National Tax Journal, March 2011, 64 (1), 193–218

THE IMPACT OF FACILITIES ON THE COST OF EDUCATION

Timothy J. Gronberg, Dennis W. Jansen, and Lori L. Taylor

This paper uses new data on school district capital stocks, stochastic frontier analysis, and a value-added measure of school quality to provide the first direct evaluation of the relationship between school facilities and school district costs. We find that the cost of education increases as the capital stock increases, suggesting either that school districts are grossly overcapitalized or that nicer facilities reflect an important, unmeasured dimension of school quality. We also find that cost function estimates for Texas are largely insensitive to the exclusion of the capital stock measures, suggesting that stochastic frontier cost function estimates without a capital stock measure are unlikely to be biased.

Keywords: school finance, school facilities, cost function

JEL Codes: H75, I22, I18, 128

I. INTRODUCTION

Cost function analysis can provide valuable information for addressing important public policy questions related to the structure and financing of K–12 school systems. It can be used to evaluate the potential benefits from major institutional design changes such as school district consolidations, expansions of charter schools, or the introduction of vouchers. Cost function analyses have been used to suggest appropriate adjustments to school funding formulas for differences in the educational environment or student demographics (Alexander et al., 2000; Duncombe and Yinger, 2005; Imazeki and Reschovsky, 2005, 2006). Recent legislative and judicial evaluations of school finance adequacy have looked to cost function analysis for guidance (Gronberg, et al., 2005; Imazeki and Reschovsky, 2005; Duncombe, Lukemeyer, and Yinger, 2008). In the current fiscal climate, policymakers may turn to cost function analyses to justify spending cuts.

Timothy Gronberg: Department of Economics, Texas A&M University, College Station, TX, USA (tjg@econmail.tamu.edu)

Dennis Jansen: Department of Economics, Texas A&M University, College Station, TX, USA (d-jansen@tamu.edu)

Lori Taylor: Bush School of Government and Public Service, Texas A&M University, College Station, TX, USA (ltaylor@bushschool.tamu.edu)

The value of information obtained from an estimated cost function rests, of course, upon the quality of the estimates. Obtaining high-quality cost function estimates for education is a particularly challenging task for at least three major reasons. First, measuring school quality is difficult. The common practice is to include various test-score-based measures of student achievement as an indicator of classroom quality. However, test scores are likely to be biased measures of school output because students from advantaged families tend to score well on standardized tests even when the school itself is mediocre, and students from disadvantaged families can score poorly even when the school is excellent.

A second major hurdle to education cost function estimation is the difficulty measuring many of the non-labor inputs used in the production of education services. This is especially true for the capital stock. A near-universal feature of education cost function estimates is the absence of good measures of school capital inputs. Although education is quite labor-intensive, there are good a priori reasons to believe that capital is an important factor in the production function for educational outcomes, and therefore an important factor in the (dual) educational cost function. Indeed, state policies for infrastructure funding are predicated on its central role in educational production. Unless one is willing to assume that existing capital stocks are at their long-run cost minimizing levels for all districts and that the price of capital is constant across all districts (rather heroic assumptions), excluding capital from cost function analyses introduces omitted variable biases.

The very real possibility that public school managers systematically fail to minimize cost creates a third obstacle to deriving valid estimates of school cost functions. There are several approaches to modeling inefficiency in a cost function context. Some researchers have chosen to add independent variables that may be correlated with inefficiency — such as education market concentration indices — as efficiency controls (e.g., Imazeki and Reschovsky, 2006; Duncombe and Yinger, 2007). Others have employed cost frontier techniques, both of the nonparametric data envelopment analysis (DEA) variety and of the parametric stochastic frontier type (Alexander et al., 2000; Gronberg et al., 2004, 2005). It is widely understood that estimates of inefficiency may reflect true inefficiency, or may be the result of omitted variables, either outputs or inputs.

In this paper, we address all three challenges. Our analysis incorporates a value-added measure of school quality that is based on changes in the performance of individual students over time. Value-added metrics greatly reduce the influence of student and family characteristics on the measure of school quality. We use a newly available series of capital stock indicators. Previous studies that have wanted to include capital stock measures have used proxies such as the number of books in the school library, school expenditures on maintenance, equipment and library books, or bond indebtedness (Jones and Zimmer, 2001; Callan and Santerre, 1990). In this paper, the capital stock indicators are based on insurance records and insurance valuations provided by individual school district superintendents. As such, our analysis provides the first direct evaluation of the relationship between capital stocks and school district costs. Finally, we directly address the questions of school inefficiency by estimating a stochastic cost frontier model that does not presume that school districts are operated efficiently.

We use our estimated model to address three important questions concerning the role of facilities in public school costs and the potential for cost savings from K–12 education. First, we look at the direct effect of capital on the variable operating costs and ask, “Do differences in school facilities lead to differences in the cost of education?” Second, we assess the effect of the inclusion of a measure of capital on the estimated marginal effects of outcomes and non-capital inputs on cost and ask, “Are cost analyses that do not include capital biased?” Finally, we ask the $64,000 question, “How much could be cut from operating expenditures if school facilities were allocated efficiently?”

II. THE COST FUNCTION MODEL

School districts produce education outcomes using a production process that combines input factors that are purchased (for example, teachers and other personnel) with environmental input factors that are not purchased (for example, student skills acquired in an earlier grade). Thus, we model school district cost as a function of the quantity and quality of outcomes produced, the prices of variable inputs, the quantity of the quasi-fixed capital, and the characteristics of students and parents that directly influence the education production process.

Our model is a district-level cost function. The underlying conceptual assumption is that resource allocation decisions are largely made at the district level, where district administrators allocate funds across campuses and across physical facilities and instructional programs. There are also practical considerations favoring a district-level cost function. First, the capital stock indicators are only available at the school district level. Second, districts have central administration costs that are borne jointly by campuses in the district. Allocating these costs among campuses, or ignoring them entirely, are issues that complicate any campus-level cost function analysis. Further, campuses and districts vary in how actual administrative services are conducted at campus or district levels, and they differ in how they apportion these administrative costs across units when they report to the state of Texas.

