PEABODY JOURNAL OF EDUCATION, 86: 28–57, 2011

ISSN: 0161-956X print / 1532-7930 online
DOI: 10.1080/0161956X.2011.539954

Routledge
Taylor & Francis Group

Are Education Cost Functions Ready for Prime Time? An Examination of Their Validity and Reliability

William Duncombe and John Yinger
The Maxwell School, Syracuse University

This article makes the case that cost functions are the best available methodology for ensuring consistency between a state's educational accountability system and its education finance system. Because they are based on historical data and well-known statistical methods, cost functions are a particularly flexible and low-cost way to forecast what each school district must spend to meet the standards in a state's accountability system. However, the application of cost functions to education must confront several challenges in both data collection and estimation methodology. This article describes the strengths and weaknesses of various ways to address these challenges and illustrates how the reliability and forecasting accuracy of cost functions can be tested using data for Missouri school districts.

INTRODUCTION

Every state faces the challenge of creating an education finance system that is consistent with its educational accountability program. The accountability program sets student performance targets that school districts are expected to meet, and, for consistency, the education finance system should provide every district with the resources it needs to meet these targets. This article explains how educational cost functions can be used to address this problem, explores the challenges facing researchers estimating educational cost functions, responds to recent criticism of the cost-function approach, provides estimates of education cost functions for the state of Missouri, and tests the accuracy of these estimates. Overall, we make the case that cost functions are the best available methodology for ensuring consistency between a state's educational accountability system and its education finance system.

EDUCATION COST FUNCTIONS

Production and cost functions are key microeconomic tools for understanding how various inputs are translated into a given output. A production function explains output as a function of input levels, and the related cost function explains a firm's costs as a function of its output and the input prices it faces. These tools have been extensively used to study education and other public

Correspondence should be sent to John Yinger, Center for Policy Research, Eggers Hall, Syracuse University, Syracuse, NY 13244. E-mail: jyinger@maxwell.syr.edu

services. At the school district level, for which a full list of inputs is difficult to observe, cost functions are the principal tool for studying educational production.[1] In this case, the key outputs are measures of student performance on state tests and graduation rates. The key input price is the salary a district must pay to attract teachers.

Any application of cost functions to education must address several challenging issues. First, it must select the outputs on which to focus. Second, any estimation of a cost equation must recognize the difference between costs and spending, which, as discussed next, is school-district efficiency. Third, educational cost depends not only on input prices but also on student characteristics, such as the poverty rate and the share of students with limited English proficiency. These characteristics are sometimes known as "fixed inputs," because they are outside a school district's control. Fourth, estimation of an education cost equation must recognize that both measures of student performance and teachers' salaries are influenced by school district decisions and are therefore, to use the statistical term, endogenous. To obtain unbiased results, studies that estimate cost functions must account for this endogeneity. Finally, a cost function study must select an appropriate functional form.

Selection of Outputs

For the problem on which this article focuses, namely, the consistency of accountability and finance systems, it is appropriate to select an output measure or measures that are central to a state's school accountability system. In most cases, these measures will indicate student performance on key state-administered tests (such as English, reading, or mathematics) and perhaps graduation rates. One approach would be to focus on the end point—high school test scores—and another would be to look at average scores over several grades.

Yet another approach, which has appeared both in the scholarly literature and in a few state accountability systems, is to focus on levels of student performance, not measured, say, by the share of students passing a state test but measured instead by the *change* in student performance over time, often referred to as a value-added measure. This approach is difficult to implement in a cost study, however, because a value-added approach requires test score information on the same cohort in different grades—information that is not generally available. Moreover, value-added measures provide noisy signals about student performance, particularly in small school districts (Kane & Staiger, 2002).

To understand why a value-added approach is difficult to implement, consider a standard Cobb-Douglas value-added production function in which student performance depends on the starting point (i.e., the previous year's performance) plus current year inputs:

$$S_T = S_0 \left(\prod_{t=1}^{T} A\,(L_t)^{\alpha}\,(K_t)^{\beta}\,e_t \right) \varepsilon = S_{T-1}\,(L_T)^{\alpha}\,(K_T)^{\beta}\,e_T, \tag{1}$$

[1]Education cost functions have been estimated for Arizona (Downes & Pogue, 1994), California (Duncombe, Lukemeyer, & Yinger, 2008; Imazeki, 2008), Illinois (Imazeki, 2001), New York (Duncombe et al., 2008; Duncombe & Yinger, 2000, 2005b), Texas (Gronberg, Jansen, Taylor, & Booker, 2004; Imazeki & Reschovsky, 2004a, 2004c), and Wisconsin (Reschovsky & Imazeki, 2001).

where S represents student performance, L is labor, K is capital, α and β are output elasticities for labor and capital, e_t is a random error associated with period t, and ε is a time-invariant error term for the district. (The subscript for the district is implicit.) The cost function associated with this production function, which can be estimated in log form, is as:[2]

$$C = a\,(S_T)^{1/(\alpha+\beta)}\,, \tag{2}$$

where

$$\alpha = (\alpha+\beta)\left(\frac{W^{\alpha}r^{\beta}}{AS_{T-1}e_T\alpha^{\alpha}\beta^{\beta}}\right)^{1/(\alpha+\beta)}, \tag{3}$$

where W is the price of labor, r is the price of capital, and $(\alpha + \beta)$ measures returns to quality scale (Duncombe & Yinger, 1993). With a value-added approach, therefore, the log of $S_{T\text{-}1}$ should be included in the estimating equation, with a predicted negative sign.

In this derivation, $S_{T\text{-}1}$ is the previous year's score for the cohort observed in year T. If S_T is an eighth-grade score in 2000, for example, then $S_{T\text{-}1}$ cannot be measured by either eighth-grade scores in 1999 or seventh-grade scores in 2000. Instead, it must be measured by the seventh-grade scores in 1999. The accumulated inputs for the previous cohort (eighth grade in 1999) and for the next cohort (seventh grade in 2000) have no impact on the current cohort (eighth grade in 2000) and are not an approximation for the term in the model. With a few exceptions, data for a correctly specified value-added cost model are not available, so this approach can be used in only a few states.[3] Implementation of "growth models" in a number of states to comply with No Child Left Behind (NCLB) may increase the feasibility of value-added cost functions over the next decade.

School-District Efficiency

The second issue at the heart of the problems addressed in this article is the difficulty of measuring school district efficiency. A cost equation indicates how much a school district would have to spend to achieve a given student performance level if it used the best available technology, that is, the best available teaching methods and management practices. "Cost" by this definition cannot be observed, however; instead, a researcher observes actual spending, which may not reflect the best available technology. To put it another way, spending may exceed cost because some districts deviate from the best available technology. Districts that deviate from the best available technology are defined to be "inefficient". No study can identify the determinants of a school district's "costs," therefore, without controlling for school-district efficiency.

More formally, education costs (C) depend on student performance (S); resource prices (W), such as teacher salaries; student enrollment (N); and student need measures (P); that is, $C = f(S, W, N, P)$. Now let e stand for school district efficiency in delivering S. Without loss of generality, we can set the value of e at 1.0 in an efficient district, so that e has a value between zero and one in a district that does not use current best practices. With this scaling, the cost/efficiency equation

[2]For a derivation, see Henderson and Quandt (1980, p. 85).

[3]In their studies of Texas, Imazeki and Reschovsky (2004a, 2004c) and Gronberg et al. (2004) used measures of value-added across 1 or 2 years.

is:

$$E = \frac{C}{e} = \frac{f(S, W, N, P)}{e}. \tag{4}$$

Moreover, a district that does not use best practices ($e < 1$) must spend more than an efficient district ($e = 1$) to achieve the same level of performance (S), all else equal.

Unfortunately, however, the concept of "efficiency" in the education context is widely misunderstood. To understand why, consider first the case of firms producing a single output (or multiple outputs produced with totally separate inputs). In this case, an inefficient firm is one that does not use the best available technology to produce the output. Inefficient practices include using an outmoded machine or sending executives on expensive vacations, and "inefficient" is a synonym for "wasteful."

When a firm produces multiple outputs with some sharing of inputs, however, the concept of "inefficiency" applies to the firm's production of each output, not to the firm as a whole, and "inefficiency" and "waste" may no longer be equivalent. Consider farms that grow corn and beans in the same fields, a practice used by Native Americans to give the beans a place to climb. One farm might use extra person power, and hence be inefficient, in growing corn because it takes great care not to trample the beans. This type of efficiency has nothing to do with waste; instead, it has to do with the farm's trade-off across the two outputs. A farm that is inefficient in this sense may, of course, also be wasteful by using an outmoded tractor, for example, but there is no logical connection between these two components of inefficiency.

Multiple outputs and input sharing are inherent features of production in public education; the same teachers and classrooms, supported by the same administrative services, provide many different outputs. These outputs include student performance on standardized tests, graduation rates, and student performance in art, music, athletics, and citizenship. In this setting, an analyst can ask if a school district is inefficient in producing English and math performance, but one must recognize that this type of inefficiency reflects both spending to promote other outputs and the use of outmoded techniques or other forms of waste. For example, a school district that is efficient in delivering student performance in mathematics might not be efficient in delivering student performance in English or art. Indeed, spending on art may have little impact on mathematics performance, so that it is a source of inefficiency in the production of mathematics. A good art or music program may, of course, contribute to students' general conceptual skills with some spillover to mathematics, but in most cases spending on these programs will not have as large an impact on mathematics scores as more spending on instruction in mathematics. No existing study has been able to separate these two types of inefficiency.

Misunderstanding about the concept of inefficiency in education leads to many incorrect statements in both the public debate and the scholarly literature. School districts with high spending levels and low performance are often called "inefficient," even if their high spending reflects the relatively high wages they must pay to attract teachers or the added costs of educating disadvantaged students (an issue discussed further in the next section). After accounting for labor market conditions and student characteristics, school districts that spend more than other districts with the same student performance on English and mathematics tests are sometimes called "wasteful." In fact, however, districts that spend more than other districts with the same English and math performance may simply be providing relatively high levels of student performance along other dimensions, such as graduation rates or music skills. Of course, these districts may

also be using outmoded techniques or being wasteful in some other way, but spending numbers alone cannot determine whether this is true.

