III.14 Measurement of Cost Differentials

William Duncombe, Phuong Nguyen-Hoang, and John Yinger

Introduction

The evaluation of education cost differentials across school districts has been an important topic in education finance research for decades (Fowler and Monk, 2001). Interest in this topic has grown in recent years with the emergence of adequacy as the primary standard in school finance litigation and the growth of state accountability systems that focus on student performance. These developments call attention to education cost differentials, which arise when some districts must spend more than others to obtain the same performance. Despite a growing literature on the link between research and policy on this topic, however, existing state aid formulas usually contain ad hoc cost adjustments that fall far short of the across-district cost differences estimated by scholars. The objective of this chapter is to synthesize the research literature on education cost differences across school districts and to discuss the implications of this literature for state education aid formulas. The material in this chapter complements the discussions of equity and adequacy in Chapters III.12 (Baker and Green) and III.13 (Downes and Stiefel).

The term “cost” in economics refers to the minimum spending required to produce a given level of output. Applied to education, cost represents the minimum spending required to bring students in a district up to a given performance level. Education costs can be affected by three categories of factors, each of which is outside of school district control: (1) geographic differences in resource prices; (2) district size; and (3) the special needs of some students. In this
chapter, we address the principal methods for estimating the cost impacts of each of these factors. As discussed below, these impacts need not be the same, of course, for every measure of student performance or in every state.

While states commonly adjust their basic operating aid programs for differences in the capacity of school districts to raise revenue, typically measured by property wealth or income, few states systematically adjust these programs for cost differences across districts. Instead, cost adjustments tend to be confined to ancillary aid programs or to be non-existent. In her recent survey, Verstegen (2011) finds that sixteen states provide no extra funding for low-income students and thirteen provide no extra funding for English Language Learners. This limitation is important because the success of a basic operating aid formula in providing the funds needed for an adequate education in each district, however defined, is linked to the accuracy of the cost adjustment. An aid formula cannot achieve an adequacy standard defined as a minimum level of student performance without adjustments for student needs, district size, and geographic variation in resource prices. Because cost adjustments may vary with the measure(s) of student performance or with conditions in a state, standard cost adjustments are not available. Instead, each state needs to estimate its own cost adjustments or else settle for approximate adjustments based on studies in similar states. Estimating cost adjustments is a challenging enterprise that requires clear judgments about state educational objectives, good data, and technical expertise. These challenges are the subject of this chapter.

The focus here is on cost differentials across school districts, not across individual schools. School districts are the primary budget decision making units taxing power and budget authority also lie with district officials. Accordingly, state school finance systems are focused on distributing education aid to school districts, not schools, in most cases. As a result, this chapter
does not cover recent research about inequity in the distribution of resources across schools within large urban school districts (Baker, 2012) or about weighted-student formulas designed to address this inequity (Baker, 2009; Ladd, 2008). Nevertheless, the measures of cost differentials due to student needs discussed in this chapter may also be appropriate for intra-district funding formulas.

The chapter is organized roughly in line with the major cost factors. We begin by discussing briefly the one method that can produce estimates for all three types of cost factors—education cost functions. We then turn to looking at other methods for estimating geographic resource price differences, the cost effects of enrollment size, and the cost impacts of various student characteristics. Each section describes the most frequently used methods, discusses their strengths and weaknesses, and provides key references for more detailed information.

**Education Cost Functions**

To estimate the relationship between spending, student performance, and other important characteristics of school districts, many education researchers employ one of the key tools of production theory in microeconomics, namely, a cost function. Cost is defined as the minimum spending required to reach a given level of student performance using current best practices. A district using best practices is said to be efficient. Cost cannot be directly observed, however, so cost functions are estimated using district spending (usually operating spending per pupil) as the dependent variable. Spending is higher than cost when school districts are inefficient, that is, when they deviate from current best practices. As a result, cost functions need to be estimated with controls for school district efficiency.

More formally, education costs, $C$, depend on (1) student performance ($S$); (2) resource prices ($W$), such as teacher salaries; (3) enrollment size ($N$); and (4) student need measures ($P$),
which are discussed in detail below; 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 it has a value between zero and one in a district that does not use current best practices. We then can write the cost/efficiency equation that scholars estimate as:

\[
E = \frac{C}{e} = \frac{f(S, W, N, P)}{e}
\]  

(1)

This formulation makes it clear that 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.

Equation (1) has been widely used in various forms in empirical work because it addresses many fundamental questions of interest to scholars and policy makers. For example, a cost function measures how much a given change in teacher salaries, district enrollment, or student needs affects the cost of achieving a particular level of student performance at a given level of efficiency. The cost function methodology has been refined over the last few decades, and cost function studies have been undertaken for several states.¹

In order to estimate equation (1) using multiple regression analysis, researchers must address several methodological challenges. The first challenge is to identify a performance objective and find data to measure it (\( S \)). One common approach, for example, is to select the performance measure or measures that are most central to a state’s school accountability system, which typically include student performance (measured by average scores or passing/proficiency rates) on state-administered tests in English, reading or mathematics, and perhaps graduation rates. In addition, these measures of student performance are determined simultaneously with district spending, so they need to be treated as endogenous when equation (1) is estimated.²

Together with a few state accountability systems, cost function studies, e.g., Imazeki and Reschovsky (2006) and Gronberg et al. (2011), may focus not on levels of student performance...
but instead on the change in student performance over time, often referred to as a value-added measure. As shown in Duncombe and Yinger (2011b), 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 and Staiger, 2002).