A. Sources of Data and Measurement Issues

Data for this analysis come from the Texas Education Agency and school district surveys. Due to concerns about cost function differences among districts with different service scope, we focus on districts that offer K–12 education. The six special-purpose school districts that have no tax base (most of which are located on military bases) are excluded because they face a different incentive environment than other school districts, as is one special purpose school district that serves as a residential facility for at-risk youth. All other traditional public school districts for which complete data are available have been included in the analysis. Charter schools are not included because capital stock data are not available for them. This analysis covers the 2006–07 school year.

1. Cost

We measure district costs as total operating expenditures per pupil, regardless of funding source. Therefore, our cost measure includes federal, state, and local dollars. We exclude debt service, and we also exclude transportation expenditures on the grounds that they are unlikely to be explained by the same factors that explain student performance. We exclude food expenditures on similar grounds. We include maintenance expenditures as they are a necessary feature of employing capital.

2. Quantity of Output/Scale

The quantity of output is conceptually measured as a student year and is practically measured as the number of students in fall enrollment. There are few dimensions on which Texas school districts differ more than scale. The five largest Texas districts have average daily attendance of more than 65,000 students; the five smallest Texas districts have average daily attendance of less than 30 students. Moreover, the smaller school districts are often located in regions where the population density is low, so that these districts may be geographically large.

3. Quality of Output

The principal quality indicator is a value-added measure of education outcomes that is based on changes in scores on the Texas Assessment of Knowledge and Skills (TAKS). We obtained math and reading scores for every student in grades three through eleven attending traditional public schools in Texas during the 2004–05 through 2006–07 school years. For all students for whom we could match test scores in the 2006 and 2007 spring test administrations, we calculated standardized tests score gains similar to the method used by Reback (2008). Thus, for each subject (math and reading) we divided the students into subgroups based on their grade level and test score in 2006. For example, all of the students in one subgroup had the same 2006 math score in the fourth grade; all of the students in another subgroup had the same 2006 reading scores in the tenth grade. We then calculated normal curve equivalent (NCE) scores for 2007 for each subgroup. The NCE is a standardized test score that is commonly used in the education literature. It is defined as $NCE = 50 + 21.06*z$, where z is the standardized test score, $z = (x_i - \mu)/\sigma$, the mean μ is calculated as the mean score for all students in, say, the subgroup in 2007 defined based on performance in 2006 (i.e., the subgroup of students with the same score in 2006), and the standard deviation σ is calculated as the standard deviation for all students in that subgroup. An NCE of 71.06 indicates that the student's post-test score in 2007 is one standard deviation above the mean for students with an identical pre-test score. Our value-added measure is the average NCE score in math and reading for this cohort-constant set of elementary and high school students (grades four through eleven).

One criticism of the TAKS is that it is a test of minimum performance. To capture the cost impact of higher levels of performance, we also consider a second district quality indicator: the percentage of students who complete an advanced course (i.e., an honors, advanced placement, or international baccalaureate course). This indicator is specifically directed at high achieving students in high school.

4. Input Prices

The education sector is highly labor-intensive, with three major sources of cost — teachers, administrators, and non-professional staff. To generate our measure of teacher price, we follow the literature and estimate a teacher salary index using a hedonic wage model and data on individual teacher earnings. The salary index for each district is the predicted wage for a teacher with five years of experience and a bachelor's degree, holding all other observable teacher characteristics constant at the statewide mean. (See the appendix for more on the teacher salary index.)

Administrator salaries respond to local market conditions in much the same way that teacher salaries do. While administrators clearly earn more than teachers with five years of experience, the differential between administrator wages and teacher wages is very consistent across districts. Thus, principal salaries track teacher salaries very closely. Because the two compensation measures are highly correlated, administrator salaries did not add much information to the cost function estimation and we do not include them in the final estimating equation.

Teachers and administrators may have very similar salary profiles across districts, but there is little reason to expect that the price of a teacher would be a good proxy for the price of other, nonprofessional school district personnel. Taylor (2008) estimated a comparable wage index for high school graduates (*HS CWI*) using the Individual Public Use Microdata Sample (*IPUMS 5-Percent*) from the 2000 U.S. Census and the Bureau of Labor Statistics' Occupational Employment survey. We use her *HS CWI* as the price index for non-professional educators in Texas.

Prices for instructional equipment and materials would ideally be included in our cost function but were not available to us. However, since prices for computer equipment and other instructional equipment and materials are largely set in a national market, any variation in prices among school districts should arise primarily from variations in transportation costs. Therefore, we include a measure of geographic isolation — the distance to the center of the nearest major metropolitan area — to capture variation in non-labor input prices.[1]

[1] Miles to the nearest major metro area for each district was calculated as the pupil-weighted average distance from each campus to the geographic center of the nearest county with a population of at least 650,000 people. (According to the 2000 Census, there are six major metropolitan areas in Texas fitting this description — Austin, Dallas, El Paso, Fort Worth, Houston, and San Antonio.) Distances are calculated using latitude and longitude information. The latitude and longitude of county centers come from the U.S. Census. Where available, latitude and longitude information for campuses are taken from the National Center for Education Statistics' Common Core Database. The remaining campuses are assigned latitudes and longitudes according to the zip codes at their street address. We censored the data at a lower bound of ten miles due to the concern that the geographic center of a county does not necessarily reflect the population center and our belief that districts in the range from zero to ten miles from the geographic center of the populous county were equally urban.

5. Capital Stock Indices

This analysis builds on a series of capital stock indicators developed by Taylor et al. (2005). Their capital stock indicators are based on survey information provided by individual school district superintendents and on insurance records provided by the Texas Association of School Boards (TASB). Superintendents were surveyed via e-mail in the fall of 2003, and asked to provide information about the age, square footage, replacement value, and value of contents for each building in their district. Although participation in the survey was voluntary, 327 of Texas's 1,039 traditional, Independent School Districts (ISDs) responded.[2] Those 327 ISDs employ 49 percent of the public school teachers in Texas.