Scholars have used a variety of methods to control for school-district efficiency in the estimation of a cost function. Because efficiency cannot, by definition, be directly observed, all of these methods control for efficiency indirectly. One method is to estimate the cost function with district fixed effects, which control for all district characteristics, including efficiency, that do not vary over time (Downes & Pogue, 1994). The limitations of this approach are that it requires panel data; that it cannot control for district efficiency that varies over time; and that, by removing all cross-section variation, it undermines a researcher's ability to estimate the impact on costs of S, W, N, and P.

The second approach begins with the estimation of a cost frontier based on the lowest observed spending for obtaining any given student performance, using a technique called data envelopment analysis (DEA). The next step is to calculate each district's deviation from this spending as an index of inefficiency and then to control for this measure in an estimated cost function (Duncombe, Ruggiero, & Yinger, 1996; Duncombe & Yinger, 2000; Reschovsky & Imazeki, 2001).[4] This approach has two key limitations. First, because the frontier is estimated with data on S and C, it relies on the functional form assumptions in DEA to identify the role of efficiency, assumptions that cannot be tested. Second, a standard DEA index of "inefficiency" reflects both cost and efficiency differences across districts. As a result, this approach may lead to underestimated coefficients of cost variables, such as student poverty, because a portion of the impact of these variables on costs may be captured by the estimated coefficient of the "inefficiency" index.

The third approach is to identify factors that have a conceptual link to efficiency and then to control for them in a cost function regression. A limitation of this approach is that these conceptual links cannot be directly tested. Nevertheless, a strong case can be made for the inclusion of two types of efficiency controls. First, some district characteristics might influence the incentives for voters to monitor school officials or for school officials to adopt best practices. For example, Imazeki and Reschovsky (2004a) controlled for efficiency using a measure of competition from other public schools, which might influence the behavior of school officials. Second, some district characteristics, such as median household income or tax price, might influence voters' demand for measures of school-district performance other than S. Because efficiency can only be defined relative to specific measures of school district performance, in this case S, any spending to obtain other measures of performance is, by definition, inefficient.[5] Income and tax prices are examples of variables that help control for this type of inefficiency (Duncombe & Yinger, 2000, 2005a, 2005b).

[4]Ruggiero (1998) showed how to separate cost and efficiency factors in DEA, but his approach requires far more observations than are available for any state because each district must be compared with other districts that have the same performance *and* the same cost factors. A multistage DEA-based approach has been used by McCarty and Yaisawarng (1993), Ray (1991), and Ruggiero (2001). Another approach is a stochastic frontier regression (Alexander et al., 2000; Gronberg et al., 2004). Ondrich and Ruggiero (2001) showed, however, that stochastic frontier regression produces the same results as an OLS regression except that the intercept has been shifted up to the frontier. As a result, this approach does not remove biases caused by omitting controls for efficiency.

[5]In a cost-function context, it is not possible to separate inefficiency associated with "wasteful" spending from inefficiency associated with spending on performance measures other than those included in S. It follows that a given school district could be deemed inefficient in providing one measure of student performance, say, math and English scores, and efficient in providing another, say, art and music.

CONTROLLING FOR STUDENT CHARACTERISTICS AND OTHER EXTERNAL COST FACTORS

Scholars have long recognized that the cost of education depends on many factors outside a school district's control. These factors include the wage environment, student enrollment, and the concentration of disadvantages among the student population. Duncombe and Yinger (2007b) have provided a detailed review of the scholarly literature on these factors.

A typical cost-function study controls for teacher wages (either measured by a private wage or treated as endogenous), student enrollment and student enrollment squared, a measure of poverty among the students (such as the percentage eligible for a free or reduced price lunch), and a measure of the share of students who speak English as a second language. In most studies, these variables are highly significant and have a large impact on education costs. A study of New York by Duncombe and Yinger (2005b) found, for example, that the cost of bringing a poor child up to a given education performance standard is more than twice as high as the cost for a nonpoor child.

Addressing Endogeneity

An estimated education cost function includes school outputs on the right side. Because these outputs are jointly determined with spending, however, they need to be treated as endogenous. This endogeneity is a great challenge for cost functions studies. The standard solution to an endogeneity problem is to estimate the model using instrumental variables. To be a valid instrument, a variable must help explain observed school outputs while not affecting school spending in any other way. Not surprisingly, variables with these two characteristics are difficult to find. Studies that use teacher wages as a cost factor also should treat this variable as endogenous.

Some early studies, including some of our own, used income and tax-price as instruments for school outputs. These variables satisfy one characteristic of a good instrument: they are demand variables that influence a community's choice of output levels. As it turns out, however, these variables do not satisfy the other main characteristic of a good instrument; because they affect the demand for many outputs, they show up as a determinant of efficiency in the cost function for any subset of outputs. In other words, they are invalid instruments because they are direct determinants of spending.

A more promising approach is to identify districts that are similar to a district and use their exogenous characteristics of instruments. The decisions of voters and school officials set their targets for good performance in part by observing comparison districts, but the exogenous characteristics of these districts do not influence the district's spending in any other way.[6] Because any cost analysis focuses on specific performance measures and other performance measures show up in inefficiency, this approach is most compelling when the performance measure in the cost

[6]Scholars face a trade-off with this approach. If the "similar" districts are literally neighbors, then their exogenous characteristics might be correlated with unobservable sorting factors shared with the district that defines the observation, thereby creating an endogeneity problem. If the "similar" districts are too far away geographically, however, they may not be part of the district's comparison group so that their traits have no explanatory power in the first stage regression. In this article we compromise between these two extremes by selecting the traits of districts in the same labor market area.

function either covers the most publicized measures in the state, such as those included in the state's accountability program. This approach is used (and, to the extent possible, tested) in the regressions presented next.

Functional Form

A final challenge is to select a functional form for the cost model. This form reflects underlying assumptions about the technology of production, such as the degree of substitution between inputs, economies of scale, and the interaction between school and nonschool factors. Most education cost studies have used a simple multiplicative cost function, which works well in practice but which imposes limits on both factor substitution and economies of scale.[7] By contrast, Gronberg et al. (2004) used a flexible cost function that does not impose significant restrictions on production technology. This approach adds many variables to the cost model, however, which makes it more difficult to identify cost effects with precision.[8]

CRITICISM BY COSTRELL, HANUSHEK, AND LOEB

The article by Costrell, Hanushek, and Loeb (2008; henceforth CHL) criticizes cost functions on several grounds. Many of these criticisms can also be found in Hanushek (2005) and Loeb (2007). This section presents our responses to the CHL analysis.

Lack of an Alternative

Our main disagreement with CHL is that they do not propose an alternative approach to solving the problem posed at the beginning of this article, namely, the compatibility between a state's education finance and school accountability systems. We do not claim that cost functions are perfect or that the challenges they pose are easily met, but we do claim that they are a logically compelling and empirically reasonable way to address this problem. We do not know of any other method that provides an alternative approach that is nearly as complete or compelling. No such alternative is identified by CHL. In fact, one of the authors in previous work (Hanushek, 2005) argued that decisions "on the right balance among different government programs and between public and private spending along with structuring the schools and their incentives is rightfully the province of the democratic appropriations process and not consultants hired by interested parties" (p. 2). In other words, Hanushek believes that better designs of school finance systems

[7]Most studies use a variant of the Cobb-Douglas function, which is multiplicative in form. The Cobb-Douglas function assumes that the elasticity of substitution between all inputs is equal to one and that the elasticity for economies of scale is constant at all levels of output.

[8]One of the most popular flexible cost functions used in empirical research is the translog cost function. A translog cost model includes squared terms for each input price and outcome, and adds interaction terms between all factor prices, and outcomes. Gronberg et al. (2004) also included a number of interaction terms between outcomes, teacher salaries, and nonschool factors. In all, they have more than 100 variables in their cost function for Texas compared to 18 variables in the Texas cost model estimated by Imazeki and Reschovsky (2004a).

to support adequacy will emerge from the political process, not from research on the cost of adequacy.

Complex empirical challenges appear in many types of policy problems. To prepare a budget, for example, every state needs a macroeconomic forecasting model and method for forecasting revenue. Economists certainly do not all agree on the best way to estimate a model of this type. Nevertheless, states all have procedures for creating and debating such a model—and for using its forecasts in their budgets. According to Voorhees (2004), states vary widely in methods and institutional arrangements for revenue forecasts. Some states set up a council of economic advisers, others rely on university experts, and a large number of states produce their forecasts using econometric methods. While state revenue forecasts by state agencies are influenced by politics (Bretschneider, Gorr, Grizzle, & Klay, 1989), they are informed by technical forecasts developed using a variety of methods (Kuo & Liang, 2004; Mocan & Azad, 1995; Shkurti, 1990).

The estimation of education costs, which is required to make education finance and accountability systems compatible, is also a fundamental problem for state policymakers, and cost functions provide the best available estimating method. States need to find ways to develop cost function studies, supplemented with appropriate other methods, which can inform the design of their state aid systems, just as they have found ways to incorporate macroeconomic forecasting models into their budgeting process. Rejecting cost functions because they are imperfect without providing an alternative is like rejecting the use of macroeconomic models for revenue forecasting and thereby turning the forecasting process back over to politicians.

We should make it clear that we have no trouble with the development of other methods for estimating what it would take to meet a state's accountability standards. Formal evaluations of various programs that look at costs and impact on performance, for example, would be helpful contributions. In fact, we have argued that states should take a more active role in initiating evaluations and helping school districts determine which programs are the most cost effective (Duncombe et al., 2008). But existing evidence about program impacts and costs is limited, and this type of evidence cannot replace a comprehensive cost measure obtained through a cost regression.[9]

We also do not believe, as CHL (2008) seem to suggest we do in their last paragraph (p. 221), that cost function studies indicate the amount of money that will "guarantee student success or at least the opportunity for student success" according to the state's standards. Cost functions can be used to estimate the spending required to meet a performance standard if a district uses best practices, but the district may not use best practices, and even if it does, a cost function only provides an estimate, not a guarantee. Nevertheless, we do believe that cost function studies are the best currently available method for estimating the cost of reaching any given performance target.

Interpretation of an Expenditure Equation

CHL also argue that a regression with expenditure as the dependent variable cannot be given a cost interpretation. On this point our disagreement with them is profound. Equation 1 in this

[9]On this point we are at least in agreement with Loeb (2007), who said that "the primary drawback of this method," which she called the "evidence-based" approach, "is that the research base is not strong enough to support it" (p. 13).

article shows that spending equals costs divided by efficiency. If efficiency is not correlated with the variables that determine cost, then a regression of spending on cost variables yields unbiased estimates of the cost parameters. If efficiency is correlated with these variables, then the methods discussed earlier are needed to insulate the coefficients of the cost variables from omitted variable bias. In either case, the cost variables have clear, legitimate cost interpretations. The coefficient of the poverty variable, for example, indicates the impact of an increase in poverty on the spending that is required to achieve any given level of student performance.

