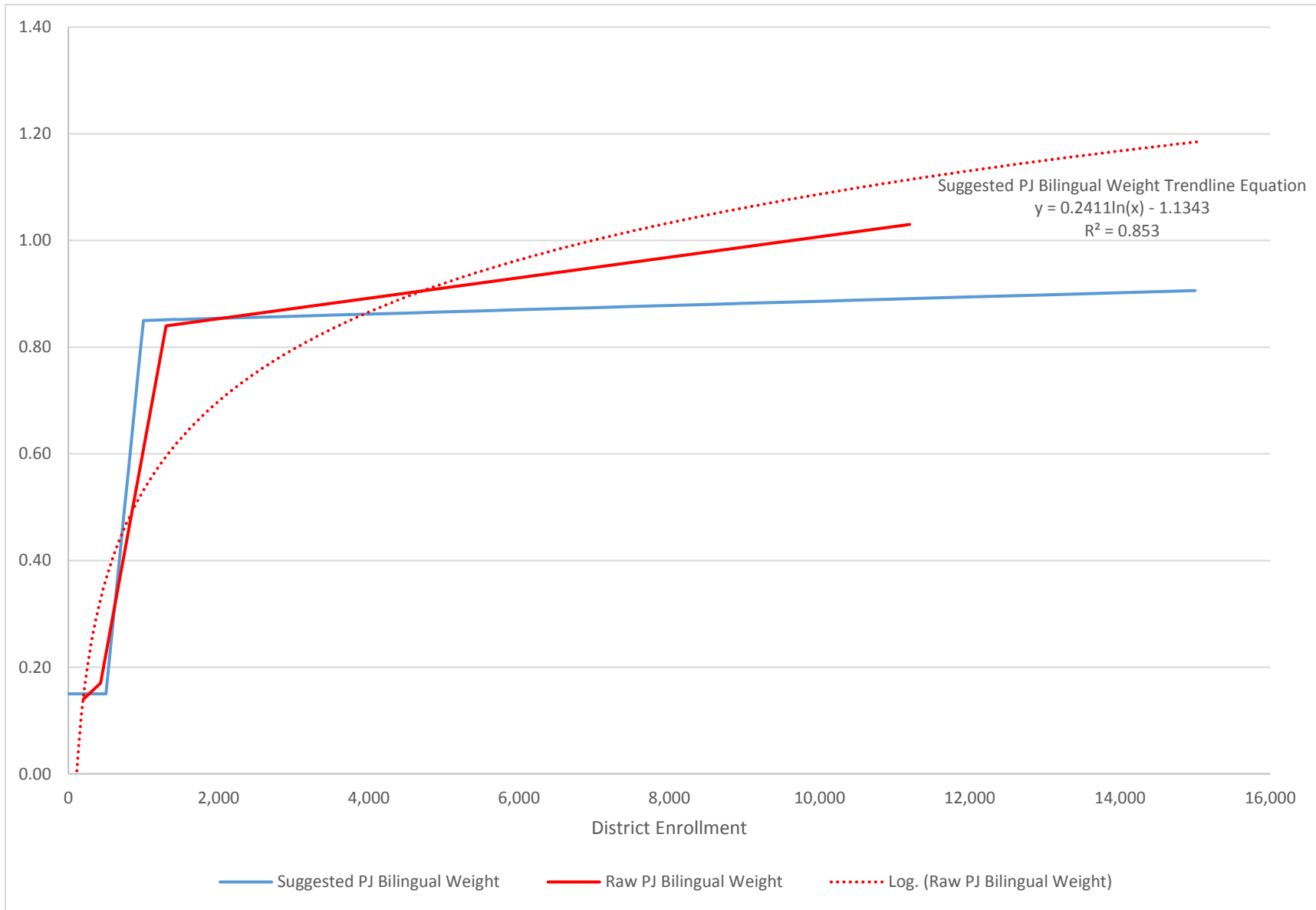
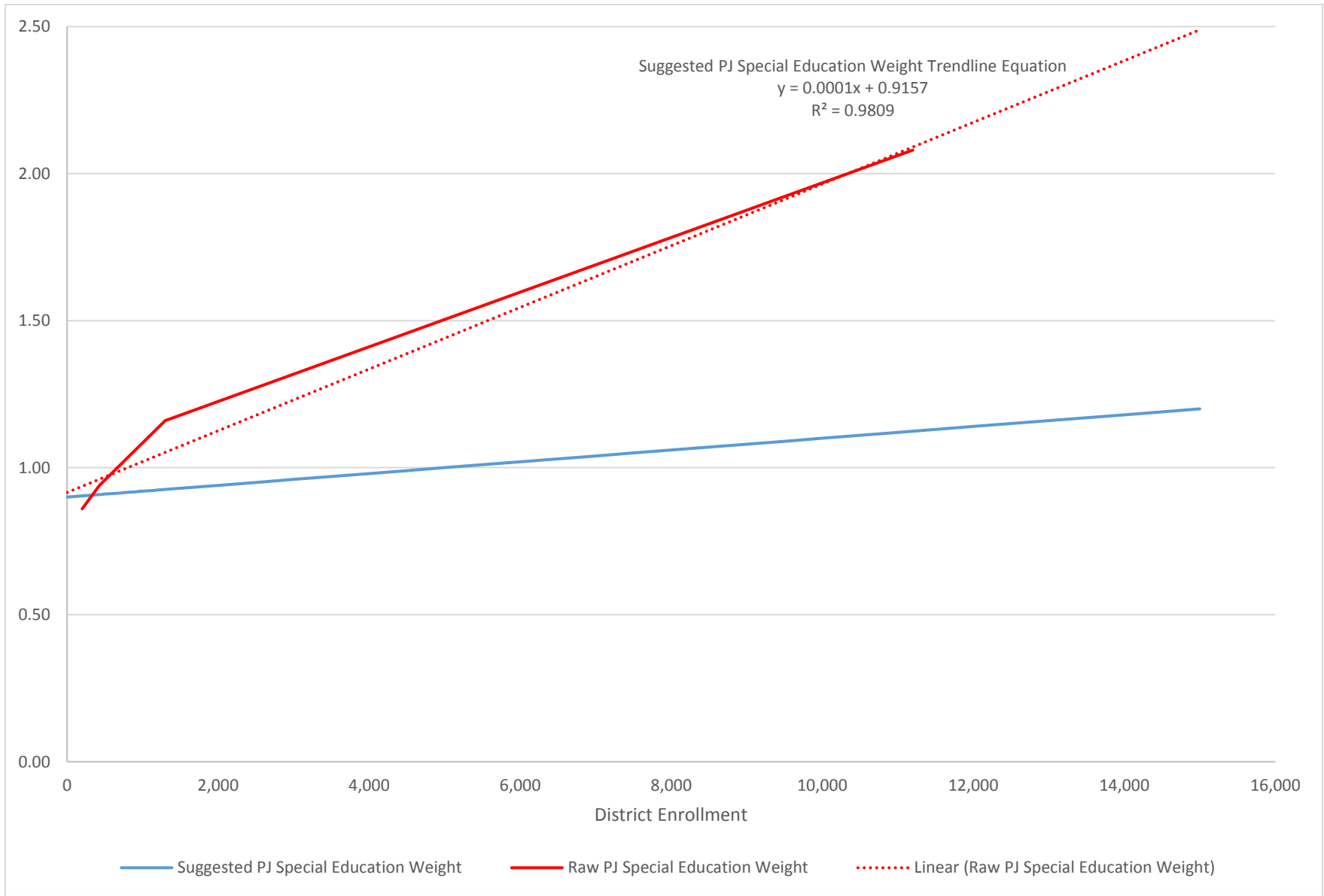


Exhibit 5 – A&M Suggested Special Education Weights by District Enrollment



The argument could be made, however, that the degree to which this is an underestimate will depend on the extent to which special education students' Individualized Education Programs (IEPs) include levels of support that constitute a suitable education (and the extent to which these services are actually provided). Again, similar concerns raised above for the other weights apply here, but the existence of a concentration effects seems less likely, but perhaps apparent given the large increase in the numbers of students in high-incidence special education categories (such as those who are specific learning disabled) and the potential disproportionate identification of these students in Moderate and Large sized districts. Again, as an alternative to the weight schedule I would propose that implementation of a constant special education weight calculated as the pupil-weighted average across the district size specific prototypes be considered (1.55).