A second methodological challenge is to control for school district efficiency ($e$). The problem is that efficiency cannot be directly observed, so a researcher must select a method to control for efficiency indirectly. Several approaches, each with limitations, have appeared in the literature. One approach 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 and Pogue, 1994). The limitations of this approach are 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 of $S$, $W$, $N$, and $P$ on costs.

Another efficiency approach is to estimate a cost frontier based on the lowest observed district spending for obtaining any given student performance, 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 and Yinger, 2000; Reschovsky and Imazeki, 2003). A limitation of this approach is that this 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 final efficiency approach in the literature 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 (2008) and Imazeki and Reschovsky (2004a; 2006) control 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 $S$, any spending to obtain other performance measures is, by definition, inefficient. Income and tax price are examples of variables that help control for this type of inefficiency (Baker, 2011; Duncombe and Yinger, 2011a; 2011b; Nguyen-Hoang and Yinger, 2014).

Because efficiency is defined by the selected performance measures, the division between costs and efficiency may be quite different in different states. States with different performance measures in their accountability systems—and hence in their cost analyses—should not expect to have the same variation in costs across districts. The factors influencing efficiency could also vary from state to state even if the performance measures are the same.

A third 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 non-school 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. By contrast, Gronberg et al. (2011) use 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.\(^7\)

The cost function approach has been criticized by Costrell, Hanushek, and Loeb (2008) and Hanushek 2007). These scholars do not offer a better alternative to cost function estimations, however, and their criticisms are addressed in Baker (2006), Duncombe (2006), and Duncombe and Yinger (2011b). Despite empirical challenges involved in estimating cost functions, therefore, they have some clear advantages over other methods of estimating cost differentials. First, they use actual historical data with appropriate statistical procedures to separate the impact of factors outside and within district control on the cost of reaching student performance levels. Second, they can provide measures of overall cost differentials across districts as well measures of individual cost factors (resource prices, enrollment, and student needs) that can be used in state aid formulas. Some scholars have criticized cost functions on the grounds that their technical complexity makes them difficult for state policy makers to understand (Guthrie and Rothstein, 1999).\(^8\) One of the objectives of this chapter is to explain the intuition behind cost functions to help make them more accessible to policy makers. After all, complex statistical procedures are accepted in some policy arenas, such as revenue forecasting and program evaluation, and we see no reason why they could not become accepted in the design of state education aid formulas.

In the following sections we describe the use of cost functions and other methods to estimate cost differentials for resource prices, economies of size, and student needs.

**Geographic Variation in Resource Prices**

The impact of geographic variation in the prices of goods and services on the purchasing power of school districts has been recognized for decades (Brazer and Anderson, 1975;
Chambers, 1978). Ten states incorporate geographic cost of education indices (GCEIs) into their school funding formulas (Lofgren, 2007), and the National Center for Education Statistics (NCES) has sponsored the development of GCEI for all school districts in the country using two different methods (Taylor and Fowler, 2006).

Controlling for the compensation a district must offer to attract personnel of a given quality is particularly important for accurate cost estimation, because personnel compensation (salaries plus benefits) made up over half of current spending in the average school district in 2011 (Dixon, 2013, Table 6). In this section, we discuss the reasons for variation in resource prices and review the four most common approaches for estimating a GCEI. Each of these approaches attempts to measure the extent to which the cost of personnel varies across districts based on factors outside of districts’ control—not on variation in districts’ generosity. Variation in the price of inputs other than personnel has been largely ignored in the literature. This remains an important topic for future research.

**Reasons for Geographic Variation in Resource Prices**

The prices school districts must pay for resources can differ across school districts for several reasons: (1) cost-of-living; (2) labor market conditions; (3) local amenities; and (4) working conditions for employees. The higher the cost-of-living in an area, defined as the resources required to purchase a standard bundle of goods and services, the more school districts in the area must pay to attract employees of a given quality. Local labor market conditions can also affect the salaries districts are required to pay. If an area’s unemployment rate for professionals is high relative to the rest of a state, for example, then teachers and school administrators in that area may have relatively limited choices of alternative jobs and thus be more apt to accept school district offers with lower salaries and benefits. This type of situation
may not persist in the long run, however, if teachers are mobile.

School employees, like other employees, may also be willing to sacrifice some compensation to have ready access, or proximity, to amenities including natural sites (coastline, lakes, mountains, or parks), transportation (highways, airports, or railway), cultural events, and other state or local public services. Finally, districts may trade off spending on factors related to working conditions against increased teacher compensation. Specifically, the salary required to attract instructional and administrative personnel may depend on the working conditions in the school district, which reflect both school policies (e.g., school size, class size, professional development spending, availability of instructional materials, school leadership and culture) and factors outside district control (e.g., student characteristics such as the socio-economic background). Research has found that teacher mobility may negatively affect student achievement (Ronfeldt, Loeb, and Wyckoff, 2013) and may be influenced by working conditions (Boyd et al., 2011; Isenberg, 2010; Hanushek, Kain, and Rivkin, 2004; Ondrich, Pas, and Yinger, 2008).