CHL give this coefficient, which they label β_3, a different interpretation. As they put it (where FRL stands for free and reduced price lunch, a common poverty variable),

> The estimated coefficient β_3 represents the additional spending, on average, among districts with higher percentages of FRL students, holding other variables constant. In essence it indicates what districts with different levels of poverty are spending. It does *not* represent the extra *cost* required to achieve any given performance level for FRL students. All a positive β_3 coefficient in equation (1) [not presented here] would reflect is a tendency of either the state or the district to spend more heavily when there is a greater proportion of students in poverty, while any similar tendency to spend less on poor students would yield a negative coefficient. This interpretation of β_3 holds regardless of whether extra spending is required to increase performance or is effective at doing so. (pp. 205–206)

This is simply not correct. CHL have forgotten the end of their own first sentence in this quotation. Performance is held constant. Moreover, in the estimating framework presented earlier, efficiency is held constant as well. So this coefficient does have a cost interpretation.

The cost equation presented by CHL is linear. This specification assumes that the impact of an increment in student poverty on the cost of education is constant, controlling for performance and efficiency. If this assumption is not correct, then the coefficient of a poverty variable estimated with their equation might give a biased estimate of the impact of poverty on costs. The problem, however, is with their linear specification, not with the cost function method. In fact, most studies estimate a multiplicative equation, which is much more reasonable on conceptual grounds; in this case, the coefficient of the poverty variable indicates the percentage increase in cost that accompanies an increase in the share of students from poor families. No published cost function study of which we are aware estimates a cost function that is linear.

CHL (2008) go on to argue that "deviations from average spending of comparable districts—are simply redefined as deviations from 'cost.' That is why the 'cost' estimates carry the logically incoherent implication that half the districts spend less than is necessary to achieve what they have achieved" (p. 212). We disagree. Cost based on best practices is defined as the minimum spending to achieve a given student performance, not as average spending. No cost study either claims or implies that "half the districts spend less than is necessary to achieve what they have achieved." Instead, cost studies control for performance and for differences in the cost environment in different school districts, and then assume that remaining variation in spending reflects variation in efficiency. With this approach, it is possible to construct an estimated efficiency index, which is 1.0 in the most efficient district (which is the one with the lowest spending after accounting for cost variables) and lower than 1.0 in other districts.

Of course, errors may arise in the efficiency index if the regression is misspecified, if the efficiency controls are inadequate, or if some cost factors have been omitted from the regression. But these are the types of problems that confront any estimating problem, not just this one. In

other words, this approach only yields an estimate of efficiency, but, contrary to the claims of CHL, it is a logically coherent one.

Production Functions as an Alternative to Cost Functions

Finally, CHL argue that cost functions are not credible because a "production function" estimated with the same data yields extremely different results.[10] Normally, a production function relates output, in this case student performance, to inputs. The great challenge facing this application of production functions is that it requires a comprehensive set of inputs, which is difficult to observe at the district level. In fact, we know of no study that has attempted to estimate a district-level production function with measures of teachers and teacher quality, administrators and administrator quality, maintenance staff, and supplies. To avoid measuring these inputs, CHL use student performance as the dependent variable and spending per pupil as the key explanatory variable. In effect, this approach assumes that spending is a proxy for the bundle of inputs that a school district selects.

This approach, namely, using spending as a measure of inputs, runs into two serious problems: extreme assumptions about production technology and measurement error.

Extreme Assumptions about Production Technology

To see what assumptions are required for this approach to make sense, consider a simple production function for student performance, S, with two inputs, K and L. The prices of these inputs are P_K and P_L, so spending, E, equals $(P_K K + P_L L)$. Hence the production function is:

$$S = f\{E\} = f\{P_K K + P_L L\}. \tag{5}$$

With this formulation, spending on K is assumed to have the same impact on S as spending on L. In other words, the marginal products of the two inputs are

$$\frac{\partial S}{\partial K} = f' P_K \quad \text{and} \quad \frac{\partial S}{\partial L} = f' P_L. \tag{6}$$

Now the slope of an isoquant is found from the total differential of the production function with dS set equal to zero. In symbols,

$$dS = 0 = \frac{\partial S}{\partial K} dK + \frac{\partial S}{\partial L} dL = f' P_K \, dK + f' P_L \, dL. \tag{7}$$

This equation leads directly to the slope of an isoquant:

$$\frac{dK}{dL} = -\frac{P_L}{P_K}. \tag{8}$$

[10]Imazeki (2008) made this argument, too. She claimed that "if the data and model were perfect (i.e., correctly specified with no unobservable variables or measurement error), the final cost estimates from the cost function and production function should be similar" (p. 102). This is simply not correct. As discussed next, spending and inputs are not the same thing and there is no reason at all to expect the final cost estimates from a misspecified production function to yield cost estimates that are similar to those from a cost function.

This is the first sign of trouble. Isoquants are supposed to depend only on technology, not on prices.

The slope of the government's iso-cost lines are found by totally differentiating the cost equation given earlier, namely, that $E = (P_K K + P_L L)$, and setting $dE = 0$.

$$dE = 0 = P_K\, dK + P_L\, dL \quad \text{or} \quad \frac{dK}{dL} = -\frac{P_L}{P_K} \tag{9}$$

This equation signals more trouble. The isoquants and the iso-cost lines have exactly the same slope; in other words, they are tangent everywhere! This means that, with this approach, all combinations of K and L are assumed to be equally efficient.

These problems extend to the case of more than two inputs. This approach requires the assumptions that spending of every input is equally productive and that any combination of inputs, including a value of zero for every input except one, is equally productive.

A cost function, in contrast, can be directly derived from standard production functions without imposing any such unrealistic assumptions on production technology. A cost function does require the measurement of input prices. If resource prices for some types of resources, such as materials and supplies, exhibit little variation across districts, however, they can be dropped from a cost-function estimation without imposing any constraints on the underlying production technology.

Measurement Error in the Key Explanatory Variable

As discussed earlier, efficiency, e, in Equation 1 cannot be directly observed, so it is not possible to separate cost and spending. This fact causes trouble for a production function, which is intended to describe the best available technology. The fact that the measured spending variable reflects waste as well as best-practices spending implies that there may be a large measurement error in this variable. It would not be surprising, therefore, to find an estimated impact of spending on performance that was close to zero.[11]

One cannot avoid this problem by saying that the regression picks up the average link between spending and performance. With errors in the key variable, the estimated coefficients are biased and inconsistent. One way to solve the problem would be to use an instrumental variables technique. To the best of our knowledge, this has never been attempted. Moreover, an instrument might be impossible to find. To isolate the cost component of spending, which is the effect one is trying to estimate, one needs an instrument that helps explain the link between inputs and spending. However, any instrument that meets this first test for a good instrument inevitably fails the second test, which requires no direct link between the instrument and the dependent variable, in this case district performance. An empirical implementation of Equation 5 also could include controls for variables associated with inefficiency, but this approach does not provide a general solution to the errors-in-variables problem. There may be a way out of this conundrum, but we do not know what it is, and it certainly has not been provided by people using this approach.

[11]In addition, we find it ironic that CHL, who are so critical of the treatment of efficiency in cost function studies, do not find fault with the production function approach, which faces much more severe challenges in dealing with efficiency.

Cost function studies may also face an errors-in-variables problem in measuring their key explanatory variable, namely, performance. This problem has an entirely different source, of course, namely, the random component in test-based measures of student performance. The difference is that this problem can be accounted for by the instrumental-variables procedures that are common in these studies. In the case of cost functions, instrumental variables that meet the above two tests can be identified. Moreover, policymakers are often interested in the performance index itself, which is not measured with error, instead of in the underlying student skills. Cost functions face no errors-in-variables problem for performance measures interpreted in this way.

Finally, cost function studies also face the possibility of omitted-variable bias if they do not control for efficiency. Because efficiency cannot be directly measured, no cost function study can definitively rule out this possibility. As discussed earlier, however, cost function studies have used a variety of different approaches to minimize this type of bias.

Conclusion

In short, it is not possible to estimate a production function at the school district level using spending as a measure of "inputs" without making the assumption that spending on every input is equally productive and that all input combinations are equally efficient. These are extreme, indeed, ridiculous assumptions, which the cost function approach does not have to make. Moreover, the production function approach magnifies the problem of accounting for efficiency because it automatically incorporates inefficiency into the definition of the key explanatory variable, namely spending per pupil. This creates a difficult errors-in-variables problem that has no obvious solution. To the best of our knowledge, the people who have used this approach have made no attempt to address this problem. In contrast, cost function studies do not employ an explanatory variable that reflects inefficiency by definition, and, as discussed earlier, they have used a range of methods to incorporate efficiency into the estimating equation.

COST FUNCTION ESTIMATES FOR MISSOURI

In the remainder of the article we illustrate how the validity and reliability of cost functions can be tested using cost function estimates for school districts in Missouri. The cost function estimates in this article are based on data from 2000 to 2005, which include three years prior to NCLB and three years after its passage. The implementation of NCLB during the middle of our time series poses a significant challenge to forecasting accuracy because pre-NCLB years will generally be used to fit the model and post-NCLB years will be used to test the forecasts. In this section we describe the data and measures, methodology, and cost function results.