Ensuring That PJ Models Are Efficient

As mentioned above, a key criticism of the PJ approach is that the specification of staffing and non-personnel resources by panelists may not represent efficient allocations of resources. That is, the contention is that the lists of resources specified through the panels' deliberations do not provide combinations that will achieve the outcomes put forth in definition of a suitable education at a minimum cost. To this end, more recent studies have incorporated safeguards to minimize the likelihood that the resource specifications and the corresponding estimates of sufficient cost might be deemed inefficient.²³

Caliber of Panelists and Transparency of Their Work

The objectivity and expertise of the educators involved in the PJ process is critical to the strength of the final product. In turn, PJ studies should ideally employ a highly selective recruitment process in which nominations are solicited from a wide group of educational organizations to identify potential PJ panel candidates. This has been done in previous studies through various processes such as the following (Chambers et al., 2004a,b; Chambers, Levin & Delancey (2006); and Chambers et al., 2008a,b):

- Soliciting nominations at town hall meetings or other forms of public engagement, or by directly contacting district superintendents, school boards, and professional education associations throughout the state.
- Soliciting nominations from schools identified as being extraordinarily successful through a beating-the-odds analysis (described earlier).

Ideally, nominators or candidates themselves will be required to complete a questionnaire asking about their educational experience and preparation, job histories, and special areas of expertise. The questionnaires should then be reviewed by the study team and selected from districts located in all parts of the state. Furthermore, the names of the panelists should be made a matter of public record by being published in the final report. Sometimes, panelists are required to present their work in public to stakeholders and that other higher-level panels will be reviewing their work, which adds an important element of accountability to the process. In light of this effort to be transparent, panelists were instructed to treat this effort seriously, base their deliberations upon their expert professional judgment, and fulfill their charge to develop school program designs and resource specifications that

²³ For specific details on comprehensive costing-out studies that include these safeguards, the reader is referred to Chapter 4 – The Comprehensive Costing-Out Study Component 2: Specifying and Costing Out Programs and Resources in Chambers & Levin (2009).

would achieve the goals statement objectives at a minimal cost. By utilizing a selective recruitment process and putting into the public light individual educators' professional reputations helps assure that panelists complete their work in a responsible manner and develop appropriate efficient models.

The A&M study states that panelists were chosen in consultation with the KSDE and LEPC, but goes no further in describing *how* the panelists were chosen. Exhibit 6 provides a map of the school-site panelists, which shows there seems to have been sufficient panelist representation of the state. In addition, the names of the panelists were made public (listed in the study in Appendices C-1A, C-1B, and C-1C).

Multiple Independent PJPs Performing Duplicate School/District Prototypes

Cost analysis making use of PJ relies heavily on resource specifications developed by one or more panels of educators. However, the importance of assembling multiple panels whenever possible cannot be stressed enough.²⁴ The use of multiple panels increases the reliability of the results by preventing the dependence of the findings on the judgment of a single panel. The panels should be instructed to work independently from one another and their deliberations occurring in different rooms. Moreover, they should be instructed to not communicate with individuals outside of their panels for the duration of the panel convening. Finally, each panel should include individuals representing a comprehensive range of professional roles. For example, each panel should ideally contain each of the following roles: a superintendent; principals and teachers from all three schooling levels (elementary, middle, and high); a special education specialist; a bilingual education specialist; and, a school business official.

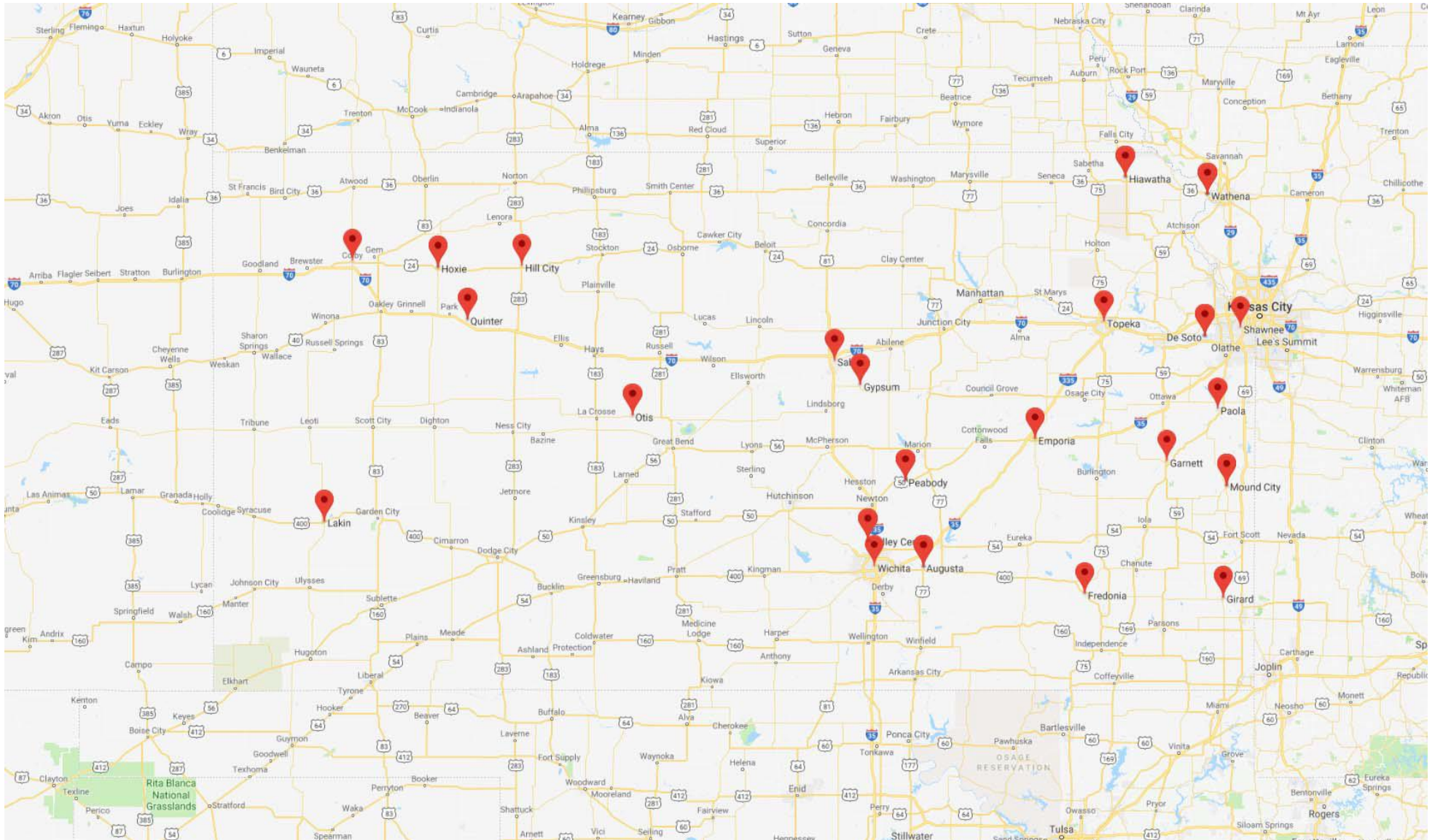
The A&M study was interesting in that it had separate school-site, district and expert panels. The A&M study lists the titles of the individuals serving on each of these. While it did not specify how these individuals were broken out into the four school panels or two district panels, from the provided list of school-site panelists we can ascertain that there were not enough panelists to develop fully comprehensive panels such as those described above.²⁵ For the 25-person school panel, there were eight teachers, six curriculum staff, five principals, three school business managers, two special education staff, and one superintendent. To this end, teachers and principals at all three schooling levels could not be represented on all school-site panels and there were not enough school business managers, special education staff or superintendents to go around for all four panels.

There were 15 staff serving on the two district panels. These two were split to review the work of the Very Small/Small panel and one of the Moderate size panels, and the Large panel and other Moderate size panel, respectively. The list of panelists was made up of (assistant) superintendents, finance officers, and teachers, and designated seven as "Avg.", three as "Lg.", two as "Sm." (understood to be coming from Average, Large and Small districts, given the cities in which they were located), and the remaining three without designation. In turn, it seems that there was more than appropriate coverage in terms of panelists to review the Moderate size panels work, but probably less than ideal numbers of panelist from Very Small/Small and Large districts.

²⁴ Previous costing out studies in New Mexico and New York that made use of six and eight independent panels, respectively, that independently developed models for identical prototypes (Chambers et al., 2008a,b; and Chambers et al. 2004a,b).

²⁵ However, it is assumed that they were allocated appropriately to the one panel working on the Very Small and Small district prototypes, the two panels working on the Moderate size district prototypes, and one panel working on the Large district prototypes.

Exhibit 6 – Map of Locations of School Site Panelists



The use of multiple panels working on identical prototype exercises limits the potential for any one panel with inefficient specifications to bias the results. Moreover, by selecting multiple panels and assigning identical exercises, the research team provides an incentive for each individual panel to be as efficient and thoughtful as possible in the design of its educational programs to achieve adequacy. The notion is that no individual panel wants their resource specifications to stand out as overly rich, while at the same time, no panel wants to be accused of omitting important design elements typical of successful schools. Ensuring that panels perform their work independently from one another will tend to prevent any bias resulting from collusion amongst panelists to develop richer specifications than they otherwise would have chosen. The extent to which each panel is made up of a well-balanced group of educators with respect to their roles also contributes to limiting the potential for panel over-specification of resources.