Cost-of-Living (COL) Index

The cost-of-living (COL) approach estimates price differences for a “market-basket” of goods and services across geographic areas (Duncombe and Goldhaber, 2003). For each factor in the market basket, price data is collected by geographic area and a market basket is identified using data on consumer expenditure patterns. The final COL index is the spending required to purchase the market basket in each location relative to the state or national average. The use of a COL index as an education cost adjustment is based on the assumption that teachers compare real wages across districts, not nominal wages. This assumption implies that a high-COL district cannot attract the same quality teachers as a low-COL district without paying higher real wages.
The principal strengths of the cost-of-living approach are its conceptual simplicity and the fact that COL indices are based on private sector prices outside of district control (McMahon, 1996). This simplicity comes at a price, however. Even if a COL index accurately captures variation across locations in consumer prices and in the wages required to attract teachers, school personnel do not necessarily shop or live where they work. In addition, COL indexes do not capture variation across districts in working conditions and local amenities, which can affect the compensation required to attract equal quality teachers. Moreover, COL data at the school district level are surprisingly difficult to obtain; existing national and state-level COL indexes provide no insight into within-state COL variation (Nelson, 1991; McMahon, 1996). As a result, some states, including Colorado, Florida, Oregon, and Wyoming, have developed their own regional COL indices and incorporated them into their school aid calculations (Bureau of Economic and Business Research, 2013; Corona Insights, 2012; Rothstein and Smith, 1997; Wyoming Department of Administration & Information, 1999).

**Competitive Wage Index (CWI)**

Another approach to estimating a GCEI is to use information on variation in private sector salaries (or full compensation), particularly for occupations similar to teaching, typically professional, managerial, or technical occupations. Some states, including Ohio, Massachusetts, New York, and Tennessee, have used measures of average private wages as cost adjustments in their education aid formulas. One disadvantage of this approach is that it assumes that private employees in these occupations are comparable on experience, education, and demographic factors across geographic areas.

A more appealing approach is to use detailed individual-level data on private employees to construct a private wage index that controls for employee characteristics. Taylor (2011) has
applied this approach to Texas and Washington school districts. Combing data from the
{
\textit{Occupational Employment Survey} (OES) published by the U.S. Bureau of Labor Statistics and
the Census, Taylor and Fowler (2006) developed a CWI for all school districts in the country by
regressing salaries of college graduates on demographic characteristics (age, gender, ethnicity,
education, and hours worked), occupational categories, and indicator variables for labor market
areas. The regression results are used to “predict the wages that a nationally representative
person would earn in each labor market area” (Taylor and Fowler, 2006, p. 9). The CWI is
obtained by dividing the predicted wage by the state or national average wage. This CWI can be
updated for other years by using the OES to estimate changes in wages across years by
occupation and labor market area.

The comparable wage methodology is straightforward, and a carefully constructed CWI
should capture the impact of cost-of-living, local amenities, and labor market conditions on the
salary a district must pay to attract teachers of a given quality. The principal drawback to this
methodology is that average private sector salaries are not likely to reflect differences in working
conditions for teachers across districts. For example, private sector salaries in professional
occupations are not likely to reflect the demographics of the student body in a district, the age
and condition of school buildings, the extent of overcrowding in classrooms, and so on, which
could be very important to the job choices of teachers.

\textbf{Hedonic Teacher Cost Index (TCI)}

A GCEI can also be estimated by separating the impact on teacher (or other employee)
compensation of factors within and without district control using statistical methods—and then
to determine costs based only on external factors. What sets this approach apart from the others
is that, to the extent possible with available data, it directly accounts for the effects of school
working conditions on the salaries required to attract teachers to a district. It has been conducted for several states, including Alaska, Maryland, New York, and Tennessee (Chambers et al., 2004; Chambers, Taylor, and Robinson, 2003; Duncombe and Goldhaber, 2003). However, only Texas and Wyoming, presently use a hedonic-based cost adjustment in their aid formulas (Taylor, 2010; 2011).

The hedonic salary approach involves estimating a multiple regression model in which employee salary (or salary plus fringe benefits) is regressed on teacher characteristics, working conditions under district control (such as school or class sizes), and factors outside district control that are related to cost of living, labor market conditions, local amenities, and school working conditions. Characteristics of teachers typically include education, experience, gender, race, type of assignment, and certification status. Some studies include other measures associated with teacher quality, such as certification, test score performance, and ranking of the college a teacher attended (Duncombe and Goldhaber, 2003). Amenity variables typically include distance to a central city, and crime rates, and working-conditions variables include district enrollment and student characteristics (e.g., disability, language proficiency, and poverty).

Hedonic models are typically estimated with individual teacher-level data using standard multiple regression methods. To construct a personnel cost index, the coefficients for discretionary factors are multiplied by the state average value for that factor, while coefficients for external (i.e., non-discretionary) factors are multiplied by actual values for that district. The sum of these terms is the predicted salary required to attract an employee with average characteristics to a particular district. Because they have the most complete controls, hedonic salary models are likely to produce the most accurate estimate of the salary required to attract teachers with given characteristics to work in a district. However, even hedonic estimates face
several difficult challenges (Goldhaber, Destler, and Player, 2010).

Perhaps the most difficult challenge is to fully control for teacher quality. Teacher characteristics included in existing studies capture several important dimensions of teacher quality, but these characteristics predict only a small share of variation in teacher quality as directly measured from teachers’ impacts on the test scores of their students (Harris and Sass, 2011; Rivkin, Hanushek, and Kain, 2005). Moreover, teacher quality is likely to be negatively correlated with concentrated student disadvantage, so imperfect controls for teacher quality will bias the coefficients of the student disadvantage variables toward zero. As a result, hedonic studies may systematically understate the impact of concentrated student disadvantage on the compensation a district must pay to attract teachers of a given quality.