Data and Measures

The cost function estimates provided in this article are based on a number of databases. Most of the data are produced by the Missouri Department of Elementary and Secondary Education (DESE). This section is organized by the type of variables used in the cost model, and summary

TABLE 1
Descriptive Statistics for Variables Used in Cost Model, Missouri School Districts (2005)

Variables	Average	Standard Deviation	Minimum	Maximum
Per pupil spending	$6,112	$1,513	$3,907	$16,434
Student Performance Measure	25.6	7.2	4.8	53.5
Cost variables				
Teacher salaries	$27,460	$3,290	$21,201	$40,055
Student poverty (percent subsidized lunch students)	46.1	16.0	4.7	95.1
Race-poverty interaction (Poverty variable multiplied by percent African American)	2.9	10.1	0.0	96.2
Percent special education students	16.7	4.	3.4	43.4
Enrollment	1607.7	3568.4	49.8	38682.5
K12 distircts (1 = yes; 0 = no)	0.86	0.34	0.00	1.00
Efficiency-related variables				
Fiscal capacity				
Per pupil property values	$61,631	$41,074	$20,399	$410,166
Per pupil income	$63,962	$35,112	$2,652	$394,463
Statel aid/income ratio	0.0594737	0.062326	0.00	1.144402
Other monitoring variables				
Percent of adults that are college educated (2000)	0.13	0.08	0.03	0.67
Percent of population 65 years or older (2000)	0.15	0.04	0.04	0.32
Percent of housing units that are owner occupied (2000)	0.77	0.08	0.42	0.95
Local tax share (median housing price/average property values)	1.23	0.49	0.25	2.91
Sample size		516		

Sources: Missouri Department of Elementary and Secondary Education; U. S. Census Bureau.

statistics for key variables in 2005 are reported in Table 1 for most school districts in Missouri (sample size in 2005 was 516).[12]

Per-Pupil Spending

The dependent variable used in the cost function is per-pupil operating spending. To broadly reflect resources used in the production of education in Missouri school districts, the spending measure is based on "current operating cost" (COC) developed by DESE. COC includes total instructional and support spending minus total capital outlay and several revenue categories (food service sales, state food service aid, federal food service aid, and receipts from other districts). In

[12]Data for approximately five school districts are not available. See Duncombe (2007) for a more in-depth discussion of the variables used in this cost model.

addition, transportation spending was removed since it is affected by factors, such as sparsity of the population, weather conditions, and road conditions, which are not likely to affect instructional spending. Our measure of per pupil spending in the average district in the state was approximately $6,100 in 2005.[13]

Student Performance Measure

The student performance measures used in the cost function are based on Missouri Assessment Program exams in math and communication arts administered by DESE. These are criterion-referenced exams in three grades for each subject area (Grades 4, 8, 10 for math, and Grades 3, 7, and 11 for communication arts). The information reported on these exams is the percent of students reaching certain thresholds in performance: (Step 1, progressing, near proficient, proficient, and advanced). The measure used in the cost model is the average of the share of students reaching proficiency for each exam. Although the majority of districts in Missouri serve the full range of grades, there are 73 districts serving kindergarten to eighth grades. We impute high school exam results for students attending K-8 districts.[14]

Teacher Salaries

Costs of providing education services vary across districts, in part, because of differences in prices that districts have to pay for resources, such as teachers. Some districts may have to pay more than other districts to attract similar teachers, because of a higher cost of living, fewer amenities in the area, and more difficult working conditions. Because teachers are the principal resource used to produce education, we include a measure of teacher salaries in the cost model.[15] To ensure that the teacher salary measure is comparable across districts, we use data on individual teachers to develop a salary measure that controls for differences in average education

[13]Special education services in St. Louis County and Pemiscot County are provided by special school districts serving these counties. Total spending, counts for special education students, and counts of students receiving subsidized lunch in these two special school districts are assigned to the regular school districts in each county using the share of county enrollment in each regular school district. For example, if a regular school district had 10% of St. Louis County enrollment, then it would be assigned 10% of the spending, 10% of special education students, and 10% of the subsidized lunch students in the special district serving St. Louis County.

[14]K8 students attending only one K12 district are assigned the high school proficiency rates for math and communication arts in this K12 district. In a few cases students in a K8 district attended two K12 districts for high school. To assign a high school performance measure to a K8 district, we constructed a weighted average of proficiency for high school math and communication arts exams, where the weight is based on relative enrollment. For example, assume students in the K8 district A attended K12 Districts B and C, where the enrollment in District B is 6,000, and enrollment in District C is 4,000. Then the high school performance assigned to District A is based on a weighted average of high school performance in Districts B and C, where the weights are 60% and 40%, respectively.

[15]Although other professional staff are key resources as well, variation in other professional salaries across districts are typically highly related to variation in teacher salaries (correlation over 0.75). We include only teacher salaries in the cost model.

and experience across districts.[16] We treat this salary variable as endogenous using private sector salaries and student enrollment in the labor market area as instruments.

Student Measures

A key variable in a cost model is the number of students served by the district. Student counts are used both directly as a variable in the cost model and to transform other variables into per-pupil measures. The student count measure used in this article is an average of the enrollment estimates in September and January. We use the average of these two enrollment counts to provide a measure of average enrollment for the year. Average enrollment provides a better estimate of the underlying enrollment of the district during the year and is less sensitive to unusual results associated with a single enrollment count. To account for the nonlinear relationship between enrollment and per-pupil spending we include enrollment and enrollment squared.[17]

One of the key factors affecting the cost of reaching performance targets is the number of students requiring additional assistance to be successful in school. Poverty has consistently been found to be negatively correlated with student performance (Ferguson & Ladd, 1996; Haveman & Wolfe, 1994). Poverty measures should accurately capture the percentage of a district's students living in low-income households. The most commonly used measure of poverty in education research is the share of students receiving free or reduced price lunch in a school, because this measure is produced annually.[18] To reduce the potential instability in this measure, especially in small districts, we use a two-year average of the subsidized lunch percentage.

The one change in specification for cost variables across the cost models we estimate (Models 1 and 2) is the inclusion of an interaction term between the share of subsidized lunch students and the share of African American students. To improve forecasting accuracy we tried several interactions of other variables with the subsidized lunch rate and as well as a quadratic specification.[19] The interaction with the share of African American students produced the most accurate forecasts of those we tried. This variable accounts for the possibility that the social disadvantages students bring to school are more severe among low-income African American students than among low-income White students.

---

[16]To control for variation in education and experience across districts, the natural logarithm of teacher salaries is regressed on the log of total experience, and an indicator variable for whether the teacher has a graduate degree. We use the regression to estimate average salaries for teachers in each district with the statewide average experience (between 0 and 5 years) and the statewide average percentage of teachers with a graduate degree.

[17]In other cost function estimates we have used a series of dummy variables to capture different enrollment size categories. Although this is a flexible way of measuring the relationship between spending and enrollment, the forecasting accuracy of models using these variables was slightly worse than using a quadratic relationship for enrollment. We also checked forecasting accuracy when a cubic term is included in the model and did not find it improved accuracy (and it was statistically insignificant).

[18]Another measure of child poverty is the child poverty rate produced by the Census Bureau every 10 years as part of the *Census of Population*. Although this measure is updated on a biennial basis, the updates are based on the decennial Census estimates, which implies that they may be quite inaccurate by the end of every decade. We found that the subsidized lunch rate in 2000 had a correlation of over 0.7 with the Census child poverty rate.

[19]We also tried interactions of the subsidized lunch rate with pupil density, enrollment, and percent college educated adults. Only the interaction with percentage African American and with pupil density were statistically significant. When a quadratic specification for subsidized lunch was tried, the coefficient on the squared term was not significant.

Students with special needs often require additional services and support, which can substantially increase school spending. Counts of special education students with individualized education programs are collected by DESE. To measure special education, we calculated total special education students as a share of enrollment. Another student characteristic that can affect the cost of bringing students up to a given performance level is a lack of English fluency, often called limited English proficiency (LEP). Unfortunately, the LEP data for Missouri do not appear to be accurate enough to use in this study.[20]

Efficiency-Related Measures

Costs are defined as the minimum spending of school resources required to provide students an opportunity to reach a given level of student performance. However, the dependent variable in the cost model is per-pupil spending. As discussed earlier, inefficiency in the cost function context can include both waste and a district's choice to focus on nontested subject areas (e.g., art, music, or athletics). Although it is not possible to measure efficiency directly, it is possible to control for it indirectly, and thereby to minimize the possibility of omitted variable bias in the cost coefficients.

Our approach is to include in the cost function variables that have been linked to efficiency in previous research. The literature on managerial efficiency in public bureaucracies suggests three broad factors that might be related to productive inefficiency: fiscal capacity, competition, and other factors affecting voter involvement in monitoring government (Leibenstein, 1966; Niskanen, 1971; Wyckoff, 1990). Research on New York school districts indicates that taxpayers in districts with high fiscal capacity, as measured by property wealth, income, and state aid, may have less incentive to put pressure on district officials to be efficient, or may be more apt to spend money on nontested subjects (Duncombe, Miner, & Ruggiero, 1997; Duncombe & Yinger, 2000). Property values are measured by assessed value for real property (residential, agricultural, and commercial and industrial) and personal property. The measure of income used in the analysis is adjusted gross income, which is provided by the Missouri Department of Revenue to DESE based on information from Missouri income tax returns.[21] We use a measure of state aid per pupil supporting basic operations divided by per-pupil income.[22] Previous studies have also found that voter's incentive and capacity to monitor operations in school districts and their demand for a broad set of school performance measures may differ depending on the education level of residents, the share of senior citizens in the population, the share of owner occupied housing, and

[20]Unlike subsidized lunch, there are no federal standards on how LEP students are measured, and typically no auditing process to assure that the data are accurate. In Missouri, student language data are collected in the Limited English Proficient Student Census (or English Language Learners Census) in October of each year. To evaluate the accuracy of the LEP data collected by Missouri, we compared this data to an alternative measure available in the *2000 Census of Population*—the percentage of students, who live in a household where English is not spoken well at home. The LEP measure supplied by school districts in Missouri is not highly correlated ($r = .30$) with the Census measure, suggesting that there are inconsistencies in how districts are classifying and reporting LEP students.

[21]The income data lags several years, so that the income data from the 2002 calendar year are used for the 2004–05 school year.

[22]The state aid measure includes minimum guarantee aid (basic formula) and aid for free and reduced price lunch students.

the share of school taxes paid by the typical voter (local tax share). Per-pupil property values are published by DESE, and the other variables are from the *2000 Census of Population*.

Cost Function Estimates

We estimate a constant-elasticity cost model using log linear 2SLS regression to account for the potential endogeneity of student performance and teacher salaries. To select instruments we use the average of exogenous variables related to student performance (percentage of Hispanic and percentage of African American students) and salaries (private sector salaries and enrollment size) in other districts in the same labor market area. The strength of instruments is tested using a weak instrument test based on partial F statistics (Bound, Jaeger, & Baker, 1995). A threshold of 10 is often recommended for acceptable instruments, although lower thresholds are acceptable in some cases (Stock & Yogo, 2005). For the models presented in Table 2, all of the partial F statistics are above 10. For the models in Table 3, the partial F statistics on teacher salaries are between 6 and 9.[23] To check for the possible effect of weak instruments on the accuracy of 2SLS estimates, we reran the cost function models using two of Fuller's k-class estimators ($a = 1$, $a = 4$), which are considered to be better estimators when instruments are weak (Murray, 2006). Because the results using the Fuller estimators are very similar to those with 2SLS, we report the 2SLS results. Second, we tested for overidentifying restrictions using a Hansen J-test (Baum, Schaffer, & Stillman, 2007), and found that we could not reject the null hypothesis that the instruments are uncorrelated with the error term at the 5% level.