Unfortunately, the A&M study was somewhat lacking with respect to employing multiple panels working on identical exercises. There were only four panels, one working on the Very Small and Small district prototypes, one working on the Large district prototypes, and only two that I assume worked in parallel independently developing models for two sets of identical Moderate size district prototypes. Although it was not made clear in the study, I further assumed that the Moderate school prototype model presented was some sort of average of the individual panels' work.

Charge of PJs to Develop Efficient Models

The charge of PJ panels is to develop schooling models that will achieve the definition of a suitable education *at a minimum cost*. This should be made clear to panelists both through the written materials they were given and through the facilitation given during their deliberations. As an example, for the AIR study conducted in New Mexico the requirement that they develop efficient programs is stated clearly in the written PJ panel instructions (Chambers, 2008b) as shown in Exhibit 7.

To relay the importance of providing high-quality models that minimized costs the New Mexico study team also developed the acronym GEER (**G**oals, **E**vidence, **E**fficient and **R**ealistic) representing the following four questions that were continually asked of the PJ panels throughout their meeting.

- **Goals:** Will your program designs and resource specifications allow students to achieve the objectives in the goals statement?
- **Evidence:** Is there research evidence that supports your program designs and suggested use of resources?
- **Efficient:** Will your program designs and resource specifications achieve the goals at a minimum cost?
- **Realistic:** Can your program designs and resource specifications realistically be implemented?

In the earlier study conducted by A&M for Kansas, I could find no mention of developing *efficient* resources in the panel instructions. However, this is not to say that this important point was not discussed in person with the panels at the meetings.

Exhibit 7 – Excerpt from New Mexico Professional Judgment Panel Instructions

Statement of Purpose

The ultimate purpose of this work is to help us estimate the cost of providing an *adequate* education for all public school students in New Mexico. There are four components required to achieve this objective:

- **Define adequacy.** First, we are providing the PJPs with a *Goals Statement* (Exhibit A.1) that will define what is meant by the term “adequate education.” The *Goals Statement* incorporates input from a Stakeholder Panel established for this project and from a series of public engagement meetings held throughout the state in the Fall of 2006.
- **Design programs.** Second, we are asking each PJP to work independently to design educational programs at the elementary, middle, and high school levels that, in the judgment of the panel members, will provide an adequate opportunity for students in schools with varying demographics to have access to the learning opportunities specified in the *Goals Statement* (see Exhibit A.1) and to achieve the desired results.
- **Specify resources.** Third, each PJP will be asked to specify the resources and services necessary to deliver those programs in elementary, middle, and high schools in New Mexico.
- **Estimate costs.** Fourth, the AIR research team will use the information provided by each PJP to estimate the cost to deliver “adequate” educational programs in each and every public school and district in the state.

The charge of the PJPs is to complete components 2 and 3, above. Please note that we are **not** asking PJPs to create a “one size fits all” model to be implemented in all New Mexico public schools. Rather, we are asking panels to design instructional programs and specify the resources that they believe will deliver the desired results as **efficiently as possible** (i.e., at the lowest possible cost to taxpayers). These program designs and resource specifications simply provide us with a basis from which to estimate the costs of achieving the goals and to show how these estimates might be used to modify the existing school funding formula. By developing cost estimates for an adequate education from the work of six independent panels, we can measure how sensitive the cost estimates of the panels are to alternative assumptions of what resources are required to deliver an adequate education.

Professional Judgment Review Process

As part of PJ studies, the research team will often incorporate a formal review of the PJ panel models. The express purpose of this review was to ensure that the final models are both efficient and based upon a realistic and grounded set of specifications and cost estimates. The A&M research team explicitly included a review process in their design by appointing both a district-level panel and an expert panel. In turn, there were two sets of reviews incorporated into the study design. In addition, they report that these higher-level panels played an active role noting that they suggested additional school-level resources and modifications to certain resource prices. That being said, it should be noted that the expert panel only reviewed one of the four panel-specific models (one of the two Moderate district size models) that had been developed. It clearly seems like this was not enough time to perform a thorough review of the work of the panels developed each of the four district size categories.

This relates to a more general issue with the study in terms of the amount of time provided to the school, district and expert judgment panels to develop and review their models. The school and district panels had 1.5 days to complete their work, while the expert panel only was given 1 day. In my opinion, this is not enough time for panelists to become sufficiently familiar with their charge, engage in in-depth deliberations as to the resource needs for each of the prototypes, etc. Studies I have personally been engaged in have allocated three days to in-person PJ panel meetings, which is often followed up by telephone engagements.

Validating Results of PJ Results

The validity of cost study results is important to consider. Specifically, it is important to answer the following question:

Does the cost estimate really estimate the costs of producing the desired level, depth and breadth of educational outcomes, including whether and how those costs vary from location to location and child to child?

Far too little attention has been paid to methods for improving validity in education cost analysis (Baker & Levin (2014)). Moreover, validating cost studies using input-oriented approaches such as PJ is inherently difficult because the suggested spending is for hypothetical districts and schools. In contrast, outcome-oriented approaches such as cost functions, which are based on existing data that describe the relationships between spending, outcomes and cost factors (student needs, scale of operations and price levels of inputs) are easier to validate. Nevertheless, despite the costing-out approach that is used, it is important to be confident that any suggested funding increases deemed necessary to provide a suitable education would be targeted to districts and schools according to their needs. I could find no attempt on the part of the A&M study authors to do this. However, the following provides an example of how the results of previous PJ studies have been validated.

Clearly, to provide an equal opportunity for all students to achieve a state's educational goals, regardless of their circumstances, funding must be provided in an equitable manner. This calls for a check of the projected distribution of sufficient funding generated by a costing-out study to make sure that funding is properly aligned with needs. To this end, it is important to validate the results of a costing-out study by evaluating the relationship between the projected additional funding necessary to provide a suitable education and outcomes such as student achievement. If the model is working as intended so that adequate funding is provided in an equitable manner that affords all students an equal opportunity to achieve regardless of their needs or location, then we should see a systematic relationship between a district's relative need (how much more/less they need to provide a sufficient education) and student outcomes such as achievement on standardized tests.

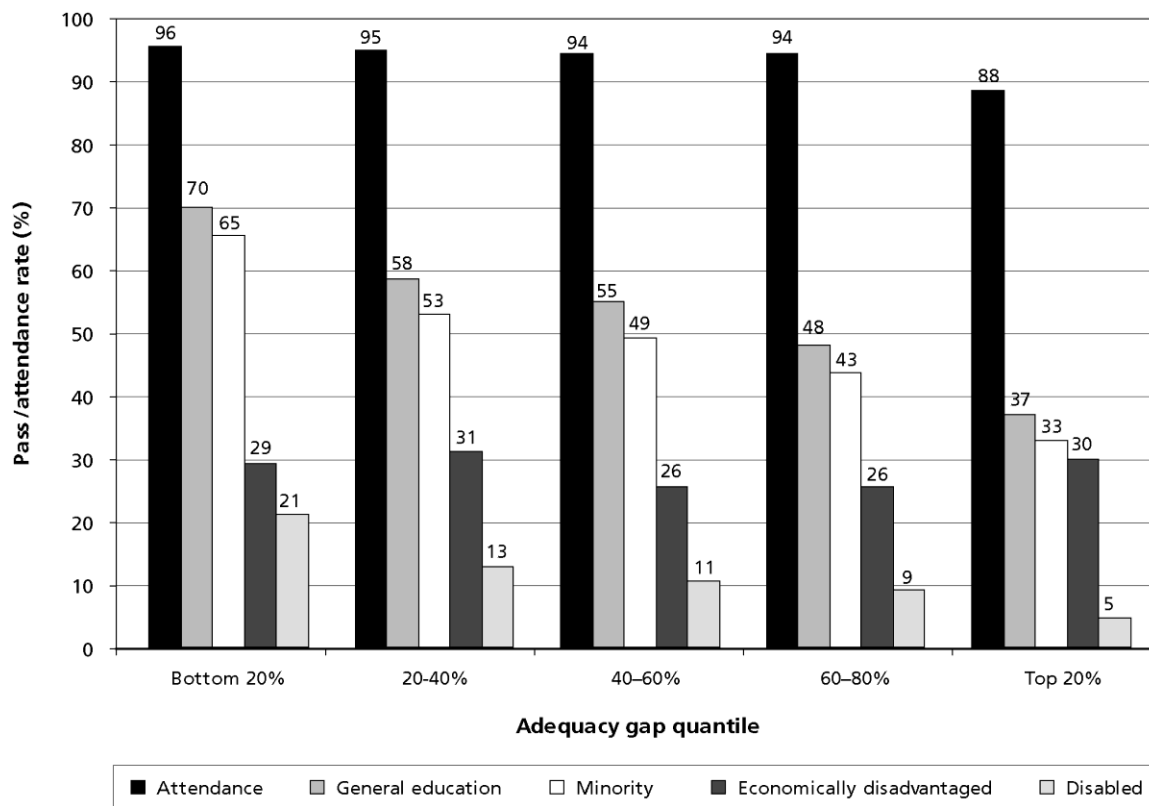
As an example, previous studies have performed this type of validation analysis for large-scale costing-out studies in New Mexico (Chambers et al., 2008a) and New York (Chambers et al., 2004a; Chambers, Levin & Parrish, 2006). The analysis involves calculating the funding shortfall or *Adequacy Gap*, which is a district-level measure defined as the relative difference between the projected necessary per-pupil funding to provide a sufficient education and actual per-pupil funding. Mathematically, it is simply the ratio of projected adequate to actual per-pupil funding for a given district:

$$(5) \quad \text{Adequacy Gap} = \text{Adequate Per-Pupil Funding} / \text{Actual Per-Pupil Funding}$$