In addition, actual teacher salaries may not correspond in all districts to the minimum salaries required to attract teachers with certain characteristics into the district. Some districts could be overly generous or particularly inept in bargaining, for example. Differences between actual salaries and minimum required salaries are signs of district inefficiency, and they could lead to biased results in hedonic salary models if the (unobserved) factors that lead to inefficiency are correlated with the explanatory variables in the model.

Another challenge is that readily available COL measures may reflect discretionary district decisions. For example, housing prices often account for most of the variation in private prices across geographic areas but they may partially reflect differences in perceived education quality across districts (Nguyen-Hoang and Yinger, 2011). Using MSA level housing prices reduces this endogeneity (Duncombe and Goldhaber, 2003). Some hedonic studies have used unimproved agricultural land as a COL measure to avoid the potential endogeneity of housing prices; however, agricultural land in central cities or inner ring suburbs often does not exist, and
has to be imputed (Chambers et al., 2004). Private sector salaries can serve as a proxy for cost-of-living, labor market conditions, and some amenities, but are likely to be influenced by housing prices (and education quality) as well.

Several studies have attempted to address potential biases in hedonic salary models. Teacher fixed effects models have been estimated to control for unobserved teacher quality differences (Chambers et al., 2003). To account for the possibility of omitted compensation or working condition variables, some studies have included an estimate of the turnover rate in the model (Chambers et al., 2004; Duncombe and Goldhaber, 2003). Moreover, a few hedonic studies include variables to control for school district efficiency (Duncombe and Goldhaber, 2003). Finally, Boyd et al. (2013) develop a “two-sided matching model,” which separates school and teacher preferences and leads to the conclusion that “teachers prefer schools that are closer to home, have fewer poor students, and, for white teachers, have fewer minority students” (p. 106).

The TCI calculated from hedonic salary models tend to display relatively little variation, because most of the variation in teacher salaries is explained by key factors in the models, usually education and experience, and because information on other determinants of teacher quality and on working conditions is incomplete. The limited impact of working conditions on hedonic TCI runs counter to recent research on teacher labor markets, which as discussed earlier, finds that teacher mobility is influenced by the characteristics of the students they teach. More research is needed to resolve this apparent contradiction.

**Teacher Cost Indices from Cost Functions**

Resource prices, particularly teacher salaries, are key variables in education cost functions. The coefficient on the teacher salary variable indicates the increase in costs required to
maintain student performance levels when teacher salaries increase (holding other variables in the model constant). Using this coefficient and measures of teacher salaries by district it is possible to construct a teacher cost index (relative to the state average), which reflects variation in teacher salaries weighted by the impact of teacher salaries on spending.

Two different types of salary measures have been used in education cost functions: (1) private sector wage indices, such as a CWI (Reschovsky and Imazeki, 2001; Imazeki and Reschovsky, 2004a); and (2) actual teacher salaries for teachers with similar education and experience levels. Recognizing that teacher salaries can be set simultaneously with spending levels in the annual budget process, studies using actual teacher salaries often treat them as endogenous variables in estimating the cost function (Duncombe and Yinger, 2000; 2011a; 2011b; Nguyen-Hoang and Yinger, 2014). In these studies, the teacher salary index may be based on actual salary or on salary predicted on the basis of factors outside a district’s control.

A teacher cost index derived from a cost function is similar in some respects to a CWI and should capture variation in employee compensation due to differences in cost-of-living, labor market conditions, and amenities across school districts. Because student characteristics are also included in the cost model, the teacher cost index is not likely to reflect the impact of working condition differences across school districts on the wages required to attract teachers. This impact may appear, however, in the estimated coefficients of these student characteristics; if so, it will appear when teacher cost indexes and pupil weights (discussed below) are combined.

The strength of this approach is that it produces a teacher cost index that both reflects variation in key factors affecting teacher salaries and is weighted by the impact of teacher salaries on spending. The accuracy of this approach depends, however, on the quality of the cost-model controls for student disadvantage and school-district efficiency.
Enrollment Size and Education Costs

The 90-percent drop in the number of school districts in the United States since 1938 represents one of the most dramatic changes in education governance and management in the twentieth century. While the pace of school district consolidation has slowed since the early 1970s, some states still provide financial incentives to encourage school district consolidation (Verstegen and Jordan, 2007) amid potentially strong opposition from local citizens (Weldon, 2012). However, operating aid formulas in a number of states compensate districts for small size or sparsity, thereby discouraging consolidation (Baker and Duncombe, 2004). In this section we briefly review the reasons for and the evidence on the relationship between costs and district size before discussing methods to estimate the cost effects of size.

Reasons Costs May Vary with Enrollment

Economies of scale are said to exist when the cost per unit declines as the number of units goes up. In education, the focus has been on economies of size, which refer to a decline in per-pupil expenditure with an increase in district enrollment, controlling for other cost factors. Several explanations have been offered for economies of size in education. First, some district services, such as central administration, are relatively fixed in the sense that the same central administrative staff may be able to serve a significant range of enrollment without a degradation of service. Economies of size might exist if larger school districts are able to employ more specialized labor, such as science or math teachers, which could improve the quality of instruction at no additional cost. Furthermore, teachers may be more productive in a large school district because they can draw on the experience of many colleagues. In addition, large districts may be able to negotiate relatively low prices for bulk purchases of supplies and equipment, or
use their monopsony power to negotiate lower wages for their employees.