Table 2 presents results of the basic cost function (Model 1) estimated for four periods—the full sample (2000–05) and several 3-year periods used to check forecasting accuracy (2000–02, 2001–03, 2002–2004). Table 3 presents the results for an alternative model (Model 2), which was specified based on analysis of forecasting errors associated with the basic model.

In general, per-pupil spending has the expected relationship with the independent variables in the cost function (Table 2) and the estimated coefficients are statistically significant from zero. The coefficient on the student performance measure ranges from .24 to .30 and indicates that a 1% increase in performance (as measured by a composite of proficiency rates for communication arts and math tests) is associated with a .24 to .3% increase in per pupil spending, controlling for the other variables in the cost function. Teachers' salaries are positively related to per-pupil spending; a 1% increase in teachers' salaries is associated with a 1.19 to 1.41% increase in per pupil expenditures holding other factors constant. The share of students in poverty and in special education is positively related to spending, and the poverty measure is significantly different from zero at the 5% level. The coefficients on the enrollment measures suggest significant economies of size with the cost minimizing enrollment level between 8,000 and 9,000 students.

[23]The model was estimated with xtivreg2 in STATA (Schaffer, 2005). Another weak instrument test involves comparing Kleibergen-Paap rk statistic to critical values established by Stock and Yogo (2005). Although this comparison is not technically correct given non-i.i.d errors, Baum, Schaffer, and Stillman (2007) argued that this is a reasonable approximation. The Kleibergen-Paap rk statistic is generally below the critical values established by Stock and Yogo (2005) for 10% relative bias. In other words, this test would suggest the potential for weak instruments in both Models 1 and 2.

TABLE 2
Cost Function Estimates for Different Years for Missouri School Districts (Model 1)

Variables	2000–2005	2000–2002	2001–2003	2002–2004
Intercept	−6.07793	−5.51048	−4.70453	−6.57974*
Student Performance measure[a]	0.30329*	0.28074*	0.24449*	0.30029*
Cost variables				
Teacher salaries[a]	1.41207**	1.27558**	1.19364**	1.41729**
Student poverty (percent subsidized lunch students)	0.00519**	0.00453**	0.00464**	0.00525**
Percent special education students	0.00188	0.00215	0.00202	0.00129
Enrollment[a]	−0.69238**	−0.60763**	−0.61445**	−0.68291**
Enrollment squared[a]	0.03830**	0.03331**	0.03380**	0.03782**
K12 distircts (1 = yes; 0 = no)	0.12652**	0.11427**	0.12468**	0.13081**
Efficiency-related variables				
Fiscal capacity				
Per pupil property values[a]	0.00203	0.01290	0.02656	0.00992
Per pupil income[a]	0.18073**	0.21961**	0.21876**	0.20799**
Statel aid/income ratio	1.27080**	1.94095**	1.88658**	1.57542**
Other monitoring variables				
Percent of adults that are college educated (2000)	0.26049	0.20842	0.26867	0.19820
Percent of population 65 years or older (2000)	−0.27082	−0.26372	−0.32357	−0.29627
Percent of housing units that are owner occupied (2000)	−0.19776**	−0.18511*	−0.17242*	−0.17325*
Local tax share (median housing price/average property values)[a]	−0.07631**	−0.07936**	−0.07424**	−0.07969**
2001	−0.00156	−0.00141		
2002	−0.01508	−0.01403	−0.01129	
2003	−0.03204		−0.02738*	−0.01824**
2004	−0.05741**			−0.04437**
2005	−0.03711			
Instrument tests:				
Partial F-statistics:				
Student Performance measure	25.93	26.04	24.28	21.90
Teacher salaries	15.53	16.74	13.23	12.13
Overidentification test (p-value)	0.09	0.07	0.03	0.11
Sample Size	3068	1520	1538	1546

Note: Estimated with linear 2SLS regression with the log of per pupil current operating cost (minus transportation spending) as the dependent variable. The performance measure and teacher salaries are treated as endogenous variables with instruments based on exogenous variables for other districts in the same labor market area and census district type (see text). Robust standard errors are used for hypothesis testing (controlling for clustering at the district level).
*indicates a coefficient significantly different from zero at the 5% level.
**indicates a coefficient significantly different from zero at the 10% level.
[a]Expressed as a natural logarithm.

Turning to the efficiency-related variables, the fiscal capacity measures have the expected positive relationship with per-pupil spending, which may indicate greater inefficiency or more demand for a broader array of educational services. The coefficients for state aid and per-pupil income are statistically significant from zero at conventional levels. A higher share of senior

TABLE 3
Cost Function Estimates for Different Years for Missouri School Districts (Model 2)

Variables	2000–2005	2000–2002	2001–2003	2002–2004
Intercept	−4.14437	−3.53610	−2.00864	−5.06573
Student Performance measure[a]	0.37845**	0.41108**	0.29732**	0.34819**
Cost variables				
Teacher salaries[a]	1.28134**	1.19568**	1.06783**	1.38386**
Student poverty (percent subsidized lunch students)	0.00407**	0.00389**	0.00345**	0.00424**
Poverty variable multiplied by percent African American	0.00004**	0.00006**	0.00005**	0.00004**
Percent special education students	0.00199*	0.00255*	0.00249**	0.00152
Enrollment[a]	−0.67788**	−0.64316**	−0.62510**	−0.68362**
Enrollment squared[a]	0.03743**	0.03540**	0.03458**	0.03768**
K12 distircts (1 = yes; 0 = no)	0.13825**	0.12209**	0.13833**	0.14873**
Efficiency-related variables				
Fiscal capacity				
Per pupil property values[a]	−0.00575	−0.02046	0.00327	−0.00880
Per pupil income[a]	0.12808**	0.14464**	0.12579**	0.13064**
State aid/income ratio–low[b]	−0.06880**	−0.06517**	−0.07083**	−0.06756**
Statel aid/income ratio–high[b]	0.19155**	0.20328**	0.16506**	0.19913**
Other monitoring variables				
Percent of adults that are college educated (2000)	−0.12248	−0.37317	−0.15399	−0.00828
Percent of adults college educated squared	0.61936	0.96208**	0.88190**	0.42890
Percent of population 65 years or older (2000)	−1.57258*	−1.36373	−1.66517*	−1.89970**
Percent of population 65 years or older squared	4.33299*	3.58088	4.51787*	5.28648**
Percent of housing units that are owner occupied (2000)	−0.20794*	−0.18841	−0.19921*	−0.18570
Local tax share (median housing price/average property values)[a]	−0.09174**	−0.10032**	−0.09421**	−0.09629**
Local tax share squared[a]	0.03805	0.04170	0.05111**	0.03997
2001	0.00196	0.00194		
2002	−0.00249	−0.00054	0.00100	
2003	−0.01484		−0.00906	−0.01323*
2004	−0.04103			−0.03731**
2005	−0.02453			
Instrument tests:				
Partial F-statistics:				
Student Performance measure	19.35	20.08	17.33	14.97
Teacher salaries	8.01	7.99	6.95	6.57
Overidentification test (p-value)	0.09	0.07	0.03	0.11
Sample Size	3068	1520	1538	1546

Note: Estimated with linear 2SLS regression with the log of per pupil current operating cost (minus transportation spending) as the dependent variable. The performance measure and teacher salaries are treated as endogenous variables with instruments based on exogenous variables for other districts in the same labor market area and census district type (see text). Robust standard errors are used for hypothesis testing (controlling for clustering at the district level).
*indicates a coefficient significantly different from zero at the 5% level.
**indicates a coefficient significantly different from zero at the 10% level.
[a]Expressed as a natural logarithm.
[b]Low ratio is a ratio below 0.05 and high ratio is above 0.2.

citizens in the population, a higher share of owner-occupied housing in the district, and a higher local tax share are associated with lower spending, with the latter two variables statistically significant. The share of college-educated adults is associated with higher spending but is not statistically significant.

EVALUATING RELIABILITY AND PREDICTIVE VALIDITY OF COST FUNCTIONS ESTIMATES

Cost functions, like production functions, are a general tool, which can be used for several types of empirical analyses. For example, cost functions have been used for program evaluation (Duncombe & Yinger, 2007a); policy analysis (Eom, Duncombe, & Yinger, 2007; Wang, Duncombe, & Yinger, 2010); and to forecast spending required to meet performance targets, such as those associated with NCLB (Duncombe et al., 2008; Imazeki & Reschovsky, 2004b, 2006). It is in this later role, as a tool for forecasting the costs of reaching adequate student performance, or so-called cost of adequacy, that cost functions have been recently criticized (Costrell et al., 2008; Hanushek, 2005). In this section, we describe the criteria and measures for evaluating the reliability and predictive validity of cost functions and illustrate their application using our cost function estimates for Missouri school districts (Duncombe, 2006).

Reliability

An important criteria for evaluating any measure is reliability, which can be defined as the "consistency and repeatability" of a measure (Trochim, 2001, p. 88). Reliability is typically estimated by comparing consistency of measures of the same phenomenon at different times (test–retest reliability), by different raters (interrater reliability), or using different items measuring the same phenomenon (internal consistency). Of the three types of reliability, test–retest reliability seems the most appropriate for forecasting. Test–retest reliability could be evaluated by comparing forecasts derived from cost functions estimated from different periods. Barring some major change in the education system (e.g., NCLB), we should expect that the regression coefficients on the key cost variables and the composite cost indices that were derived from these variables, should be highly related.

To examine test–retest reliability we estimate cost indices from the last three regression models presented in Tables 2 and 3. A cost index indicates how much more or less a particular district needs to spend compared to a district with average characteristics to provide its students an opportunity to reach the same performance level.[24] For example, a cost index of 120 indicates that a district will require 20% more spending than the average district to reach any student

[24]For each variable a district can influence (outcome measure and efficiency-related variables), the estimated coefficient of the cost model is multiplied by some constant, typically the state average for that variable. For each cost factor outside of district control, the estimated coefficient from the cost model is multiplied by the actual values for the district. The sum of the products for factors outside and within district control is used to predict costs in a district with average outcomes and efficiency. Predicted costs are also calculated for a hipothetical district, which has average values for all variables in the cost model. Predicted spending in each district is divided by spending in this average district (and multiplied by 100) to get the overall cost index.