To increase the coverage of the capital stock indicators, the survey responses were supplemented with additional information from TASB's insurance files. TASB serves as an insurer for many Texas school districts, and was the original source for the data provided by some of the survey respondents. Together, the insurance files and the survey responses provided data on 718 school districts which employ 63 percent of Texas' public school teachers.

This sample of 718 Texas school districts, though not random, is generally representative of Texas school districts. Districts in the sample are similar in size and growing at the same average rate as districts outside the sample. Differences in property wealth per pupil are negligible, as are differences in both the levels and the budget shares of capital outlays, plant maintenance, and debt service. However, as discussed in Taylor et al. (2005, p. 5), "sample districts have fewer students who are economically disadvantaged, a lower share of students with limited English proficiency, and a lower share of minority students than out-of-sample districts."

From the survey responses and TASB records, Taylor et al. (2005) developed a composite index of educational capital. The composite index indicates the total value of general purpose facilities and equipment in the district. This indicator excludes teacherages (residential structures that districts own for the use of district personnel) and capital devoted to athletics, but includes all other insured property. It ranges from less than $5,000 per student to more than $100,000 per student, with an un-weighted mean of $20,000 per student.

We use those cross-sectional estimates of educational capital as the foundations for perpetual inventory estimates of school district capital stocks in the years 2002–03 through 2006–07. We follow the approach used by the U.S. Bureau of Economic Analysis (BEA) to generate national capital stock estimates. We use the survey responses as the best available estimate of the capital stock at the end of year 2002–03 (K_0) and build the capital stock for following years using the annual capital expenditures by each school district as our estimate of capital investment.[3] We use the annual geometric depreciation rate used by the BEA for educational structures, 0.0182. Thus we calculate the net capital stock in 2003–04 through 2006–07 (K_1 through K_4) as:

[2] Charter schools were not included in the survey.

[3] We adjusted the capital outlays for inflation using the implicit GDP deflator for the Texas construction industry. Annual expenditures for athletics capital were excluded from the estimates of investment.

$$(1)\quad K_t = \sum_{i=1}^{t} (1-\delta/2)(1-\delta)^{t-i} I_i + (1-\delta)^t K_0,$$

where δ is the depreciation rate and I_j is investment in period j, where $j = 0$ in 2003–2004, $j = 1$ in 2004–05, etc. The net educational capital per pupil is this net capital stock divided by the number of students in fall enrollment.

6. Other Environmental Factors

In addition to facilities, we consider several other environmental factors that influence district costs but which are not purchased inputs. In principle, a range of student, family, and neighborhood characteristics would be included in the cost model. In practice, school district matched socioeconomic variables are hard to come by in non-Census years. Because a change in a student's test scores from year to year is due to both the quality of the schooling and to the quality of the student and home inputs, changes in scores are still impure measures of school input contributions to performance. To address this concern, we incorporated into our analysis the percentages of students in each district who were limited English proficient (LEP), received free or reduced-price lunches, were enrolled in high school, or were classified as special education students. To capture variations in the cost of education that arise because the student population is geographically dispersed and the population density is low, we include the log of the square miles in the school district.

For LEP, special education, and high school students, the required resource intensity per student is increased (for example, due to more specialized teachers and supplies and smaller required class sizes), thus driving up per pupil costs relative to regular students. The free lunch measure is chosen to account for the lower level of home inputs which are expected to obtain, on average, in households in official poverty status and in migrant households. The importance of home inputs in the production of student achievement is quite well established, and districts will have to increase other inputs to substitute for lower levels of parental-supplied inputs in order to produce any given quality of output.

B. Model Specification

The core functional form is a translog cost function. A primary advantage of the translog is its flexibility. Other popular functional forms — including the Cobb-Douglas specification used by Imazeki and Reschovsky (2004a, 2004b) — are restricted, special cases of the translog function. There is also a strong theoretical basis for preferring the translog cost function to its Cobb-Douglas counterpart. In a multi-output model setting, the Cobb-Douglas specification imposes the wrong curvature on the output cost frontier.[4] The primary estimation disadvantages of the translog specification are its complexity in evaluating marginal effects and statistical concerns with multi-collinearity

[4] The fundamental theory underlying all cost function analyses requires that there be a convex relationship among the outputs in a multiple-output cost function; the Cobb-Douglas specification forces the relationship to be quasi-concave, as discussed in Gronberg, Jansen, and Taylor (forthcoming).

and over-parameterization due to the presence of many interaction terms involving the explanatory cost factors.

The dependent variable is operating expenditures per pupil (E). The right-hand-side variables are the n_1 output variables (enrollment and quality measures, the q_i, the n_2 input prices w_j, and the n_3 environmental factors x_k. All variables are in natural logs.[5] Because school district size varies so greatly within Texas, other researchers have chosen to exclude the largest Texas districts from analysis (e.g., Imazeki and Reschovsky, 2004a). Rather than take such an approach, we also include a cubic term for log enrollment.

The model for district expenditures per pupil is:

$$(2)\quad \ln(E_i) = a_0 + \sum_{i=1}^{n_1} a_i q_i + \sum_{i=1}^{n_2} b_i w_i + \sum_{i=1}^{n_3} c_i x_i + .5\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} d_{ij} q_i q_j + .5\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} e_{ij} q_i w_j$$
$$+ .5\sum_{i=1}^{n_2}\sum_{j=1}^{n_2} f_{ij} w_i w_j + .5\sum_{i=1}^{n_3}\sum_{j=1}^{n_2} g_{ij} x_i w_j + .5\sum_{i=1}^{n_1}\sum_{j=1}^{n_3} h_{ij} q_i x_j + .5\sum_{i=1}^{n_3}\sum_{j=1}^{n_3} k_{ij} x_i x_j$$
$$+ l q_1^3 + v_i + u_i$$

We impose the symmetry restrictions that $d_{ij} = d_{ji}$, $f_{ij} = f_{ji}$, and $k_{ij} = k_{ji}$.