The existence of economies of size in education has been challenged for several reasons. First, some studies claim that the potential cost savings from consolidation are seldom realized because districts seldom lay off staff, salaries are often leveled-up across the merging districts, and transportation costs actually increase (Hanley, 2007). Second, market concentration from consolidation may increase cost inefficiency, and thus reduce potential cost savings (Gronberg et al., 2013). Third, large school districts tend to have large schools, which, according to some studies, lead to lower student performance (Kuziemko, 2006), by hurting staff morale, student motivation and involvement in school, and parental involvement (Cotton, 1996; Howley, 1994).

**Evidence on Economies of Size in Education**

A large literature on economies of size in education has emerged over the last four decades. Since this literature has been covered in depth in existing literature reviews (Andrews, Duncombe, and Yinger, 2002; Fox, 1981), we only summarize the main findings. The vast majority of evidence on economies of size has come from the estimation of education cost functions. The early evidence on economies of size found sizeable economies with the cost-minimizing size for an urban district as high as 30,000 students (Fox, 1981). Recent cost function research, which have addressed a number of methodological limitations with early studies (Andrews et al., 2001), has also found that there may be sizeable economies of size in education, but that most of the cost savings from an increase in district enrollment are exhausted once enrollment levels of 2,000 to 4,000 pupils are reached. A few formal evaluations of the effects of school district consolidation on costs have been conducted. Zimmer, DeBoer and Hirth (2009) find evidence of potential cost efficiencies from consolidations of school districts with enrollment below 2,000. Evaluating school district consolidations in New York from 1985 to
1997, Duncombe and Yinger (2007) find that doubling enrollment cuts operating costs per pupil by 61.7 percent for a 300-pupil district and by 49.6 percent for a 1,500-pupil district. Contrary to these studies, Gordon and Knight (2008) find that consolidation in Iowa “did not change pupil-teacher ratios, enrollments, or dropout rates,” but they cannot control for student test scores.

**Methods for Estimating Cost Effects of Enrollment Size**

A common method used by states to construct scale adjustments is to estimate average district costs by enrollment class and then compare the average cost in a class to average costs in relatively large districts. Kansas used this strategy, for example, to develop “low enrollment weights,” which are used in calculating operating aid (Duncombe and Johnston, 2004). The problem with this approach is that it does not consider factors other than enrollment size or sparsity, such as student performance, resource prices, topography, and student needs, each of which might affect spending differences across districts.

The cost function method provides the most direct way to determine the relationship between enrollment and costs. By controlling for student performance, resource prices, student needs, and efficiency, cost functions have the potential for isolating the effects of enrollment size on cost differences. The key decisions in estimating economies of size in a cost function is selecting measures of student counts, and the functional form of the relationship between cost and enrollment.

Student counts used in aid formulas generally are of three types: 1) enrollment, which is the count of all students at one point in time (usually the fall); 2) average daily membership (ADM), which is an estimate of the average enrollment over the course of the year; and 3) average daily attendance, which measures the average number of students actually attending school. States adopt either one or a hybrid form of these three. In general, the difference between
these student counts is quite small except in the large cities where attendance rates are often lower than enrollment.

The existence of economies of size implies a negative relationship between per pupil spending and enrollment at least over some range of enrollment. However, it is likely that the rate of decline in per pupil spending occurs more quickly at low enrollment levels than at higher enrollment levels because of relatively fixed costs, such as central administration. Several different functions have been used to account for the possible non-linear relationship between enrollment and per pupil cost. The most common approach is to use a quadratic function (the natural log of enrollment and its square) to model the relationship. Quadratic functions allow for the relationship between enrollment and per pupil costs to go from negative to positive (that is to be U-shaped), and cost function studies have found diseconomies of scale as well as economies of scale (Reschovsky and Imazeki, 2001; Imazeki and Reschovsky, 2004b). In states with a few high enrollment school districts (e.g., New York) the quadratic function can lead to estimates of large diseconomies of scale; some studies have used cubic functions to reduce the effects of these large districts (Duncombe and Yinger, 2000). To allow for a more flexible relationship between enrollment and per pupil costs, several cost function studies have used enrollment “groupings” (such as 0 to 300, 301 to 2000 etc.) instead of a quadratic function (Duncombe and Yinger, 2005). Flexible cost functions, such as translog functions, provide another alternative for specifying the enrollment-spending relationship by including both a quadratic term and a number of interaction terms between enrollment and other variables in the cost model (Gronberg et al., 2004).

Professional judgment studies can also be used to estimate the effects of size on costs. (See Chapter III.13 by Downes and Stiefel for a more detailed discussion of this approach.) In
professional judgment studies, panels of education professionals are asked to estimate the resources required to produce a particular set of student performance results. Panels are typically asked to do estimates for prototypical schools or districts with different characteristics, such as enrollment size or poverty rates (Baker, 2005). The estimates for the prototypical districts can then be extrapolated to districts of different sizes to develop an economies-of-size estimate for all districts in a state. Using the results of professional judgment studies in several states, Baker (2005) found that the shape of the per-pupil cost curve relative to enrollment was very similar to that found in cost function studies.