Values that are greater than 1.00 indicate that the district needs more than it is currently receiving to provide an adequate education, while values that are less than 1.00 imply that the district is getting more than it needs to achieve sufficiency. Note that the adequacy gap is a direct measure of *relative need* (i.e., it represents in percentage terms the amount necessary to achieve adequacy compared to what is received). As an example of this type of analysis, consider Exhibit 8 taken from Chambers, Levin & Parrish (2006) based on the results of the New York Adequacy Study.

In the exhibit, the leftmost group of bars corresponds to districts in the bottom 20 percent of the adequacy gap distribution (i.e., those with the lowest need for funding to achieve adequacy). In contrast, the rightmost group of bars in each chart denotes districts in the top 20 percent of the sufficiency gap distribution—that is, those districts that are most in need of funding to achieve sufficiency. Each bar represents an average outcome for districts within each adequacy gap category (quintile), where outcomes are 8th grade attendance rates and pass rates for various student populations on the New York standardized tests (specifically, the minimum pass rate out of the English and math tests).

Exhibit 8 – 2001–02 Student-Weighted District Average 8th Grade Attendance/Pass Rates across New York Districts by Adequacy Gap Quintile



Note: Pass rate is defined as the lower of the percentages of test takers scoring at level 3 or above on the English and mathematics CTB tests.

Putting the performance measures on the vertical axis, we would expect that districts with the poorest performance levels (represented by lower column heights on the chart) would exhibit the largest adequacy gaps. Indeed, with few exceptions, one observes that districts with larger adequacy gaps

exhibit lower average attendance and pass rates for virtually every group of students including general education, minority, economically disadvantaged, and disabled students. As an example, the pass rate for general education students drops from 70 percent for districts with the lowest relative need by almost half, to 37 percent, for those districts with the greatest relative need.

Use of Public Engagement

More recent applications of the PJ approach (Chambers et al., 2004a,b; Chambers et al., 2008a,b) have used extensive engagement efforts to better understand public sentiment concerning the public education system. Chambers and Levin (2009) cite several served by an in-depth public engagement effort. First, the process directly involves the public promoting “buy in” from those with an interest in public education. Second, it helps capture the public’s educational priorities in terms of both the outcomes they feel are important as well as the types of programs they think are most appropriate to deliver services, which can be incorporated into the development of the standards defining a suitable education. Finally, it sheds light on public willingness to commit funding to public education and the types of revenue streams (e.g., taxes, lotteries, etc.) they feel are most appropriate to support a suitable education. While the A&M engaged in outreach through administration of interviews and questionnaires, it is not clear that any of this information was used to develop the definition of a suitable education that the PJ panelists responded to.

Lack of Transparency

As a final note, the A&M study lacked transparency surrounding the deliberations of the PJ panels and the justification of their resource allocation decisions. While the quantities of different personnel and non-personnel resources chosen for the various school/district prototypes are necessary to calculate the costs of implementing these models, they do not capture *how* the combinations of resources will translate into coherent schooling programs capable of achieving the standards put forth in the definition of a suitable education. Transparent documentation decisions behind the specified resources also serves to keep the panelists accountable for their work and counter the common argument by critics of the PJ approach that the process is simply an educator wish list that necessarily results in inefficient decisions on the part of panels. Other more recent PJ studies (e.g., Chambers et al., 2008a,b) have carefully documented the resource allocation decisions of panels, which are then included in the final report.

4 – Review of Elementary and Secondary Education in Kansas: Estimating the Costs of K-12 Education Using Two Approaches (Kansas Legislative Post Audit Committee)

Study Methodology

The study by the LPA made use of both input-oriented and output-oriented approaches to investigate how much it would cost to provide various levels of educational services to suffice two different purposes. The following chapter describes each of the approaches, their main results, and discussion.

Expenditure Analysis (Input-Oriented Approach)

The input-oriented approach attempts to estimate an accurate cost of providing regular K-12 education defined as educational curricula, programs and services that are either mandated by statute or specified as high school graduation and State scholarship/college admissions requirements. The analysis was performed with the following steps:

- 1) Determine Mandated Requirements – The researchers compiled a list of requirements related to attendance (days and hours per year), curriculum subject areas and required high school credits, student assessments and health exams.
- 2) Develop District Prototypes – They next created eight prototype districts defined by the following enrollment sizes: 100, 200, 300, 400, 600, 1,100, 2,000 and 15,000. The prototypical numbers of schools by schooling level and enrollments were determined by a sample of 94 comparison districts with enrollments near each of the of the prototype sizes (the districts were sorted into individual comparison groups around each prototype size).
- 3) Determine Staffing Levels – Both the types and numbers of staff were selected for the prototype districts. To determine the types of staff that should be included in the prototypes, a survey was administered to 80 school districts. The numbers of different types of staff were determined for regular education teachers and other staff separately. Quantities of regular education teachers were assigned to the prototypes under the following three different scenarios:
 - a. Average class sizes of 20 students.
 - b. Average class sizes of 25 students.
 - c. Average class sizes of 18 students for grades K-3 and 23 students for grades 4-12.

The quantities of other staff were determined using accreditation standards (for principals, assistant principals, library specialist and counselors). For other staff positions the researchers made use of extant staffing data on the comparison districts and in order to be “efficient”, selecting the FTE level for each staff type that was associated with the 33rd percentile of the within-comparison group distribution (i.e., the level at which two-thirds of the districts have higher staffing levels and one-third have staffing levels below).²⁶ Operations and maintenance

²⁶ It is unclear whether the researchers calculated the 33rd percentile of raw FTEs of other staff or the 33rd percentile of their staffing ratios (defined as the number of staff divided by enrollment) for each staff type and then used the ratios to allocate various types of other staff FTEs to the different district prototypes. The latter would have been more accurate in the cases where there was significant variation in staffing levels across districts within a comparison group.

staff were excluded because they are sometimes contracted out, so instead the 33rd percentile of the five-year historical average spending per-pupil on these functions was used.²⁷

- 4) Determine Average Salaries – Extant salary data was used to calculate Statewide average salaries for teachers and other staff including superintendents, assistant superintendents, principals and assistant principals. For other positions, average salaries were derived from a survey of 90 districts. The final compensation rate for each staff type was calculated using a 17 percent benefit rate.
- 5) Determine Non-Salary Resources – Extant district-level fiscal data was used to calculate the five-year inflation adjusted averages of non-salary expenditure per student. To create “efficient” estimates of spending to apply to the prototypes the researchers calculated the 33rd percentile of non-salary spending per-pupil within each district comparison group.
- 6) Calculating and Projecting Overall Spending Per Student – The overall spending per student was then calculated for each of the eight prototype districts and a cost curve developed (i.e., a schedule showing the relationship between per-pupil spending and district enrollment), with which projected spending per pupil for each district could be determined.
- 7) Developing Enrollment Weights – Weights from the generated cost curve for low- and high-enrollment were calculated and compared to the low- and high- (correlation) weights in the current State formula.

The LPA study also performed calculations of the additional costs of special education spending, vocational education, and transportation. The additional costs of special education spending (i.e., costs spent on special education students above and beyond those dedicated to their regular education) were based on the reported expenditures of 19 districts and the interlocals or cooperatives serving these districts that claimed to have both recorded all identified needs for their students with IEPs and provided all specified services included in these programs. Additional costs of vocational education were calculated by identifying through a survey 21 districts that could differentiate expenditures that were part of an approved program and examining their spending data. Additional transportation costs were calculated by a careful review of the current formula used and how closely it adhered to the assumption that students who live more than 2.5 miles from their schools are on average twice as costly to transport as are those who live within a 2.5-mile proximity.