TABLE 4
Comparison Between Cost Indices for Different Years for Missouri School Districts (Model 1) Averages by Census Region

	2000–2002	2001–2003	2002–2004
Model 1:			
Large central cities	112.3	110.0	114.0
Medium cities	92.0	90.6	91.9
Urban fringe of large cities	96.6	95.3	96.3
Urban fringe of medium cities	84.2	84.3	83.7
Large town	86.9	86.6	86.8
Small town	91.5	91.5	91.0
Rural metro	94.3	93.9	93.7
Rural non-metro	106.1	106.6	106.6
Model 2:			
Large central cities	150.2	138.0	140.6
Medium cities	88.0	88.1	89.0
Urban fringe of large cities	103.9	101.2	102.1
Urban fringe of medium cities	82.1	83.2	82.3
Large town	88.0	88.2	87.3
Small town	92.2	92.0	91.5
Rural metro	93.0	93.6	93.5
Rural non-metro	105.4	105.5	105.6

outcome level. The correlations between cost indices from the different cost models are very high (more than 0.99) suggesting high consistency in cost function estimates across different years. When average cost indices are compared across Census district type there is also a high degree of consistency (Table 4). As expected large central cities have the highest costs as well as sparsely populated rural districts (rural nonmetro). The largest difference across Models 1 and 2 are for large central cities and districts on the urban fringe of large cities. Several districts in these categories have very high concentrations of low-income and African American students.

Predictive Validity

The appropriate validity criteria to evaluate cost functions depend on the purpose for which they are used. If cost functions are used in program or policy evaluation, then the appropriate criterion is internal validity, which involves ruling out alternative explanations for the causal connection between the program and outcome (Barrow & Rouse, 2005). If cost functions are used for prediction purposes, then they should be evaluated on predictive validity. Predictive validity measures how well a measure predicts "something it should theoretically be able to predict" (Trochim, 2001, p. 68). Predictive validity is closely related to the concept of forecasting accuracy, which measures how well a forecast of a phenomenon fits the actual values. The predictive validity criteria focuses on the accuracy of the bottom-line spending estimate associated with a particular level of student performance, not on identifying "successful" education strategies.

In selecting a forecasting method, it is important to consider the timeframe and objectives of the forecast (Armstrong, 2001, 2005; Bretschneider, 1985). Forecasting the spending associated

with a particular performance level are typically made for the medium term (2–5 years), and the estimates need to be able to adjust to relatively large changes in key factors affecting student performance and the education environment (e.g., poverty, enrollment size, etc.). In addition, most of the variation used to estimate the cost function forecasting model is cross-sectional because long time-series are not generally available. Bretschneider labeled these types of forecasts as "prediction" forecasts (p. 6), because "they focus on specific policy issues, tend to be one-shot in nature, or . . . re-occur irregularly, and have a variable or arbitrary lead time" (p. 6). For prediction-type forecasts where the underlying theoretical model is known and data are cross-sectional, econometric models are typically recommended (Armstrong, 2005, 2006; Bretschneider, 1985).

Forecasts are principally evaluated on forecasting accuracy and bias (Armstrong, 1985, 2001; Makridakis, Wheelwright, & McGee, 1983). "The real test of a forecasting model is its out-of-sample forecasting ability" (Chatfield, 1996, p. 506). Out-of-sample tests of forecasting accuracy generally involve dividing the sample into two (Tashman, 2000). The first period (fit period) is used to estimate the forecasting model and the second period (test period) is used to test the accuracy of the model (Tashman, 2000, p. 438). Multiple tests of forecasting accuracy can be done depending on the timeframe of the forecast (how many years in the future) and the length of the time series. For example, if six years of data are available and three years of data are needed to estimate the forecasting model, then it is possible to estimate three one-year-out forecasts, two two-year-out forecasts, and one three-year-out forecast.[25] The accuracy measures from similar forecasts can be averaged and compared across forecasts of different time horizons. Normally, we would expect that the longer the time horizon, the less accurate the forecast.

To evaluate forecasting accuracy (and bias) the actual value for a test period is subtracted from the forecasted value and divided by the actual value to produce an estimate of the percent error (PE) of the forecast. Forecasting bias can be measured by taking the mean or median PE. If the average PE is primarily positive (negative), then the forecast is overestimating (underestimating) actual values. Forecasting accuracy can be assessed by taking the absolute value of the percent error for each observation and calculating the mean or median absolute PE.[26]

The accuracy of a forecasting model is commonly compared to a simple alternative model, usually called a naïve forecast. The most common version of a naïve forecast is to assume that the forecast this period is equal to the actual value for the previous period (no change). Thiel's *U*-statistic compares a forecast to this simple naïve forecast (Armstrong, 1985).[27] Since cost of adequacy forecasts have to be able to adjust with changes in student performance standards and key cost variables, a more appropriate naïve forecast would be based on a simple model

[25]Another decision is whether to keep the 1st year of the model fit period fixed (rolling origin evaluations) or to keep the period used to estimate the forecasting model fixed (rolling window evaluations). Rolling origin evaluations use the maximum data available to fit the model, whereas rolling window evaluations create a more equal comparison across different forecasts (Tashman, 2000). We use a rolling window forecast based on 3 years of data for this article.

[26]The average absolute PE has been criticized as having a bias favoring underestimates compared to overestimates (Armstrong, 1985). An alternative is to divide the error by the average of the forecast and actual values rather that the actual values. Armstrong (1985) called this the adjusted mean absolute PE.

[27]The Theil *U*-statistic is commonly measured as the square root of the ratio of the square of the PE divided by the square of the PE for the naïve forecast. The PE of the naïve forecast is actual value in this period minus the actual value for the last period divided by the actual value for this period. If the Theil *U*-statistic is below 1, then the forecast does better than the naïve forecast. See Collopy and Armstrong (2000) for a discussion of alternative versions of the Theil *U*-statistic.

TABLE 5
Estimates of Forecasting Bias and Accuracy (Difference Between Predicted and Actual as a Percentage of Actual) Missouri School Districts

	Bias (percent error)			Accuracy (absolute percent error)		
Distribution	1-year	2-year	3-year	1-year	2-year	3-year
Naïve Forecast:						
Mean	−2.5	0.0	−1.9	13.7	14.6	15.1
Median	−1.5	1.7	−0.2	11.2	12.4	12.4
Model 1:						
Mean	5.7	5.5	4.8	11.2	10.9	11.3
Median	5.3	5.3	4.2	9.0	8.8	9.4
1st percentile	−24.9	−22.5	−30.0	1.0	0.1	0.1
10th percentile	−9.6	−9.3	−11.9	2.7	1.8	1.8
25th percentile	−3.5	−3.3	−4.5	4.7	4.3	4.4
75th percentile	13.3	12.9	12.6	14.4	14.8	14.8
90th percentile	20.0	19.7	19.4	20.7	20.3	20.3
99th percentile	32.6	31.8	32.4	32.6	33.6	33.6
Model 2:						
Mean	0.6	0.7	−0.5	9.0	8.9	9.0
Median	0.2	0.7	−0.9	7.5	7.3	7.5
1st percentile	−25.8	−25.0	−28.9	0.9	0.7	0.2
10th percentile	−12.8	−13.1	−14.9	2.4	2.4	1.4
25th percentile	−6.3	−6.4	−8.0	4.1	4.2	3.5
75th percentile	8.0	7.8	7.3	12.4	12.2	12.8
90th percentile	14.5	14.0	14.4	17.8	17.5	18.7
99th percentile	26.4	25.8	25.4	28.7	28.1	30.8

Note: Average of forecasts by time horizon. For 1-year horizon, this is an average of 3 forecasts, for 2-year horizon an average of 2 forecasts, and for 3-year horizon this is just one forecast.

of spending regressed on the student performance index, teacher salaries, and a student poverty measure.

Tables 5 and 6 report estimates of forecasting bias and accuracy for several different forecasts for Model 1 (second panel). For 1-year-out forecasts we averaged three different forecasts, and for 2-year-out forecasts we averaged two forecasts. These forecasts were compared to those from a naïve model where per-pupil spending was regressed on student performance, teacher salaries, and the subsidized lunch rate (first panel).[28] The forecasts from Model 1 reported in Table 5 overestimated spending on average by 5% (first two lines of panel 2), which compares to a much lower level of bias using the naïve forecasts (panel 1). The forecasts from Model 1 are more accurate than the naïve forecasts with average error of 11% compared to 14% to 15% for the naïve forecasts. Approximately 90% of districts have errors of 20% or less using Model 1, which compares to errors of 30% using the naïve forecast (not reported).

Table 6 indicates that forecasts from Model 1 are particularly inaccurate and biased for two types of districts—large towns and medium cities. These districts tend to have low poverty

[28]Per-pupil spending, student performance and teacher salaries are logged. The fit of the model was weak (adjusted $R^2 = .064$).

TABLE 6
Estimates of Forecasting Bias and Accuracy (Difference Between Predicted and Actual as a Percentage of Actual) Missouri School Districts, by Type of District

	Bias (percent error)			Accuracy (absolute percent error)		
Distribution	1-year	2-year	3-year	1-year	2-year	3-year
Model 1:						
Large central cities	−5.8	−5.9	−7.5	5.8	5.9	7.5
Medium cities	19.3	19.0	18.2	19.3	19.0	18.2
Urban fringe of large cities	2.3	2.0	1.0	11.2	11.2	11.4
Urban fringe of medium cities	7.4	7.4	6.5	10.2	10.3	9.7
Large town	31.2	30.5	30.7	31.2	30.5	30.7
Small town	6.0	5.7	5.0	10.1	9.9	9.8
Rural metro	6.6	6.3	5.8	10.6	10.3	10.5
Rural non-metro	5.6	5.4	4.8	11.3	11.0	11.6
Model 2:						
Large central cities	0.1	0.4	−1.4	2.6	2.9	2.2
Medium cities	11.0	11.3	10.2	11.7	12.2	11.1
Urban fringe of large cities	−2.7	−2.9	−3.8	10.0	9.9	9.9
Urban fringe of medium cities	0.0	0.5	−1.2	7.2	7.3	6.6
Large town	20.8	20.7	21.5	20.8	20.7	21.5
Small town	0.1	0.2	−0.6	7.5	7.5	7.5
Rural metro	1.2	1.1	0.1	8.6	8.2	8.8
Rural non-metro	0.8	0.9	−0.4	9.2	9.1	9.3

Note: Average of forecasts by time horizon. For 1-year horizon, this is an average of 3 forecasts, for 2-year horizon an average of 2 forecasts, and for 3-year horizon this is just one forecast.

rates and costs as indicated by cost indices well below 100 (Table 4). They also have high income and property wealth and low levels of state aid. It is not surprising that these districts have student performance levels well above average. Using PE as the dependent variable it is possible to assess whether there are factors that are systematically associated with forecasting error. Table 7 presents the results when PE is regressed on the variables in the cost model. PE is positively (and significantly) associated with enrollment (over most enrollment ranges) and the fiscal capacity measures—per-pupil income, property wealth, and the state aid-income ratio. These results suggest that the cost function may be overestimating the efficiency effects from higher fiscal capacity and the higher costs associated with diseconomies of scale. PE is negatively related to the performance index, teacher salaries, subsidized lunch, and share of adults with a college education. Because several cost function variables are systematically related to forecasting error, modifications to the cost function could potentially reduce error. We tried several nonlinear specifications for the variables previously identified and interaction terms particularly between the cost variables. Although most of these changes to the model were not statistically significant, we have included in Model 2 (Table 3) changes that were statistically significant in some of the models and helped improve forecasting accuracy.