There exists a sizeable literature that finds that school districts do not all operate in an efficient, cost-minimizing fashion and that the degree of inefficiency varies considerably across districts. For example, Imazeki and Reschovsky (2004a, p. 41) find that "[t]he average district in Texas is 59 percent as efficient as the most efficient district in the state." Gronberg, Jansen, and Taylor (forthcoming) find that the average Texas district is 89 percent efficient while Taylor, Grosskopf, and Hayes (2009) find that the average Texas school district is 78 percent efficient.

We explicitly allow for the possibility of school district inefficiency using standard stochastic frontier techniques. We model the error term as being composed of a standard two-sided term to capture random un-modeled differences across districts (v_i) and a one-sided term to capture inefficiencies (u_i). We specify the one-sided error term as having a half-normal distribution. Jensen (2005) finds that specifying a half-normal distribution for the inefficiency term generates more reliable estimates of technical efficiency than other assumptions about the distribution of inefficiency.

Heteroskedasticity is always a concern for researchers estimating educational cost functions. The nature of the potential heteroskedasticity in the stochastic frontier context is more complicated than in the classical linear regression model due to the asymmetric

[5] All variables that are expressed as percentages or percentage points (such as the percent Limited English Proficient or the lagged passing rate) are assumed to enter the model in exponential terms, so that the natural log of the variable returns the percentage point values. This approach is necessary because these variables frequently take on the value of zero. Note that this strategy was also used by Eakin and Kniesner (1992) in their cost function study of hospitals.

composed error term in the frontier model. Heteroskedasticity could arise with respect to either or both error terms. If the heteroskedasticity only appears in the symmetric noise error component, then the kernel cost function parameter estimates are still unbiased. The estimates of district-specific efficiency are, however, potentially biased in the presence of heteroskedasticity in v. If the heteroskedasticity only appears in the one-sided error component, then both the cost function parameter estimates and the efficiency estimates can be biased. If heteroskedasticity is present in both error terms, then both potential types of bias exist.

We model heteroskedasticity in both the one-sided and two-sided error terms, with the one-sided variance function specified as a linear combination of three indicator variables, one for whether a district has less than 1,600 students, another for whether a district has less than 5,000 students, and a third for whether a small district serves more than 300 square miles. We choose these three indicators because the Texas school finance formula provides additional funding for small (less than 1,600 students), midsized (less than 5,000 students) and sparsely populated (small with more than 300 square miles) school districts, and those supplemental funds could foster inefficiencies. We model the two-sided variance as a function of whether the district receives a small size adjustment because measurement error is frequently a function of school district size.

The possible endogeneity of school quality indicators is also a common concern for researchers estimating educational cost functions. Unfortunately, the econometric literature provides little guidance as to the proper way to address these concerns in a stochastic frontier setting. Furthermore, the translog specification means that not only do we need instruments for the two quality indicators (the average NCE score and the share of students completing an advanced course), we also need instruments for all of the quality interaction terms — a total of 25 variables in all. The large number of potentially endogenous variables compounds the usual problems associated with weak instruments.

Following the literature, we explored using likely determinants of the local demand for education — the percentage of the adult population with at least a bachelors' degree, the percentage of households with school age children, the percentage of the population over age 65, the percentage of households with school age children, and the percentage of households that are owner occupied — and their interactions with the exogenous variables as potential instruments for the school quality indicators and their interactions with the exogenous variables. We also explored using two measures of yardstick competition — the share of students taking advanced courses in surrounding districts two years previously and the percentage passing the TAKS in surrounding districts two years previously — and their interactions as possible instruments. No combination of potential instruments passed a weak instruments identification test (based on the Cragg-Donald Wald F statistic from traditional, two-stage least squares estimation). Because weak instruments are often worse than no instruments at all, we treat all of the independent variables as exogenous in our estimation.

III. ESTIMATION RESULTS

Texas has 964 traditional school districts that serve students in kindergarten through twelfth grade (K–12). Capital stock data are available for 659 of these districts.[6] Three of those districts are missing one or more explanatory variables and must be excluded. Therefore, the estimation sample includes 656 Texas school districts.

Table 1 compares descriptive statistics for all K–12 districts and for the 656 districts in the estimation sample. As the table illustrates, the school districts in this analysis are generally representative of Texas as a whole. On average, the percentage of students who were economically disadvantaged or limited English proficient is somewhat lower for the districts in the estimation sample than for other districts, and the percentage of high school students is somewhat higher. None of the other differences in means are statistically significant at the 5-percent level, and none of the variables are significantly different at the 1-percent level.

Table 1

Summary Statistics

	Estimation Sample			Remaining K-12 Schools		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Per pupil operating expenditure (log)	656	9.008	0.214	308	9.000	0.225
Enrollment (log)	656	7.048	1.455	308	7.311	1.504
NCE score (log)	656	–0.711	0.110	308	–0.713	0.104
Percent taking advanced courses	656	0.183	0.089	305	0.180	0.086
Percent of students who are						
Low income	656	0.525	0.181	308	0.551	0.183
Limited English Proficient	656	0.072	0.084	308	0.087	0.108
Special Education	656	0.125	0.033	308	0.123	0.031
High school	656	0.303	0.041	308	0.298	0.036
Miles to major metropolitan area (log)	656	–0.231	0.861	308	–0.262	0.876
Square miles (log)	656	0.466	1.007	308	0.513	1.034
Teacher salary index (log)	656	–0.002	0.072	307	0.005	0.076
Non-professional salary index (log)	656	0.098	0.124	308	0.090	0.118
Educational capital per pupil (log)	656	0.641	0.364	3	0.337	0.142

[6] Capital stock data are also available for 62 school districts that serve grades K–8. These districts are not included because they lack a high school and therefore have a different educational technology.