Finally, the LPA study performed an analysis of regional variations in the cost by estimating a Hedonic wage model (Chambers, 1981), which uses a statistical model to explain variation in teacher salaries using factors that are within and outside of the control of districts including measures related to teacher characteristics, fiscal capacity, cost of living, community amenities and working conditions. An index measuring how much more or less than the state average it costs to hire and retain similarly qualified staff in each district is then derived using the estimates corresponding to those model factors deemed outside of district control.

Cost Function Approach (Outcome-Oriented Approach)

The cost function approach attempts to answer a different research question than the input-oriented approach. Here the purpose was not to cost out a collection of inputs that meet statutory requirements, but rather to estimate what it would cost districts to meet performance outcomes

²⁷ It is assumed the five-year average was based on inflation-adjusted (real rather than nominal) per-pupil spending.

specified by the State Board of Education. To do this, a cost function approach was employed in which statistical (regression) analysis was conducted to estimate the relationship between district per-pupil spending and an outcome (defined as the district average proficiency rate on six grade-specific math/reading and graduation rate) holding constant a host of educational cost factors including: student needs (percent FRL, bilingual headcount), district enrollment (defined across eight categories), teacher salary level, and indirect proxies for efficiency (district property wealth and income per pupil, ratio of state/federal aid to income, local tax share, percent of college-educated adults, percent of population 65 and over, and incidence of owner-occupied housing). The outcome used in the cost function was a composite defined as the average of district-level proficiency rates on the six criterion-referenced tests in math and reading used for accountability purposes (see Table 2 for the different grade/subject combinations) and the graduation rate defined on a cohort basis (i.e., percent of newly entering 9th graders that graduate four years later). The estimated cost function was then used to derive a base per-pupil cost and weights corresponding to the student needs and enrollment cost factors.

Key Results and Discussion

Key Results

Some key results from the input-oriented approach are displayed in Table 7. The first three columns of the table show the estimated per-pupil costs across the eight district prototypes for each of the three class size scenarios. The authors find that the per-pupil spending estimated from the prototypes most of the time were lower than actual funding. For example, for prototypes associated with 200 through 1,100 student districts the amount by which current funding per pupil exceeded the estimated per-pupil spending ranged from \$132 (for district size prototype 1,200 and scenario equal to a class size of 25) to \$1,248 (for district size prototype 400 and scenario equal to a class size of 25). Only in the smallest and largest district prototypes was current funding shown to be less than what the input-oriented approach estimated. For example, for district size prototype 2,000 and scenario equal to a class size of 20 the amount by which the estimated per-pupil spending exceeded current funding per pupil was \$595.

The special education analysis generated estimated an additional spending per special education pupil FTE equal to \$14,232, which was \$3,496 more than was currently being funded (\$10,736). The estimated additional cost for vocational education was \$1,375 in 2005-06 dollars or 32.3 percent of the base per-pupil funding for that year (equal to a weight of 0.32). This is less than what the current funding formula provided for each vocational pupil FTE (\$2,129, equal to a weight of 0.50).

The transportation analysis found that the current formula at the time (2005-06) was overfunding transportation. While the original system was supposed to fund transportation for students under the premise that those living over 2.5 miles from their school are twice as costly as those living within a 2.5-mile radius of their school. The authors showed that the existing formula was not funding districts in a manner that was consistent with this premise; a disproportionate amount of funding was being allocated for the transportation of students living more than 2.5 miles from their schools. As a result, the formula was providing \$13.9 million more in funding (\$80.8 million) than the LPA analysis estimated it should have (\$66.9 million).

Table 7 – Main Results from LPA Input-Oriented Approach: Estimated Per-Student Expenditures for Regular Education Using the Input-Oriented Approach (a), Compared with Current Funding Formula (b) 2005-06 School Year and Differences

Prototype District Size	Estimated Per-Student Expenditures				Difference Between Current Funding Formula and Input-Oriented Approach (Relative Difference in Parentheses)		
	Model Class Size 20^a	Model Class Size 25^a	Model Class Size 18/23^a	Current Funding Formula^b	Model Class Size 20	Model Class Size 25	Model Class Size 18/23
100	\$9,286	\$9,286	\$9,286	\$8,575	-\$711 (-7.7%)	-\$711 (-7.7%)	-\$711 (-7.7%)
200	\$7,098	\$7,098	\$7,098	\$7,447	\$349 (4.9%)	\$349 (4.9%)	\$349 (4.9%)
300	\$5,834	\$5,352	\$5,634	\$6,318	\$484 (8.3%)	\$966 (18.0%)	\$684 (12.1%)
400	\$5,464	\$4,926	\$5,251	\$6,174	\$710 (13.0%)	\$1,248 (25.3%)	\$923 (17.6%)
600	\$5,399	\$4,840	\$5,182	\$5,884	\$485 (9.0%)	\$1,044 (21.6%)	\$702 (13.5%)
1,100	\$5,029	\$4,466	\$4,838	\$5,161	\$132 (2.6%)	\$695 (15.6%)	\$323 (6.7%)
2,000	\$4,943	\$4,375	\$4,748	\$4,348	-\$595 (-12.0%)	-\$27 (-0.6%)	-\$400 (-8.4%)
15,000	\$5,062	\$4,497	\$4,886	\$4,348	-\$714 (-14.1%)	-\$149 (-3.3%)	-\$538 (-11.0%)

*Notes: Table derived from LPA Appendix 11.
(a) 2004-05 input-oriented approach estimated per-student expenditures inflated to 2005-06 school year.
(b) 2005-06 school year Base State Aid Per Pupil, plus low enrollment and correlation weighting.*

The regional cost analysis conducted by the authors generated a salary index that ranged from 95.7 to 109.6. That is, the cost of hiring and retaining teachers was 9.6 percent more than the Statewide average in the highest cost district and 4.3 percent less in the lowest cost district. In addition, the authors calculated a regional cost index that effectively only applies half of the salary index adjustment to each district. The authors claim this is logical because teacher compensation (salaries and benefits) make up only 50 percent of a school district's operating costs.

The cost function approach generated an estimated regression that estimated an equation capturing the relationships between per-pupil cost and a host of variables described including a composite outcome, student needs, enrollment, measures of district efficiency, and year indicators. The equation was then used to predict district-level spending capable of producing a suitable education defined as the State performance outcome standards in 2003-04, 2004-05 and 2005-06 (which had the same standards) and 2006-07 at a minimum cost. These results were used to calculate cost indices and weights for poverty, bilingual and enrollment. The cost function at-risk (FRL) and enrollment weights varied significantly across districts; the at-risk weights ranged from 0.65 to 1.15 with a median of 0.70 and the enrollment weights ranged from 0.00 to 0.77 with a median of 0.14. While the at-risk weights were higher than the 0.19 weight used in the State funding system, the enrollment weights were lower than those contained in the funding system. In contrast, there was virtually no variation in the bilingual weights, which held steady at 0.14 across all districts. The authors claim that it is likely the costs associated with at-risk students may be covering the additional costs of EL, given how close relationship (the degree of overlap) between these two student populations.

The estimated costs to reach the performance outcome standards generated by the outcomes-oriented cost function approach were higher for the four years that were costed out. Compared to the funding provided by the existing funding formula (\$2.159 billion or \$4,856 per pupil) it was estimated to cost \$115 million more (equal to \$258 per pupil) in 2003-04, \$315 million (equal \$709 per pupil) more in 2004-05/2005-06, and \$513 million more (equal to \$1,153 per pupil) in 2006-07. The corresponding relative increases for these years are 5.3, 14.6 and 23.8 percent, respectively.

The study drew upon both the input- oriented and outcome- oriented approaches taken to develop a range of estimated costs associated with providing a suitable education. Table 8 presents three estimates that drew upon the base per-pupil cost and enrollment weights estimated using the input-oriented approach and a fourth that used an adjusted base that excludes the portion covered by Federal funding and enrollment weights from the outcome-oriented approach. The remaining weights and funding adjustments applied to all four estimates were taken from the outcome-oriented approach (for the at-risk, at-risk/pupil density and bilingual weights) and the additional analyses of special and vocational education (input-oriented approach), transportation, and regional labor costs. While there were four different estimated cost figures, the general result is that all proved to be higher than what was being provided by the current funding system. Specifically, the authors found that the additional funding necessary using the base per-pupil funding and enrollment weights generated by the input-oriented approach ranged from \$316 to \$623 million or from 11.5 to 22.7 percent, depending on class size scenario. The additional funding necessary to provide a suitable education using the base and enrollment weights from the outcome-approach was \$399.3 million or 14.5 percent. Note, the outcome-oriented approach additional cost is about at the midpoint between the input-oriented approach figures for the 25-student and average 18/23-student scenarios.