**Student Disadvantage and Education Costs**

Extensive research on the determinants of student success in school indicates that peer characteristics, family composition, parental education and employment status, and neighborhood characteristics can significantly affect student success (Ginther and Pollak, 2004; Harris, 2010). In addition, student characteristics can affect the mobility decisions of teachers (Hanushek, Kain, and Rivkin, 2004)—and hence both the quality of teachers and the costs of teacher recruitment and training. Moreover, districts with a high concentration of students living in poverty or with limited English proficiency face much greater challenges than other districts in helping their students reach academic proficiency. In this section, we discuss the types of student characteristics considered in the literature, the methods available for estimating the additional costs required to bring disadvantaged students to a given performance level, and how states have accounted for student disadvantage in their aid formulas.

**Measures of At-Risk Students**

The term “at-risk” implies that a student is at a higher risk of failing to meet educational
objectives than are other students because of characteristics of the student or of his or her family or peers. The most widely used measure of “risk” or disadvantage is poverty. One key measure of poverty for education cost studies is the child poverty rate, defined as the share of a district’s school-age population (5 to 17 years old) living in a poor household. The Census Bureau also provides intercensal estimates of child poverty. An alternative poverty measure more commonly used in education research is the share of students that qualify for a free or reduced-price school lunch as part of National School Lunch Program administered by the U.S. Department of Agriculture. This measure has the advantage over the census poverty measure in that it is updated annually, but it is based in part on decisions by families to apply for participation in the program and on decisions by school districts to offer and promote this service. However, the percent of elementary students eligible for a free lunch is highly correlated with the child poverty rate, as reported in Duncombe and Yinger (2005) or Duncombe and Goldhaber (2003).

Other measures of student “risk” available in the decennial census include the share of children living with a single mother and the share of children living with a single mother who has an income below the poverty line and is not a high school graduate. States may also collect information on “Title 1” students (those eligible for Title 1 services). Students with limited English proficiency (LEP), also called English Language Learners (ELLs), may face significant challenges succeeding in school (see Chapter VII.33 by Rumberger and Gándara). Many states collect information on students who qualify for bilingual education programs or students that have been identified as needing language assistance. Unfortunately, however, there is no standard definition of LEP across states, and the LEP data in some states are of questionable accuracy. An alternative measure is available from the Census, which collects information on the number of children ages 5 to 17 who live in households where English is spoken “not well” or
“not at all” or of children that are living in households that are “linguistically isolated.”\textsuperscript{16}

One limitation of the data in most states is that they do not identify how many students fall into multiple at-risk categories. The data do not identify, for example, how many students are both from poor families and ELLs. Estimated cost differences across districts might be different if controls for these types of overlap were included. Moreover, the patterns of overlap might differ across states so that pupil weights estimated in one state do not apply elsewhere, even if the states’ performance measures are the same. More research on these overlaps is needed.

Students with disabilities or special needs generally require more resources than other students to reach the same student performance standards. To account for these extra costs, many states incorporate pupil weights or other adjustments for special needs students in their school aid formulas. In Chapter VII.32, Harr, Parrish, and Chambers provide a detailed discussion of state policies, including aid formulas, to account for students with special needs.

\textbf{Methods for Estimating Additional Costs to Educate Disadvantaged Students}

Cost functions provide a direct way to estimate the impact of student disadvantage on the cost of education, holding student performance constant. To be specific, the coefficients on the variables measuring student disadvantages can be used to calculate pupil weights for each type of disadvantage (Duncombe and Yinger, 2005). A weight of 1.0 indicates for a given type of student disadvantage that it costs 100 percent more to bring a student in that category up to given performance standards than the cost for a student without disadvantage. Poverty weights estimated from cost functions vary within a state due to differences in methodology and across states because of differences in performance measures and other issues discussed earlier. One recent study (Duncombe and Yinger, 2005) estimates a range of poverty weights from 1.1 to 2.1 for New York, depending on the method; another (Duncombe et al., 2008) finds an average
poverty weight of 0.55 for Kansas and 0.64 for Missouri. In the case of Missouri, but not Kansas, this weight is higher in central cities than in rural areas. In addition, estimated weights for students in poverty are between 0.23 and 0.31 for Texas (Gronberg et al., 2004) and between 0.3 (Imazeki, 2008) and 0.563 (Duncombe and Yinger, 2011b) for California.\textsuperscript{17}

Another approach for estimating the higher costs required to support at-risk students is to use professional judgment panels. These panels can be asked to estimate the required resources needed to reach student performance standards for schools with different levels of poverty (or LEP shares). The differential in costs across these prototype schools can be used to develop rough estimates of pupil weights by student type. Pupil weights from professional judgment panels are based on the judgments of professional educators, and may be sensitive to the instructions given to the panels and to the ability of panel participants to identify the extra programs that would be required to bring at-risk students up to the specified performance standard (Rose, Sonstelie, and Richardson, 2004).

Baker (2006) compared several professional judgment studies and cost function studies for the same state and found that pupil weights produced from professional judgment studies are generally lower than weights produced from cost function studies. In Kansas and New York, for example, poverty weights calculated using the results from professional judgment studies (Augenblick et al., 2002; Chambers et al., 2004) are half those calculated in cost function studies (Duncombe and Yinger, 2005).\textsuperscript{18} One exception is the professional judgment study in Maryland done by Augenblick and Myers (2001), which estimated pupil weights of 1.0 or higher for poverty and LEP.