A. The Baseline Estimation

The first column of Table 2 presents modified coefficient estimates from our baseline model.[7] For presentation purposes, we collapse the interaction terms by evaluating each variable at the mean of all of the other explanatory variables. For example, the translog cost function (2) can be rewritten as

$$(3)\qquad \ln(E_i) = \left(a_2 + d_{12}q_1 + d_{23}q_3 + \sum_{j=1}^{n_2} e_{kj}w_j + \sum_{i=1}^{n_3} h_{ki}x_i \right) q_2 + d_{22}q_2^2 + \theta,$$

where θ indicates all of the terms that do not involve q_2. Evaluated at the mean of all of the other variables, the term in parentheses becomes a constant (ϕ_{q_2}), yielding

$$(4)\qquad \ln(E_i) = \phi_{q_2} q_2 + d_{22}q_2^2 + \theta, \;\; where\;\; \phi_{q_2} = a_2 + d_{12}\bar{q}_1 + d_{23}\bar{q}_3 + \sum_{j=1}^{n_2} e_{kj}\bar{w}_j + \sum_{i=1}^{n_3} h_{ki}\bar{x}_i.$$

Table 2 presents the collapsed coefficient on the linear term (ϕ_{q_2}) and the estimated coefficient on the squared term (d_{22}) for each variable. Standard errors are in parentheses. Asterisks indicate a variable where one *cannot* reject the hypothesis that the variable and all of its interaction terms are jointly zero at the 5-percent level.

As the table illustrates, the baseline model is consistent with reasonable expectations about school district costs. We find that costs increase with increases in the quality of output and that the cost of reaching any particular performance level depends on educator wages, school district isolation, geographic size, and student needs. For example, the estimation suggests that when evaluated at the sample mean values for all explanatory variables, the cost of educating an economically disadvantaged student is 31.1 percent higher than is the cost for educating a student who is not eligible for free or reduced lunch.

As expected, economies of scale are substantial. The model suggests that the cost of achieving any given performance standard is 25 percent higher in a district with 100 students than it is in a district with 400 students, which in turn has costs 15 percent higher than in a district with 8,000 students.

What is the direct effect of district capital on district operating costs? As Table 2 illustrates, our evidence suggests that the relationship is nonlinear. For school districts with a relatively small capital stock, increases in the capital stocks have a negligible impact on operating costs; for districts with larger capital stocks, increases in the capital stock lead to increases in operating costs.

Figure 1 illustrates the relationship between educational capital per pupil and the cost of education, holding all other variables constant at the mean. The dashes indicate the region where the marginal effect of educational capital is not significantly different from zero at the 5-percent level. As the figure illustrates, school districts with a capital stock per pupil greater than approximately \$11,050 per pupil experience positive marginal cost.

---

[7] The full set of estimated coefficients is available from the authors.

Table 2

Cost Function Estimates

	Baseline	No capital	No Frontier	Cobb Douglas
Educational capital per pupil (log)	0.048**		0.045**	0.031**
	(0.035)		(0.044)	(0.030)
Educational capital, squared	0.049		0.075	0.068
	(0.026)		(0.037)	(0.023)
Enrollment	–0.723**	–1.058**	–0.218**	–0.827**
	(0.155)	(0.151)	(0.222)	(0.145)
Enrollment, squared	0.071	0.107	0.000	0.085
	(0.018)	(0.019)	(0.026)	(0.018)
Enrollment, cubed	–0.002	–0.004	0.001	–0.003
	(0.001)	(0.001)	(0.001)	(0.001)
NCE test change	0.108**	–0.012	0.219	0.061**
	(0.348)	(0.368)	(0.523)	(0.263)
NCE, squared	–0.003	–0.078	0.077	–0.027
	(0.239)	(0.256)	(0.368)	(0.173)
Advanced courses	0.027**	0.138**	0.373**	0.199**
	(0.16)	(0.159)	(0.195)	(0.116)
Advanced courses, squared	–0.050	–0.201	–0.789	–0.153
	(0.345)	(0.348)	(0.422)	(0.202)
Teacher salary index	0.569**	0.467**	0.434**	0.127
	(0.187)	(0.191)	(0.216)	(0.160)
Teacher salary index, squared	10.404	9.500	9.780	–1.774
	(3.216)	(3.363)	(3.836)	(0.824)
Nonteacher wage	–0.064	–0.107**	0.215	–0.080
	(0.144)	(0.147)	(0.186)	(0.103)
Nonteacher wage, squared	0.361	0.561	–0.761	0.519
	(0.566)	(0.580)	(0.709)	(0.337)
Distance	0.004**	0.015**	0.003**	0.004
	(0.009)	(0.008)	(0.011)	(0.007)
Distance, squared	0.010	0.014	–0.000	0.003
	(0.008)	(0.009)	(0.012)	(0.005)
Geographic size	0.018**	0.020**	0.032**	0.006**
	(0.006)	(0.006)	(0.007)	(0.005)
Geographic size, squared	0.010	0.011	0.007	0.008
	(0.003)	(0.003)	(0.004)	(0.003)

Table 2 (Continued)

Cost Function Estimates

	Baseline	No capital	No Frontier	Cobb Douglas
Low income percent	–0.273**	–0.385**	–0.342**	0.002**
	(0.199)	(0.212)	(0.292)	(0.105)
Low income, squared	0.557	0.616	0.627	0.308
	(0.178)	(0.188)	(0.257)	(0.095)
LEP percent	–0.243**	–0.070**	–0.282	–0.069
	(0.146)	(0.153)	(0.175)	(0.130)
LEP, squared	1.616	1.195	1.238	–0.131
	(0.608)	(0.603)	(0.725)	(0.286)
Special Ed percent	–1.112**	–1.149**	–0.300**	0.910**
	(0.894)	(0.892)	(0.924)	(0.652)
Special Ed, squared	7.309	7.142	3.808	–1.606
	(3.343)	(3.349)	(3.471)	(2.347)
High school percent	–2.005**	–1.492	–2.012**	–1.173**
	(0.998)	(1.000)	(1.016)	(0.697)
High school percent, squared	3.652	2.762	3.531	2.250
	(1.583)	(1.591)	(1.579)	(1.036)
Stochastic frontier?	Yes	Yes	No	Yes