Table 8 – LPA Cost Study Results Compared to State Funding Formula (Figure 1-1 from LPA Study)

Figure 1-1 Comparing Cost Study Results to the Current State Funding Formula 2005-06 and 2006-07					
	Current Funding Formula	Input-Based Approach (Using 3 Class-Size Models)			Outcomes- Based Approach
		Average 25 students/class	Average 18/23 students/class	Average 20 students/class	
Base-level costs per FTE student	05-06 = \$4,257 06-07 = \$4,257	05-06 = \$4,375 06-07 = \$4,519	05-06 = \$4,748 06-07 = \$4,904	05-06 = \$4,943 06-07 = \$5,105	05-06 = \$4,167 06-07 = \$4,659
Low-enrollment weight (to 3 decimals)	range: 1.014–0.021	range: 1.122–0.000	range: 0.956–0.000	range: 0.879–0.000	range: 0.773–0.008
Correlation (high-enrollment) weight (to 3 decimals)	0.021 for districts $\geq 1,662$	range: 0.000–0.028 for districts $\geq 2,000$	range: 0.000–0.029 for districts $\geq 2,000$	range: 0.000–0.024 for districts $\geq 2,000$	0.008 for districts $> 1,700$
At-Risk (poverty) weight (per free-lunch student)	0.193	0.484			
Additional Urban-Poverty weight (per free-lunch student)	---	0.726			
Bilingual weight (two different bases)	0.395 per FTE bilingual student	0.100 per headcount bilingual student			
Additional cost per FTE Special Education student	05-06 = \$10,736 06-07 = \$12,185	05-06 = \$14,232 06-07 = \$15,159			
Additional cost per FTE Vocational Education student	05-06 = \$2,129 06-07 = \$2,129	05-06 = \$1,375 06-07 = \$1,420			
Additional cost per student transported >2.5 miles	05-06 = \$594 06-07 = \$613	05-06 = \$491 06-07 = \$507			
Regional cost adjustment (applied to teacher salaries)	---	range: -2% to +5% of costs			
Given above cost estimates, additional amount needed to provide "foundation-level" funding compared with current funding levels (in millions)	---	06-07 = \$316.2	06-07 = \$519.5	06-07 = \$623.7	06-07 = \$399.3
"Hold-harmless" provision so no district would receive less than under the current funding formula (in millions)	---	06-07 = \$35.1	06-07 = \$ 7.0	06-07 = \$ 0.7	06-07 = \$9.4

Source: LPA analysis of school district and Department of Education data.

Discussion

My general impression of the LPA study is that it is an impressive piece of work that represents an immense undertaking. Furthermore, the methodology and application seemed to be carefully thought out and implemented very well. Finally, the large volume of work was documented extensively by the authors and laid out in a fairly organized manner. In what follows, I provide discussion on various points of the study methodology and implementation, illustrating potential limitations in the work.

Expenditure Versus Cost Analysis (Input-Oriented Approach)

My main concern with the LPA study is with the sizeable effort devoted to using an input-oriented approach to conduct what I would refer to as *expenditure* rather than *cost* analysis. As stated in the cost function analysis writeup:

“The term cost in economics refers to the minimum spending required to produce a given level of output.” (Page C-4, Appendix 17)

While there are certainly costs involved in the purchase of personnel and non-personnel resources, these purchases are not the penultimate outcome of interest in terms of what a public education system is expected to produce. Rather educational cost studies attempt to better understand the system by which educational outcomes are produced, which necessarily involves relating inputs to student outcomes. Influenced by economists performing research in this area, any reference to costs should be accompanied by some measure of outcome that has been produced (in the current context, a suitable education for K-12 students in the Kansas public school system). In my description of the input-oriented approach above, I have tried to refrain from referring to this as an investigation of “cost”, but rather as an analysis of “spending”.

Additionally, it must be mentioned that the input-oriented approach is not purely input based. Specifically, it makes use of base per-pupil figures and enrollment weights that are borne out of the input approach, but then adds student need weights from the outcome-oriented approach, which is rather strange. This is mixing results from the outcome-oriented approach, intended to get at the cost of providing a suitable K-12 public education to all students with those of the input-oriented approach intended to get at the spending necessary to provide levels of programming and service that might be regarded as minimally required by law or regulation. However, further additions to the educational cost estimates based on existing expenditures on programs and services such as transportation is more commonplace in adequacy studies (or these are simply not considered in the cost estimates).

Please note that there is nothing inherently wrong with analyzing how much is being spent on programs and services that are required by statute and regulation. However, doing so answers a very different research question than the one that is at the heart of educational adequacy studies. One would expect that state statute and regulation more often than not dictate minima with respect to the quantity, types and quality of programs and services that must be provided in public schools. Indeed, the results above in Table 7 showing the estimated costs of providing regular education defined by only those required programs and services seems to be in line with this contention. Here, the suggested base per-pupil costs for all three school size scenarios stemming from the input-oriented spending analysis are generally less than what the current formula provides (except for the largest and smallest district prototypes). However, it must also be realized that spending at these lower levels might be associated with lower educational outcomes, which the input-oriented spending analysis does not take into account. The

bottom line is that the base per-pupil and enrollment weight figures generated by the input-oriented spending analysis do not legitimately represent the cost of providing a suitable education as defined by the student outcomes that should be produced.

Methodology to Produce “Efficient” Prototypes in Expenditure Analysis (Input-Oriented Approach)

Another closely related concern I have with the input-oriented spending analysis is the attempt to provide more “efficiency” in the input-oriented approach. For non-teacher staff other than principals, library specialists and counselors the approach bases spending for the prototype districts on the 33rd percentile of the distributions of staff per FTE in the district comparison groups. Similarly, for both staff and non-personnel spending on maintenance and operations, as well as other non-personnel resources the approach bases spending for the prototype districts on the 33rd percentile of the distribution of per-pupil spending in the district comparison groups. This was done to ensure that the spending identified is that of a district operating at an above-average level of “efficiency”.²⁸

It is assumed that the choice of pegging resource utilization to the 33rd percentile in the input-oriented approach was adopted from the application of the same tertile cutoff to the efficiency proxy variables for calculating weights in the outcome-oriented approach (i.e., the (in)efficiency proxy variables were set to relatively (low) high levels when predicting weights). However, I would argue that this practice does not logically translate over to the input-oriented setting and is an incorrect use of the term. Efficiency, by definition, is determined by level of output produced using a given amount of resources or alternatively by the amount of resources used to produce a given level of output. As an example, in order to show that producing unit A is more efficient than B, one would have to demonstrate that A produced at least the same amount of output while using fewer resources than B. Alternatively, one could also demonstrate this by showing unit A produced more output than B while using at most the same level resources. The input-oriented spending analysis did not take into account the level of student outcome being produced by each district so that those districts using the 33rd percentile of a given resource cannot be referred to as operating at above-average efficiency, but only rather as operating at below-average spending, with unknown consequences as to what this would have on student outcomes.

Application of Regional Labor Market Cost Adjustments in Expenditure Analysis (Input-Oriented Approach)

The input-oriented approach used in the study correctly attempts to adjust for geographic variation in teacher salaries. Indeed, it seems that great effort went into developing a Hedonic wage model for the State. I found the methodological approach and implementation in line with best practice (Chambers, 1998). However, the application of the model results raises some concerns. The main result of the Hedonic wage model was the teacher salary index, a standard index centered around 100.0, representing the state average, that measures how much more or less costly it is to hire and retain a comparably qualified teacher in different districts (e.g., an index value of 110.0 indicates that teachers are 10 percent more costly than the state average). However, this is not what was applied to teacher compensation.

²⁸ As a small technical statistical side note, the 33rd percentile is not necessarily lower than the average; when a distribution be sufficiently skewed to the left (i.e., the mean is far below the median) then the 33rd percentile will be above the average.

Instead, the authors calculated what they refer to as a regional cost index, which simply reduced the absolute magnitude of the teacher index values by half as show in the following equation:

$$(6) \quad \text{Regional Cost Index} = [(\text{Salary Index}) - 100] \times 0.5 + 100$$

The justification the authors provide for the development and application of the regional cost index is that spending on teacher compensation (salaries and benefits) tends to make up approximately 50 percent of a district's operational spending. Furthermore, this regional cost index was only applied to teacher compensation, which was based on a standardized Statewide average salary.

As far as I can tell, the compensation for other staff was not adjusted, or at least directly, for the geographic variation across the state. Indirectly, however, it could be said that there were indirect adjustments made. Specifically, for superintendents, assistant superintendents, principals, and assistant principals, instead of calculating compensation rates based on Statewide average salaries, the authors chose to use average salaries within the eight comparison district groupings. This was done because the salaries seemed to be correlated with district size. However, to the extent that district size is related to the teacher salary index, the calculation of salaries for these staff types was an indirect and likely inaccurate adjustment. Similarly, for a host of other staff types for which Statewide salary data was not available, the authors surveyed 90 districts and took averages within district groups defined by three size categories.