As indicated in Tables 5 and 6, Model 2 significantly improves prediction validity over Model 1 (and the naïve forecasts). For the typical district these forecasts have less than 1% bias and the accuracy of the forecasts have improved approximately 20% compared to Model 1 and

TABLE 7
Factors Associated With Forecasting Bias for Missouri School Districts by Type of District

Variables	Model 1	Model 2
Intercept	4.366**	4.979**
Student Performance measure[a]	−0.029**	−0.026**
Cost variables		
Teacher salaries[a]	−0.775**	−0.539**
Student poverty (percent subsidized lunch students)	−0.002**	−0.001**
Poverty variable multiplied by percent African American		0.000**
Percent special education students	−0.001	0.001
Enrollment[a]	0.213**	0.079**
Enrollment squared[a]	−0.011**	−0.003**
K12 distircts (1 = yes; 0 = no)	−0.013	−0.015
Efficiency-related variables		
Fiscal capacity		
Per pupil property values[a]	0.144**	0.003
Per pupil income[a]	0.104**	0.022
Statel aid/income ratio	2.733**	
State aid/income ratio–low[b]		−0.008
Statel aid/income ratio–high[b]		−0.042
Other monitoring variables		
Percent of adults that are college educated (2000)	−0.208**	−0.286**
Percent of adults college educated squared		0.653**
Percent of population 65 years or older (2000)	0.024	0.401
Percent of population 65 years or older squared		−0.982
Percent of housing units that are owner occupied (2000)	−0.043	−0.071
Local tax share (median housing price/average property values)[a]	0.016	−0.001
Local tax share squared[a]		0.018*
Adjusted R-square	0.372	0.096
Sample size	1548	1548

Note: Estimated with OLS. Bias is calculated using cost model for 2000 to 2002. The bias is calculated for 2003 to 2005.
*indicates a coefficient significantly different from zero at the 5% level.
**indicates a coefficient significantly different from zero at the 10% level.
[a]Expressed as a natural logarithm.
[b]Low ratio is a ratio below 0.05 and high ratio is above 0.2.

40% compared to the naïve forecast. Model 2 improves forecasting accuracy particularly for medium cities (40% improvement) and large towns (30%). When we regress PEs from Model 2 on variables in the cost model, the model fit (adjusted R^2) drops from 37% for Model 1 to 10% (Table 7). Although there is still much room for improvement in forecasting accuracy, we have illustrated how systematic analysis of forecasting errors can be used to improve forecasting accuracy and reduce bias.

So far we have focused on expenditure forecasts. We now turn to an investigation of the cost component of these forecasts. As discussed earlier, efficiency cannot be measured directly, but we can use conceptual arguments to separate cost and efficiency variables. Scholars largely agree, for example, that labor market conditions, enrollment, and concentrated poverty, all of which

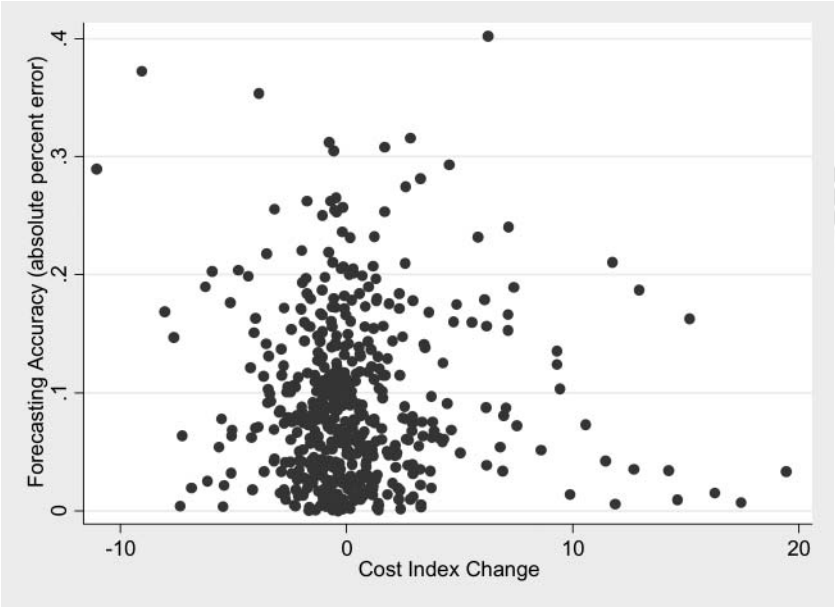

FIGURE 1 Comparison of changes in the cost index for a district with forecasting accuracy: 3-year-out forecast for Missouri school districts (2005).

are beyond the control of school officials, should be treated as cost variables. Based on this separation, we can then use our model to make forecasts of the impact of changes in a district's cost environment on its spending, accounting for change in student performance. If this cost portion of our model is working well, the forecasts for districts with large changes in cost factors should be as accurate as the forecasts for districts with small changes in cost factors.

To implement this "reasonableness" test, we estimated our cost model for 2000–02, calculated cost-index changes from 2002 to each of the following three years, and estimated forecasting accuracy for 2003 (1-year-out forecast), 2004 (2-year-out forecast), and 2005 (3-year-out forecast) for each district. If the regression is accurately capturing the role of costs, we would expect to see no systematic relationship between change in the cost index and forecasting accuracy. In other words, we would not expect districts with large changes in education costs to have lower forecasting accuracy than districts with small changes in costs. Figure 1, which presents a scatter plot of our results for the 3-year-out forecasts, shows that forecasting error does not depend on the magnitude of the cost change ($r = -.07$). Comparable figures for the 1-year-out and 2-year-out forecasts are similar. These results suggest that our estimated cost model provides not only fairly accurate forecasts of spending but also fairly accurate estimates of the cost component of spending, regardless of the magnitude of the change in costs.

CONCLUSIONS

School finance systems designed to support state accountability programs should, to be effective and fair, account for differences across school districts in the underlying cost of reaching

performance standards. Cost functions are one of several ways that have been developed to estimate these cost differentials. By utilizing historical data and strong statistical methods, cost functions are a particularly flexible and low-cost way to forecast required spending associated with a state's accountability system. However, the application of cost functions to education must confront several challenges in both data collection and estimation methodology. We describe approaches for addressing these challenges and highlight their strengths and weaknesses.

Recently cost functions have been criticized as a tool for estimating required spending associated with performance standards on several grounds (Costrell et al., 2008; Hanushek, 2005; Loeb, 2007). Critics argue that cost functions do not adequately control for efficiency differences across districts, so that their estimates cannot be given a cost interpretation. CHL point to the large differences in results between cost functions and production functions using spending to measure inputs as evidence of the inadequacy of cost functions. We argue that this criticism misses the mark in several ways. First, the cost variables in a cost function can be given a cost interpretation because the cost function controls for student performance and the omitted variable problem associated with inefficiency. Although much work remains to be done to improve efficiency controls in cost functions, it is inaccurate to suggest that no effort to account for efficiency has been made. The fact that production functions produce different estimates is also not an indicator of the weakness of cost functions but instead indicates the serious measurement problems that arise when spending is used as a composite measure of inputs in a production function. We demonstrate that to use spending in this way requires extreme, indeed, ridiculous assumptions about production technology.

Another fundamental problem with the criticisms of CHL is that they do not propose an alternative approach to forecasting the best-practice spending required to support student performance standards. Although we do not claim that the cost function approach provides perfectly accurately forecasts of required spending, we do not know of any other method that is as comprehensive and allows for low-cost testing of forecasting accuracy.

In the second half of the article we illustrate how the reliability and forecasting accuracy of cost functions can be tested using data for Missouri school districts. We show how these forecasts can be improved by examining the determinants of forecasting error and using this analysis to modify the cost model. We also provide some evidence to suggest that these forecasts accurately capture the cost component of spending.

The development of cost functions to support state school finance systems is analogous to the development of econometric models to forecast macroeconomic variables to support state revenue forecasts and budget development. States have taken steps over the years to improve these forecasts by combining forecasts of different types and developing institutional capacity within the state government to develop and evaluate forecasts. Although state and local governments have been much slower in developing forecasting models for key expenditure areas, the growing cost burdens associated with major federal programs, such as Medicaid and NCLB, suggest the expenditure forecasting may become more common. The question is not whether these types of forecasts should be done but how best to establish institutions to support the development and continued improvement of expenditure forecasts. Deschamps (2004) described the use of an independent forecasting agency and technical workgroups in the state of Washington to support the development of entitlement forecasts. We think similar steps should be taken by states to support the development and refinement of spending forecasts to support the state education accountability system.

AUTHOR BIOS

William Duncombe is Professor of Public Administration and the Associate Director, Education Finance and Accountability Program, Center for Policy Research, Syracuse University. His research specialties include educational cost analysis, school aid design, and budgeting and financial management. His research has appeared in journals in public administration, education policy, and public finance. He has worked on education finance projects in New York, California, Kansas, Maryland, Missouri, and Nebraska.

John Yinger is Trustee Professor of Public Administration and Economics at the Maxwell School, Syracuse University, and Director of the Education Finance and Accountability Program in Maxwell's Center for Policy Research. His principal research interests are education finance, discrimination in housing and mortgage markets, and urban economics, and he is the editor of *Helping Children Left Behind* (MIT Press, 2004).