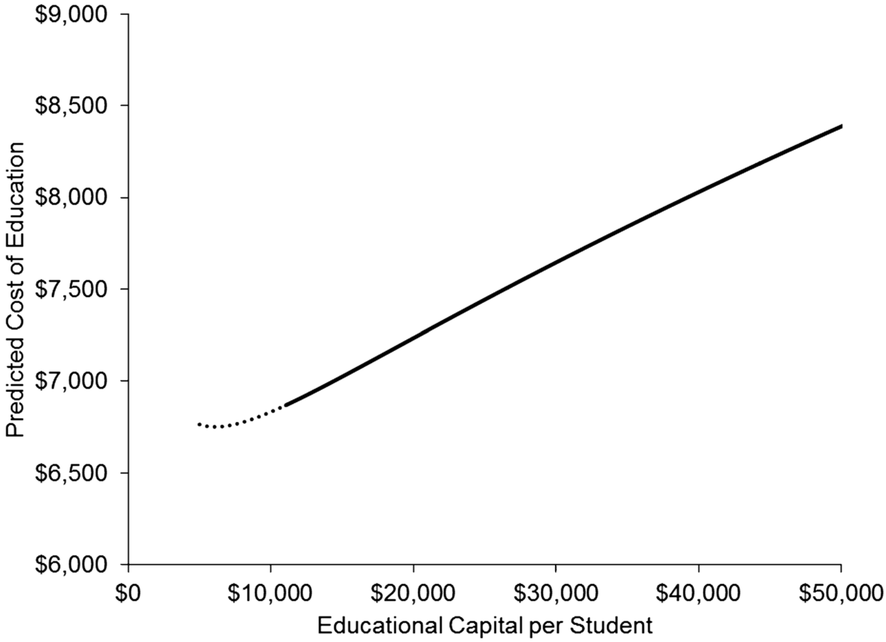
Notes: This table presents the collapsed coefficient on the linear term and the estimated coefficient on the squared term for each variable in the model. Standard errors are in parentheses and were calculated using the delta method. Asterisks indicate a variable where one cannot reject the hypothesis that the variable and all of its interaction terms (including the squared term) are jointly zero at the 5–percent level.

Most districts operate in the region where increases in capital stocks lead to increases in the cost of education. This presents a puzzle. Standard economic theory would have variable costs non-increasing, and usually decreasing, in the quantity of any quasi-fixed input. The estimated positive marginal cost implies a negative predicted shadow price on capital. Given the actual positive price of capital, the implication would be a gross overutilization of capital relative to an efficient allocation. In other words, the increased costs of maintaining a large stock of educational capital swamp any productivity gains associated with such capital.

Such a finding is not without precedent. Evidence of overutilization of capital is also found in a translog-based analysis of school districts in Connecticut by Callan and Santerre (1990). They argue the positive marginal effect is consistent with the average school capital stock being larger than its long-run cost-minimizing value. Similar results were obtained by Cowing and Holtmann (1983) in their study of short-run hospital

Figure 1

The Relationship Between Educational Capital and the Predicted Cost of Education



cost functions, and by Millimet and Collier (2008) in their recent frontier production function study of Illinois school districts.[8]

Although this is a possible interpretation, the degree of allocative inefficiency implied by our estimates seems implausible. An alternative interpretation is to posit a possible positive correlation between capital and unobserved quality variation. If a pleasant school environment is an important dimension of school quality (or is perceived to be an amenity by the local community) and if better environments are correlated with more capital, then the positive coefficient on capital may reflect the positive marginal cost of producing a higher quality school environment. In essence, under this interpretation capital serves as a proxy measure for an unobserved school amenity, and a positive coefficient is consistent with theoretical expectations. This unobserved quality is not captured by our observed quality measure, because it is not well correlated with traditional measures of student achievement.

This alternative interpretation also has roots in the literature. Cellini, Ferreira, and Rothstein (2010) examine the extent to which bond referenda for school facilities are

---

[8] Both of these studies include estimates using the translog function specification in their results.

capitalized into housing values, and find that only a small portion of the capitalized value can be explained by the relationship between academic outcomes and school facilities. They conclude that "[a] sizeable portion of the hedonic value of school facilities reflects nonacademic outputs" (Cellini, Ferreira, and Rothstein, 2010, p. 255).

Regardless of the interpretation attached to the observed positive coefficient, the estimates provide no evidence that increases in facilities capital will help districts to reduce the costs associated with maintaining or improving the test score performance of their students, or that infrastructure inequalities lead to academic disadvantages. A message that facilities capital is not a critical margin for improving test score measures of student achievement is also consistent with the evidence of a positive effect of charter school attendance on student performance from the lottery-based Gold Standard studies (Hoxby and Murarka, 2009; Abdulkadiroglu et al., 2009).[9] Charter schools do not receive the same state funding support for capital as do traditional public schools, and charters tend to operate in relatively modest facilities yet achieve student achievement results generally comparable to those achieved by traditional public schools.

The estimates also do not support the hypothesis that districts with modest capital stocks need more educational resources to compensate for their lack of capital. Within the range of capital stocks under analysis, there is no reason to expect that upgrading school district facilities would lower operating costs. Districts with ample capital stocks do not have lower operating costs; they just spend more on maintenance.

1. *Efficiency*

Our stochastic frontier model generates efficiency estimates for each district in our sample. Table 3 provides summary statistic information on the distribution of estimated efficiency values for the baseline model. The efficiency estimates indicate the percentage by which one could reduce cost without reducing measured outputs. A perfectly efficient district would have an efficiency estimate of 1.00. A school district with an efficiency estimate of 0.75 could reduce costs by 25 percent without reducing measured output. The evidence in Table 3 indicates that on average Texas school districts could reduce

Table 3

Cost Efficiency Estimates

Variable	Mean	Std. Dev.	Minimum	Maximum
Baseline	0.893	0.081	0.489	0.990
No Capital	0.881	0.092	0.439	0.991
Cobb-Douglas	0.893	0.079	0.502	0.988

[9] For a recent review of the empirical literature on charter schools, including both experimental /lottery and nonexperimental studies, see Gronberg and Jansen (2009).

their operating expenditures by 11 percent without reducing their measured outputs or changing their capital stocks.