Given that it is widely accepted that the differential level of teacher salaries across districts is a good indicator of the general cost of all educational staff, it seems that it was a mistake not to apply the teacher wage index to all staff. Moreover, I assume that the only reason the regional cost index was developed was to address the costs of teaching staff and perhaps the perception that the teacher wage index could not be legitimately applied to non-teaching staff. If this assumption is correct, then the decisions described above are rather surprising given that the cost function analysis text clearly suggests that teacher salary levels are indicative of the salary levels of all district personnel, as well as non-personnel resources:

“In addition, teacher salaries are typically highly correlated with salaries of other certified staff, so that teacher salaries serve as a proxy for salaries of all certified staff.” (Page C-13, Appendix 17)

“We find that, a one percent increase in teacher's salaries is associated with a 1.02 percent increase in per pupil expenditures. Because professional salaries typically represent 80 to 85 percent of operating spending, this result suggests that higher teacher salaries tend to be associated with higher salaries for all personnel hired by a district, as well as with higher prices for contract services.” (Page C-18, Appendix 17)

In sum, in my opinion the authors should have developed Statewide average salaries for the non-teaching staff and applied the teacher salary index (not the more compressed regional cost index) to all calculated staff expenditures. The implication of not doing so was likely significant, as compensation for non-teaching personnel Nationwide made up an additional 30 to 31 percent of current operational cost in the time period used in the study (Table 9):

Table 9 – Nationwide Total Compensation as Share of Current Operational Spending (2000-01 to 2005-06)

	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06
Compensation as Share of Total Current Expenditures	81%	81%	81%	81%	81%	80%
<i>Source: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "National Public Education Financial Survey," 1990-91 through 2005-06.</i>						

Adjustments to Cost Function Base Per-Pupil Cost and Weights (Outcome-Oriented Approach)

While the outcome-oriented approach rightfully includes all operational spending in order to calculate the cost of supporting a suitable education, which included Federal funding, the authors wanted to adjust the estimated cost so that it would only represent dollars that would have to be funded by the State. In doing so, they calculated Federal funding that could be used to support base, at-risk, and bilingual education and then downwardly adjusted the estimated base-per pupil funding, at-risk and bilingual weights, respectively, to account for these Federal dollars. Specifically, they identified Federal funding that could be used for base, at-risk and bilingual education on the order of \$71.5, \$130.0 and \$4.0 million, respectively. They then downwardly adjusted the cost-function estimated base per-pupil cost figure until the total corresponding Statewide cost decreased by the \$71.5 million and then proceeded to decrease the at-risk and bilingual weights (using the lower adjusted base) until the total cost accounted was reduced by the \$130.0 and \$4.0 million. While the authors note that an alternative might have been to first calculate the total suitable cost for each district and then to subtract off the top Federal funding to come up with the State portion, this might pose an unacceptable risk of being perceived as the State supplanting Federal funding.

Unfortunately, there is often difficulty between fulfilling the objective of identifying the overall cost of providing a suitable education, which involves estimating a total cost that will be supported by both State and Federal dollars, and applying these revenue sources to the recommended formulaic base and weights in a manner that is not perceived as undermining the supplement-not-supplant clause in the law concerning Federal education funding.

While I appreciate the delicate situation, I am not certain that the solution developed by the authors is ideal. They essentially developed a new formula for distributing base, at-risk and bilingual dollars funding from non-Federal sources. One initial concern that I have is whether the resulting adjusted at-risk and bilingual weights preserve the relative differences between the original unadjusted weights. However, fortunately this concern can be dismissed as shown by the figures in Table 10. Columns 1 and 3 of the table show the original and adjusted weights. To understand how the relative difference between the weights may have changed after adjusting them to remove federal funding from the equation, the relative differences between the original general at-risk weight have been calculated in columns 2 and 4 (e.g., the original high at-risk, inner city weight was 1.499 larger than the original

regular at-risk weight, while the original bilingual weight was 0.198 of the original regular at-risk weight). Comparing the results in columns 2 and 4 we see that the relative differences in the weights were preserved after adjusting for federal funding.

Table 10 – Original and Adjusted Estimated At-Risk and Bilingual Weights

Weight	1 – Original Estimated Weight	2 – Relative Difference from Regular Poverty Weight	3 – Weight Adjusted to Remove Federal Funds	4 – Relative Difference from Regular Poverty Weight
At-Risk				
<i>Regular</i>	0.703		0.484	
<i>High At-Risk, Inner City</i>	1.054	1.499	0.726	1.500
Bilingual	0.139	0.198	0.100	0.207

Despite there being no issue in terms of the adjustments to the weights significantly altering their relative magnitudes, this brings to light another fundamental difficulty in implementing the funding mechanism recommended by a costing out study in the context of constraints related to federal funding sources. Specifically, while the authors have devised adjusted base per-pupil costs and weights that represent how State funding will be distributed, the costing-out study dictates that a suitable education requires that the total amount of State and Federal funding be spent (according to the base cost and weights of the original model). This implies that the Federal funding should also be spent in line with a funding mechanism that is the complement of the adjusted base and weights for distributing State funding. That is, if the authors performed the same procedure but instead adjusted downward the original base per-pupil cost, at-risk weight and bilingual weight so as to eliminate the portion of total necessary funding provided by the *State*, then the resulting second adjusted formula would dictate how Federal dollars would need to be distributed in order to provide a suitable education. Clearly, there are specific rules pertaining to how different federal funding sources must be distributed and it remains an empirical exercise to best understand how this would deviate from this complementary mechanism to appropriately distribute funding to provide educational suitability. This discussion emphasizes the need for states and the Federal government to work closely in order to broker more flexibility in how federal dollars can be used in the context of state school funding reform where state funding is slated to increase and become more equitably distributed.

Definition of Outcome in Cost Function Model (Outcome-Oriented Approach)

As mentioned in the brief overview of costing-out methodologies, a drawback of the CF approach is its reliance on an outcome measure that is usually defined by one or a collection of test scores/proficiency rates that are averaged into a single composite.²⁹ Indeed, the LPA outcome-oriented approach makes use of such a composite measure; namely, the district average proficiency rate on six grade specific criterion-referenced math/reading tests and a cohort-based graduation rate. Although this outcome may seem similar in part to that used in the input-oriented study conducted by A&M there is a significant difference. The outcome in the A&M study set proficiency thresholds on the same six tests

²⁹ This is in contrast to the PJ approach where the educational objective can be more broadly defined. Note that the EB approach is also limited, but in a different manner; outcomes in EB studies are constrained by those that have been analyzed in the research literature.

included in the average composite measure used in the LPA study, all of which would be necessarily be met within five years (by the 2006-07 school year). This is contrast to the composite measure used in the LPA study, which only required that proficiency rates would be achieved on average. In this sense, with respect to proficiency rates on the math and reading tests the A&M study was technically more stringent than the LPA study.³⁰ This is because the average used in the LPA study allows lower proficiency rates on some tests to be offset by higher rates on other tests.

To illustrate this point, Table 11 provides several different hypothetical scenarios where combinations of proficiency rates on the six tests are averaged. Let us consider a target *average* proficiency rate threshold of 75 percent and a secondary target where *all* tests must individually meet the 75 percent proficiency rate.³¹ The final two rows of the table show that the first scenario meets both targets (i.e., the average proficiency rate across the six tests is 75 percent and none of individual tests exhibit a proficiency rate that falls below the 75 percent threshold. In contrast, under Scenario 2 the average is still met even though one of the six tests (5th grade reading) falls below the proficiency threshold. The remaining scenarios show further combinations where the average threshold is met with increasing numbers of individual tests that do not meet the threshold.

Table 11 - Averages of Hypothetical Combinations of Proficiency Rates

Subject	Grade	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Reading	5	75%	0%	0%	10%	50%	70%
	8	75%	90%	70%	70%	60%	70%
	11	75%	90%	80%	70%	70%	70%
Math	4	75%	90%	100%	100%	70%	70%
	7	75%	90%	100%	100%	100%	70%
	10	75%	90%	100%	100%	100%	100%
Average Proficiency Rate		75%	75%	75%	75%	75%	75%
Number of Tests Below Proficiency Rate Threshold		0	1	2	3	4	5

This demonstration does not imply that the scenarios in which the average proficiency threshold is met while proficiency rates on one or more individual tests fall below the threshold did or did not exist across the State’s districts during the study period. In fact, if there was a strong positive relationship (correlation) in proficiency rates between tests (and the graduation rate) it is less likely that this posed a problem. Nor is the comment here meant to shed a negative light on the work performed by the cost function researchers. Rather, it is meant to demonstrate a common limitation of the cost function approach and how using an average composite outcome is less stringent than requiring all components of the composite outcome to be met.