**How States Adjust for Student Disadvantage**

Almost all state governments have some type of aid program that provides additional
funds to districts with a relatively high concentration of at-risk students. Almost all states have one way or another to reimburse additional costs to educate students with special needs. Verstegen (2011) finds from her survey that twenty-one states incorporate the costs of special education students into operating aid programs through the use of pupil weights, and ten other states reimburse school districts (a certain percentage of) actual eligible costs incurred in providing educational services for these students. The weighted-pupil approach is also used to adjust the basic operating aid formula for poverty in 18 states and for students with limited English proficiency in 13 states (Verstegen, 2011). For poverty students, the lowest weight of 0.05 is used in Mississippi while the highest weight of 1.5938 appears in Georgia (Verstegen, 2011). Most states use eligibility for the National School Lunch Program as their measure of poverty, but some states use Census poverty estimates or federal welfare eligibility.

With the except of Maryland, which recently implemented an aid formula with a poverty weight from a study that used the professional-judgment method, no state uses statistically estimated pupil weights in its formula. Nevertheless, pupil weights have been estimated for many states and have been considered in many policy debates.

The cost adjustments discussed in this chapter are designed to apply to a school district’s entire operating budget. Because they almost always represent a small share of a district’s budget, categorical aid programs for at-risk students are therefore unlikely, by themselves, to provide needy districts with the funds they need to meet an adequacy standard. Most states have not yet recognized that both incomplete cost adjustments in the operating aid formula and full cost adjustments applied to small categorical aid programs are incompatible with an adequacy objective.

Conclusions and Directions for Future Research
We find a broad consensus among scholars that the cost of achieving any given level of student performance is higher in some districts than in others because of (1) differences in the compensation needed to attract school personnel, (2) differences in enrollment size, and (3) differences in the concentration of disadvantaged students or those with special educational needs. We do not find a consensus, however, on the magnitude of these cost differences or on the best methods for estimating them. Instead, we observe an active literature with many different approaches to estimating costs and a lively debate about the strengths and weaknesses of each approach.

From our perspective, the core of this topic is the estimation of education cost models. Although scholars disagree about the details, these models are now widely used and have influenced the debate about state aid formulas in many states. The most difficult issues that arise in estimating these models are how to select performance measures and then how to control for the associated school-district efficiency. No consensus on the best approach has yet emerged and far more work on these topics is needed. Questions of variable selection and functional form also deserve more attention. Because they play such a critical role in state education aid formulas, pupil weights should continue to be a focus of this research.

A second focus of the literature has been on estimating teacher cost indexes. This topic also has important links to policy because a state aid formula cannot provide the resources needed to reach any student performance target without accounting for teacher costs. As we have shown, scholars have addressed this topic using a wide range of approaches with different strengths and weaknesses. The hedonic wage approach is the most appealing conceptually, but it also requires data that are often not available, and more research developing and comparing all the approaches would be valuable.
Finally, we are struck by both the clear link between cost estimation and the objectives of most state education aid formulas and the need for more work to make cost studies accessible to policy makers. Complex statistical procedures are accepted in some policy arenas, such as the revenue forecasts used in budgeting, but are only beginning to be accepted in the design of state education aid formulas. Because careful estimates of cost differentials can help policy makers achieve their educational objectives, we believe that further efforts to make these estimates accessible would be valuable.
References


Imazeki, Jennifer, and Andrew Reschovsky. 2003. “Financing Adequate Education in Rural


Smith, Peter. 1997. “Model Misspecification in Data Envelopment Analysis.” Annals of

———. 2011. “Updating the Wyoming Hedonic Wage Index”. Submitted to the Wyoming Joint
Appropriations Committee and the Wyoming Joint Education Committee.

Cost Adjustment”. U.S. Department of Education.

Verstegen, Deborah A. 2011. “Public Education Finance Systems in the United States and
Funding Policies for Populations with Special Educational Needs.” Education Policy

State Survey of School Finance Policies.”

State Governments.

Index: Policies and Procedures”. Wyoming Department of Administration & Information.
http://eadiv.state.wy.us/wcli/policies.pdf.

Yinger, John, and William D. Duncombe. 2004. “Amicus Curiae Brief of John Yinger and
William Duncombe”. Submitted to Supreme Court of the State of New York, County of
New York, September 17.

Zimmer, Timothy, Larry DeBoer, and Marilyn Hirth. 2009. “Examining Economies of Scale in
School Consolidation: Assessment of Indiana School Districts.” Journal of Education

Endnotes
*Acknowledgements: Bill Duncombe passed away in May 2013, but he was a full partner in the
previous version of this chapter and his contributions remain central to this version. We would
like to thank Leanna Stiefel, Helen Ladd, and Ted Fiske for their excellent comments and
suggestions. We are fully responsible for any errors and omissions.