The measurement of efficiency is, of course, directly linked to the particular set of performance measures that are included in the cost model and the particular set of input measures. For example, if only a math and reading performance measure is included in the estimated cost function model but a district devotes significant resources to music and athletic programs (and if those programs have little correlation with math and reading) then this district may be identified as inefficient, even if all of its programs are being produced as efficiently as possible.

Figure 2 plots the distribution of these measures for our baseline model. The distribution is heavily skewed, and indicates that while the majority of districts have values above 0.9 and a large majority have values above 0.8, there are a non-trivial number of districts with values lower than these.

Using the baseline model, and assuming that school districts operate efficiently given their existing capital stocks, we can predict the per pupil cost of an average quality education for each school district. These estimates range from $5,497 to $20,126, with a mean of $7,384. (as shown in Table 4). We then simulated the cost of an average

Figure 2

The Distribution of Cost Efficiency

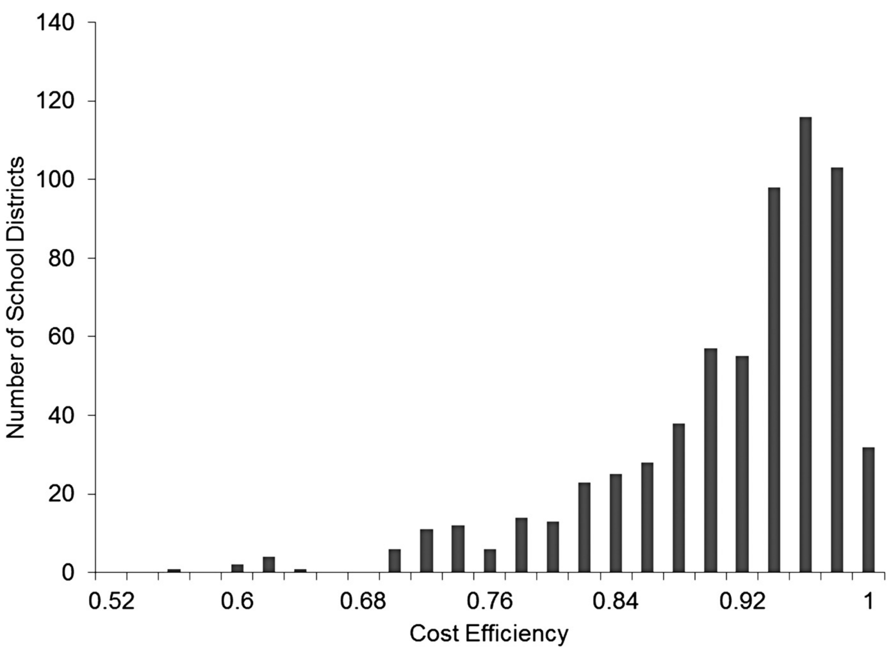


Table 4

Predicted Cost of an Average Quality Education, per Pupil (Dollars)

Variable	Mean	Std. Dev.	Minimum	Maximum
Baseline	$7,384	$1,380	$5,497	$20,126
Cost-minimizing capital simulation	6,949	1,021	5,163	13,848
No capital	7,267	1,252	5,614	17,674
No frontier	8,383	1,877	5,797	24,500
Cobb-Douglas	7,373	1,228	5,761	16,784

quality education, assuming that all school districts had a capital stock of efficient size. In other words, we assigned a capital stock of $11,050 per pupil to all districts with a larger capital stock, and left the capital stocks of all other districts unchanged. Under these alternative circumstances, average cost fell by $435 per pupil to $6,949 per pupil. Aggregating across the 656 school districts in the estimation sample, the analysis suggests that reducing the size of the educational capital stock to its efficient level could reduce educational expenditures in Texas by 3 percent, or approximately $565 million dollars per year.

B. Comparison to a Model without Capital

In the second column of Table 2 we estimate our baseline model while excluding capital. This exercise is designed to explore the sensitivity of the estimated marginal effects of traditional education cost function variables and the sensitivity of estimated school district cost inefficiencies to the inclusion of a measure of school capital.

As the results in the table illustrate, the exclusion of capital results in significant changes in the estimated coefficients. The impact of enrollment on total costs increases. The marginal effect of distance from a major metropolitan area also increases, perhaps indicating a correlation between remoteness and measures of the capital stock. The model without capital yields smaller estimates of the differential cost to educate economically disadvantaged students than does the model with capital. Evaluated at the mean of the other variables, the model without capital indicates that the cost of educating an economically disadvantaged student is 26 percent higher than the cost of educating a student who is not economically disadvantaged, rather than the 31 percent differential estimated in the baseline model.

Most importantly, although we can reject the hypotheses that the two quality measures (and all their interactions) are jointly zero, excluding the capital stock indicator leads to the mistaken conclusion that one of the quality measures — the average NCE score — is not significantly related to cost. Compared to the baseline model and evaluated at the mean of the other variables, the model without capital understates the marginal cost of an increase in NCE scores and overstates the marginal cost of an increase in the share of students taking advanced courses.

Figure 3 compares the predicted relationship between scale and cost for the models with and without capital. In both cases, the cost predictions were calculated holding all other regressors constant at the mean. As the figure illustrates, excluding the capital stock increases the predicted cost of education for districts at either extreme of the size distribution. The predicted cost of operating a school district with 200 students is 4 percent higher when capital is omitted than when it is included. The predicted cost of operating a school district with 100,000 students is also 4 percent higher in the models without capital stocks than it is in the baseline model.