³⁰ The focus on proficiency rates in this statement is important; note that the A&M study did not include graduation rate in the set of outcomes defining educational suitability.

³¹ While the simple example here uses a constant 75 percent proficiency rate threshold across all tests, it generalizes to the case where there are different thresholds for each test.

References

- Augenblick, J., Myers, J., Silverstein, J. & Barkis, A. (2002). *Calculation of the Cost of a Suitable Education in Kansas in 2000-2001 Using Two Different Analytic Approaches*. Denver: CO, Augenblick and Myers, Inc.
- Augenblick, J., Van de Water, G. & Myers, J. (1993). *Determining a Base Student Cost Figure for Use in Ohio's School Foundation Program*. Denver, CO: Augenblick, Van de Water & Associates.
- Baker, B. & Green, P. (2014). "Conceptions of Equity and Adequacy in School Finance," In Handbook of Research in Education Finance and Policy, 2nd Edition, M.E. Goertz and H.F. Ladd (eds.), New York: Routledge, pp. 231-43.
- Baker, B. & Levin, J. (2014). *Educational equity, adequacy, and equal opportunity in the Commonwealth: An evaluation of Pennsylvania's school finance system*. San Mateo, CA: American Institutes for Research.
- Baker, B., Taylor, L., & Vedlitz, A. (2008). *Adequacy Estimates and the Implications of Common Standards for the Cost of Instruction*. Washington, DC: National Research Council.
- Berne, R., & Stiefel, L. (1984). The Measurement of Equity in School Finance: Conceptual, Methodological, and Empirical Dimensions. Baltimore: Johns Hopkins University Press.
- Castellanos-Jimenez, O. and Topper, A. (2012). "The Cost of Providing an Adequate Education to English Language Learners: A Review of the Literature," *Review of Educational Research*, 82(2), 179–232.
- Chambers, J. (1981). "The Hedonic Wage Technique As a Tool for Estimating the Costs of School Personnel: A Theoretical Exposition with Implications for Empirical Analysis," *Journal of Education Finance*, 6(3): 330-54.
- Chambers, J. (1998). *Geographic Variations in the Prices of Public School Inputs*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, Working Paper No. 98-04.
- Chambers, J., & Levin, J. (2006). "Funding California's Schools, Part II: Resource Adequacy and Efficiency". In *Crucial Issues in California Education 2006: Rekindling Reform*, H. Hatami (ed.). Berkeley, CA: University of California, Policy Analysis for California Education.
- Chambers, J. & Levin, J. (2009). *Determining the Cost of Providing an Adequate Education for All Students*. Washington, DC: National Education Association.
- Chambers, J., Levin, J., & Delancey, D. (2006). *Efficiency and adequacy in California school finance: A professional judgment approach*. Palo Alto, CA: American Institutes for Research.
- Chambers, J., Levin, J., Delancey, D. & Manship, K. (2008a). *An Independent Comprehensive Study of the New Mexico Public School Funding Formula: Volume I – Final Report*. Report prepared for New Mexico State Legislature Funding Formula Study Task Force. Palo Alto, CA: American Institutes for Research.

- Chambers, J., Levin, J., Delancey, D. & Manship, K. (2008b). *An Independent Comprehensive Study of the New Mexico Public School Funding Formula: Volume II – Technical Report*. Report prepared for New Mexico State Legislature Funding Formula Study Task Force. Palo Alto, CA: American Institutes for Research.
- Chambers, J., Levin, J. & Parrish, T. (2006). “Examining the Relationship Between Educational Outcomes and Gaps in Funding: An Extension of the New York Adequacy Study,” *Peabody Journal of Education*, 81(2), pp.1–32.
- Chambers, J., & Parrish, T. (1982). *The Development of a Resource Cost Model Funding Base for Education Finance in Illinois, Volume I B, Executive Summary, and Volume II B, Technical Report* (prepared for the Illinois State Board of Education). Stanford, CA: Associates for Education Finance and Planning.
- Chambers, J., & Parrish, T. (1984). *The Development of a Program Cost Model and a Cost-of-Education Index for the State of Alaska: Final report. Volumes I–IV*. Stanford, CA: Associates for Education Finance and Planning, Inc.
- Chambers, J., Parrish, T. & Harr, J. (2002). *What Are We Spending on Special Education Services in the United States, 1999-2000?* Report. Special Education Expenditure Project (SEEP). Palo Alto, CA: American Institutes for Research.
- Chambers, J., Parrish, T., Levin, J., Guthrie, J., Smith, J. & Seder, R. (2004a). *The New York Adequacy Study: Determining the Cost of Providing All Children in New York an Adequate Education, Volume I – Final Report*. Report prepared for the Campaign for Fiscal Equity CFE. Palo Alto, CA: American Institutes for Research AIR/Management Analysis and Planning MAP.
- Chambers, J., Parrish, T., Levin, J., Guthrie, J., Smith, J. & Seder, R. (2004b). *The New York Adequacy Study: Determining the Cost of Providing All Children in New York an Adequate Education, Volume I – Technical Appendices*. Report prepared for the Campaign for Fiscal Equity CFE. Palo Alto, CA: American Institutes for Research AIR/Management Analysis and Planning MAP.
- Hanushek, E. (Ed.) (2006). *Courting Failure: How School Finance Lawsuits Exploit Judges’ Good Intentions and Harm our Children*. Hoover Institution Press: Stanford, California.
- Hanushek, E. (2007). The Confidence Men: Selling Adequacy, Making Millions. *Education Next*, 7(3), 73-78.
- Heckman, J. (2008). “Schools, Skills, and Synapses,” *Economic Inquiry*. 46(3): 289–324.
- Hollands, F., Pan, Y., Shand, R., Cheng, H., Levin, H., Belfield, C., Kieffer, M., Bowden, B. & Hanisch-Cerda, B. (2013). *Improving Early Literacy: Cost-Effectiveness Analysis of Effective Reading Programs*. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University.
- Kansas Legislative Post Audit Division. (2006). *Elementary and Secondary Education in Kansas: Estimating the Costs of K-12 Education Using Two Approaches*. Topeka, KS: Legislature of Kansas.
- Klitgaard, R. & Hall, G. (1973). A Statistical Search for Unusually Effective Schools. RAND report R-1210-CC/ RC. Santa Monica, CA: RAND.
- Levin, H. (1983). *Cost-Effectiveness: A Primer*. Newbury Park, CA: Sage Publications.

- Levin, H., McEwan, P., Belfield, C., Bowden, B. & Shand, R. (2017). Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis. Thousand Oaks, CA: Sage Publications.
- Levin, H., Belfield, C. Hollands, F., Bowden, B., Cheng, H., Shand, R., Pan, Y. & Hanisch-Cerda, B. (2014). *Cost-Effectiveness Analysis of Interventions That Improve High School Completion*. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University.
- Levin, H. & McEwan, P. (2001). Cost Effectiveness Analysis: Methods and Applications. Thousand Oaks, CA: Sage Publications.
- Odden, A., Picus, L. & M. Fermanich. (2003a). *An Evidence-Based Approach to School Finance Adequacy in Arkansas*. North Hollywood, CA: Lawrence Picus and Associates.
- Odden, A., Fermanich, M., & Picus, L. (2003b). *The State-of-the-Art Approach to School Finance Adequacy in Kentucky*. North Hollywood, CA: Lawrence O. Picus & Associates.
- Odden, A., Picus, L., Goetz, M., Mangan, M. & Fermanich, M. (2006). *An Evidence-Based Approach to School Finance Adequacy in Washington*. North Hollywood, CA: Lawrence O. Picus and Associates.
- Perez, M., Anand, P., Speroni, C., Parrish, T., Esra, P., Socias, M. & Gubbins, P. (2007). *Successful California Schools in the Context of Educational Adequacy*. Stanford, CA: Institute for Research on Education Policy and Practice.
- Reardon, S. (2011). Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children. In R. Murnane & G. Duncan (Eds.) Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children. New York City, NY: Russell Sage Foundation Press.
- Rebell, M. A. (2006). "Adequacy Cost Studies: Perspectives on the State of the Art," *Education Finance and Policy*, 1(4), 465-483.
- Somers, M.A., Corrin, W., Sepanik, S., Salinger T., Levin, J. & Zmach, C. (2010). *The Enhanced Reading Opportunities Study Final Report: The Impact of Supplemental Literacy Courses for Struggling Ninth-Grade Readers* (NCEE 2010-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Stiefel, L., Rubenstein, R. & Schwartz, A. (1999). "Using Adjusted Performance Measures for Evaluating Resource Use," *Public Budgeting and Finance*, 19(3), 67-87.
- U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "National Public Education Financial Survey," 1990-91 through 2005-06.