1 Published cost function studies have been conducted for Arizona (Downes and Pogue, 1994),
California (Duncombe and Yinger, 2011; Imazeki, 2008), Indiana (Zimmer et al., 2009), Kansas
(Chakraborty and Poggio, 2008; Duncombe et al., 2008), Kentucky and Maine (Lee, 2010),
Massachusetts (Nguyen-Hoang and Yinger, 2014), Missouri (Baker, 2011; Duncombe et al.,
2 More formally, equation (1) needs to be estimated with two-stage least squares regression, which requires “instruments.” These instruments are variables that influence $S$ but do not influence $E$ directly. Recent studies use instruments that measure the determinants of the demand for $S$, such as socio-economic characteristics, in comparable school districts, which form a point of comparison for voters and school officials (Duncombe et al., 2008; Duncombe and Yinger, 2005; 2011a; 2011b). Several studies, e.g. Imazeki and Reschovsky (2004a) and Reschovsky and Imazeki (2003), use a district’s own income and tax-price as instruments. These instruments are not legitimate, however, because, as shown below, they are determinants of efficiency and therefore influence $E$ directly. As in Nguyen-Hoang (2012), and Ross and Nguyen-Hoang (2013), an annual performance rank index, which is derived by dividing districts’ performance into equal percentile-based thirds and assigning the thirds ranks of one, two and three. could be potential instruments for performance. Overall, the choice of instruments is an important topic for future research.

3 The main cost frontier method to estimate efficiency is Data Envelopment Analysis (DEA) as in Chakraborty and Poggio (2008), Flavin et al. (2012), and Ruggiero (1998, 2001, 2007). This DEA approach is not without weaknesses: inability (as a result of no standard errors) to test hypotheses (Ray, 2004), and potential bias from measurement error and model misspecification (Smith, 1997). Another approach is a stochastic frontier regression (SFR) (Gronberg et al., 2011). Ondrich and Ruggiero (2001) show, however, that SFR and OLS yield the same results.
except for the intercept. In other words, SFR is only appropriate if efficiency is uncorrelated with any explanatory variables.

4 Ruggiero (1998) shows how to separate efficiency and cost measures in DEA, but his method requires more extensive data than are usually available.

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 some types of student performance, say math and English scores, and efficient in providing others, say art and music.

6 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.

7 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 include a number of interaction terms between outcomes, teacher salaries, and non-school factors. In all, they have over 100 variables in their cost function for Texas compared to 18 variables in the Texas cost model estimated by Imazeki and Reschovsky (2004a).

8 Downes (2004) argues that rejecting the cost function method because it is not easy to understand “means that other methodologies should be used in place of the cost function methodology, even if the cost function methodology is theoretically sound and is most likely to generate valid estimates of the spending levels needed to meet the standard. Taken to the extreme, this argument implies that, in choosing a method to determine adequate spending levels,
one is better off choosing a method that is easy to understand but wrong rather than a method that is difficult to explain but produces the right answers.” (p. 8)

9 Colorado has recognized this possibility by calculating cost of living for “labor pool areas,” which are designed to reflect where teachers in the district live, rather than where they work.

10 Some hedonic salary studies have not included any measures of student characteristics (Chambers et al., 2004), and a number of studies do not include measures of student poverty (Chambers et al., 2003; Taylor, 2004).

11 Specifically, the cost models are estimated with two-stage least squares regression, and “instruments” are identified for teacher salaries. Instruments have included private sector salaries, county population density, and teacher wages in surrounding or similar districts.

12 Economies of scale can occur either because of increasing returns to scale, which arise when a one percent increase in inputs leads to a more than one percent increase in output, or because prices for certain resources used by the school district decline with the amount of resources used.

13 Children with incomes at or below 130 percent of the federal poverty line are eligible for a free lunch, and students between 130 and 185 percent of the poverty line are eligible for a reduced price lunch. In addition, households receiving Food Stamps or Temporary Assistance to Needy Families (TANF) are also eligible for a free lunch. A description of the program and eligibility requirements is available at fns.usda.gov/nslp/national-school-lunch-program.

14 A recent study by Dahl and Scholz (2011) found that in 2002-2003 school year, only 72 percent of eligible children nationwide received free or reduced-price lunches while about 12 percent of ineligible received free or reduced-price lunches this school year.
Skin color is not literally a cost factor, of course, but in our society skin color is often highly correlated with student disadvantages. In this context, Baker (2011) provides a thoughtful discussion of the use of race in a cost function regression.

Language ability is estimated from the long form of the Census, in which individuals are asked if they speak a language other than English and are asked their ability to speak English. A household in which all members of the household 14 years or older do not speak English well and speak at least one other language than English are classified as linguistically isolated.

These weights are our estimates based on the estimated range in marginal effects for free lunch students in Table 3 in Gronberg et al. (2004) divided by average spending per pupil. They should not be attributed directly to the authors of that study. These relatively low weights may be due in part to the fact that this study interacts the child poverty rate with several other variables. Several of these interaction terms are not statistically significant. An alternative way to separate cost and efficiency is developed by Ruggiero (1998; 2001) based on the two-step methods developed by McCarty and Yaisawarng (1993). In the first stage, Ruggiero compares the spending per pupil of each district to a cost frontier (districts with equivalent outcomes and lower spending) using data envelopment analysis (DEA) discussed earlier in Endnote 3. The index produced from the first-stage DEA captures both inefficiency and cost differences across districts. Ruggiero then regresses this index on a set of cost factors, and the predicted value from this regression is his estimate of cost of education index.

Yinger and Duncombe (2004) developed poverty weights for New York using information in the professional judgment study (Chambers et al., 2004); these estimates, which should not be attributed to the authors of the study, range from 0.37 in middle school to 0.81 in elementary school.