REFERENCES

Alexander, C. D., Gronberg, T. J., Jansen, D. W., Keller, H., Taylor, L. L., & Treisman, P. U. (2000). *A study of uncontrollable variations in the costs of Texas public education* (Summary report prepared for the 77th Texas Legislature). Austin: Charles A. Dana Center, The University of Texas at Austin.

Armstrong, J. S. (1985). *Long-range forecasting: From crystal ball to computer*. New York, NY: Wiley & Sons.

Armstrong, J. S. (2001). Standards and practices for forecasting. In J. Armstrong (Ed.), *Principles of forecasting: Handbook for researchers and practitioners* (pp. 679–732). Norwell, MA: Kluwer Academic.

Armstrong, J. S. (2005). Forecasting principles and methods. *Foresight*, *1*, 29–35.

Armstrong, J. S. (2006). *Findings from evidence-based forecasting: Methods for reducing forecast error*. Unpublished paper, Wharton School, University of Pennsylvania, Philadelphia.

Barrow, L., & Rouse, C. (2005). *Causality, causality, causality: The view of education inputs and outputs from economics* (Federal Reserve Bank of Chicago Working Paper, WP 2005–15). Chicago, IL: Federal Reserve Bank of Chicago.

Baum, C., Schaffer, M., & Stillman, S. (2007). Enhanced routines for instrumental variables/GMM estimation and testing. *Stata Journal*, *7*, 465–506.

Bound, J., Jaeger, D. A., & Baker, R. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, *90*, 443–450.

Bretschneider, S. (1985). *Forecasting: Some new realities* (Occasional Paper No. 99). Syracuse, NY: Syracuse University, Metropolitan Studies Program.

Bretschneider, S., Gorr, W., Grizzle, G., & Klay, E. (1989). Political and organizational influences on the accuracy of forecasting state government revenue. *International Journal of Forecasting*, *5*, 307–319.

Chatfield, C. (1996). Model uncertainty and forecast accuracy. *Journal of Forecasting*, *15*, 495–508.

Collopy, F., & Armstrong, J.S. (2000). *Another error measure for selection of the best forecasting method: The unbiased absolute percentage error*. Unpublished paper. Available from http://www.forecastingprinciples.com/paperpdf/armstrong-unbiasedAPE.pdf

Costrell, R., Hanushek, E., & Loeb, S. (2008). What do cost functions tell us about the cost of an adequate education? *Peabody Journal of Education*, *83*, 198–223.

Deschamps, E. (2004). The impact of institutional change on forecast accuracy: A case study of budget forecasting in Washington State. *International Journal of Forecasting*, *20*, 647–657.

Downes, T., & Pogue, T. (1994). Adjusting school aid formulas for the higher cost of educating disadvantaged students. *National Tax Journal*, *67*, 89–110.

Duncombe, W. (2006). Responding to the charge of alchemy: Strategies for evaluating the reliability and validity of costing-out research. *Journal of Education Finance*, *32*, 137–169.

Duncombe, W. (2007, January). *Estimating the cost of meeting student performance standards in the St. Louis public schools*. Report prepared for the Board of Education for the City of St. Louis. Available from http://cpr.maxwell.syr.edu/efap/Costing_Out/Duncombe_technical%20report4.pdf

Duncombe, W., Lukemeyer, A., & Yinger, J. (2008). Dollars without sense: The mismatch between the No Child Left Behind Act accountability system and Title I funding. In R. D. Kahlenberg (Ed.), *Improving on No Child Left Behind: Getting education reform back on track* (pp. 19–102). New York, NY: The Century Foundation.

Duncombe, W., Miner, J., & Ruggiero, J. (1997). Empirical evaluation of bureaucratic models of inefficiency. *Public Choice*, *93*, 1–18.

Duncombe, W., Ruggiero, J., & Yinger, J. (1996). Alternative approaches to measuring the cost of education. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 327–356). Washington, DC: The Brookings Institution.

Duncombe, W., & Yinger, J. (1993). An analysis of returns to scale in public production, with an application to fire protection. *Journal of Public Economics*, *52*, 49–72.

Duncombe, W., & Yinger, J. (2000). Financing higher student performance standards: The case of New York State. *Economics of Education Review*, *19*, 363–386.

Duncombe, W., & Yinger, J. (2005a). *Estimating the Costs of meeting student performance outcomes adopted by the Kansas State Board of Education* (Report prepared for The Kansas Legislative Division of Post Audit). Syracuse, NY: Syracuse University.

Duncombe, W., & Yinger, J. (2005b). How much does a disadvantaged student cost? *Economics of Education Review*, *24*, 513–532.

Duncombe, W., & Yinger, J. (2007a). Does school district consolidation cut costs? *Education Finance and Policy*, *2*, 341–375.

Duncombe, W., & Yinger, J. (2007b). Measurement of cost differentials. In E. Fiske & H. F. Ladd (Eds.), *Handbook of education finance and policy* (pp. 238–256). Mahwah, NJ: Erlbaum.

Eom, T. H., Duncombe, W., & Yinger, J. (2007, January). *The unintended consequences of property tax relief: New York's STAR program* (Center for Policy Research Working Paper). Syracuse, NY: Syracuse University.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama Schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 265–298). Washington, DC: The Brookings Institution.

Gronberg, T., Jansen, W., Taylor, L. L., & Booker, K. (2004). *School outcomes and school costs: The cost function approach*. College Station, TX: Texas A&M University. Available from http://www.schoolfunding.info/states/tx/march4%20cost%20study.pdf

Hanushek, E. (2005, October). *The alchemy of 'costing out' and adequate education*. Paper presented at the Adequate Lawsuits: Their Growing Impact on American Education conference, Cambridge, MA.

Haveman, R., & Wolfe, B. (1994). *Succeeding generations: On the effects of investments in children*. New York, NY: Russell Sage.

Henderson, J., & Quandt, R. (1980). *Microeconomic theory: A mathematical approach* (3rd edition). New York, NY: McGraw-Hill.

Imazeki, J. (2001). *Grade-dependent costs of education: Evidence from Illinois* (Draft paper). San Diego, CA: San Diego State University.

Imazeki, J. (2008). Assessing the costs of adequacy in California public schools: A cost function approach. *Education Finance and Policy*, *3*, 90–108.

Imazeki, J., & Rechovsky, A. (2004a). *Estimating the costs of meeting the Texas Educational Accountability Standards* (Report for Texas Joint Select Committee on Public School Finance). Available from http://www.investintexasschools.org/schoolfinancelibrary/studies/files/2005/january/reschovsky_coststudy.doc

Imazeki, J., & Reschovsky, A. (2004b). Is No Child Left Behind an un (or under) funded federal mandate? Evidence from Texas. *National Tax Journal*, *57*, 571–588.

Imazeki, J., & Reschovsky, A. (2004c). School finance reform in Texas: A never ending story. In J. Yinger (Ed.), *Helping children left behind: State aid and the pursuit of educational equity* (pp. 251–281). Cambridge, MA: MIT Press.

Imazeki, J., & Reschovsky, A. (2006). Does No Child Left Behind place a fiscal burden on states? Evidence from Texas. *Education Finance and Policy*, *1*, 227–246.

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, *1*, 91–114.

Kuo, Y., & Liang, K. (2004). Human judgments in New York State sales and use forecasting. *Journal of Forecasting*, *23*, 297–314.

Leibenstein, H. (1966). Allocative efficiency vs. x-efficiency. *American Economic Review*, *56*, 392–415.

Loeb, S. (2007, August). *Difficulties of estimating the cost of achieving education standards* (Working Paper 23). Seattle: Center on Reinventing Public Education, Evans School of Public Affairs, University of Washington.

Makridakis, S., Wheelwright, S., & McGee, V. (1983). *Forecasting: Methods and applications*. New York, NY: Wiley & Sons.

McCarty, T. A., & Yaisawarng, S. (1993). Technical efficiency in New Jersey school districts. In H. O. Fried, C. A. Knox Lovell, & S. S. Schmidt (Eds.), *The measurement of productive efficiency: Techniques and applications* (pp. 271–287). New York, NY: Oxford University Press.

Mocan, H. N., & Azad, S. (1995). Accuracy and rationality of state general fund revenue forecasts: Evidence from panel data. *International Journal of Forecasting*, *11*, 417–427.

Murray, M. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, *20*, 111–132.

Niskanen, W. A. (1971). *Bureaucracy and representative government*. Chicago, IL: Aldine-Atherton.

Ondrich, J., & Ruggiero, J. (2001). Efficiency measurement in the Stochastic Frontier model. *European Journal of Operational Research*, *129*, 432–442.

Ray, S. C. (1991). Resource use in public schools: A study of Connecticut. *Management Science*, *37*, 1520–1628.

Reschovsky, A., & Imazeki, J. (2001). Achieving education adequacy through school finance reform. *Journal of Education Finance*, *26*, 373–396.

Ruggiero, J. (1998). Non-discretionary inputs in data envelopment analysis. *European Journal of Operational Research*, *111*, 461–468.

Ruggiero, J. (2001). Determining the base cost of education: An analysis of Ohio school districts. *Contemporary Economic Policy*, *19*, 268–279.

Schaffer, M. (2005). *XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models* (Statistical Software Components S456501). Boston, MA: Boston College Department of Economics.

Shkurti, W. (1990). A user's guide to state revenue forecasting. *Public Budgeting & Finance*, *10*, 79–94.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge, UK: Cambridge University Press.

Tashman, L. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting*, *16*, 437–450.

Trochim, W. (2001). *The research methods knowledge base* (2nd ed.). Cincinnati, OH: Atomic Dog.

Voorhees, W. R. (2004). More is better: Consensual forecasting and state revenue forecast error. *International Journal of Public Administration*, *27*, 651–671.

Wang, W., Duncombe, W., & Yinger, J. (2010). *School district responses to matching aid programs for capital facilities: A case study of New York's building aid program*. Unpublished paper, Syracuse University, Syracuse, NY.

Wyckoff, P. (1990). The simple analytics of slack-maximizing bureaucracy. *Public Choice*, *67*, 35–67.

Fonte: Peabody Journal of Education, v. 86, n. 1, p. 28-57, 2011. [Base de Dados]. Disponível em: <http://web.ebscohost.com>. Acesso em: 14 mar. 2011.