The second row in Table 3 presents the district efficiency measures from a model excluding capital. As these results illustrate, the without-capital distribution of estimated efficiency values is similar to the with-capital distribution. As in the baseline model, the distribution of efficiency measures is heavily skewed, and the estimates are concentrated above 0.8.

The correlation between the efficiency estimates for the two models is 0.95 (as shown in Table 5.) Not surprisingly, the difference in estimated efficiency is a function of the size of the school district capital stock. Figure 4 plots the differences between the efficiency estimates from the two models against educational capital per pupil. As the figure illustrates, compared to the baseline, the model without capital indicates that

Figure 3

The Relationship Between Scale and Cost for Two Alternative Models

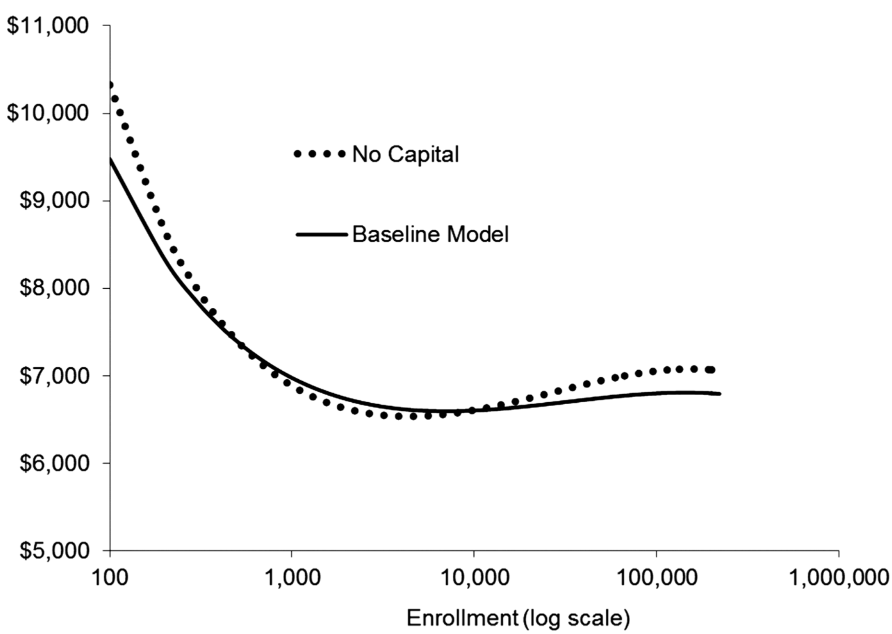


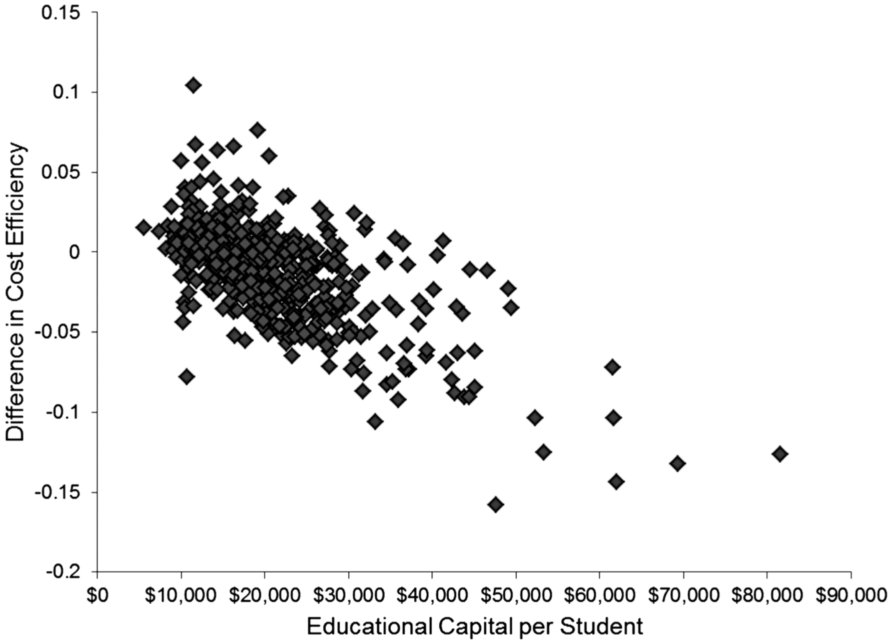
Table 5

Pearson Correlations Between Cost Efficiency Estimates

	Baseline	No Capital	Cobb-Douglas
Baseline	1.000		
No Capital	0.956	1.000	
Cobb-Douglas	0.928	0.911	1.000

Figure 4

The Difference in Cost Efficiency Between the Model With and Without Capital



districts with larger capital stocks are less efficient than districts with smaller capital stocks. In other words, the costs associated with maintaining larger capital stocks are attributed to inefficiency in the model without capital.

The third row in Table 4 presents the predicted per pupil cost of an average quality education, based on the model without capital. As the table illustrates, excluding the capital stock from the model has a modest impact on the predicted cost of producing an average quality education. The Pearson correlation between the two sets of per-pupil cost projections is 0.97, and they differ in their average estimate of cost by $117 per pupil (1.6 percent).

Table 6

Pearson Correlations Between Cost Projections

Variable	Baseline	No Capital	No Frontier	Cobb-Douglas
Baseline	1.000			
No Capital	0.968	1.000		
No Frontier	0.976	0.930	1.000	
Cobb-Douglas	0.957	0.935	0.933	1.000

Thus, in response to our second key question about the importance of capital measures in estimating educational cost functions, our evidence suggests that omitting the capital stock variable from the analysis has a modest influence on the key estimated marginal effects and a somewhat larger effect on the estimated pattern of differences in efficiency across districts. The finding of marginal effect insensitivity is particularly relevant for policy, because predictions based on marginal effects underlie all cost function-based estimates of the cost of educational adequacy. The insensitivity of relative efficiency rankings is also of potential policy importance if cost efficiency estimates are used to identify relatively efficient and inefficient district operations. Our results suggest the absence of a measure of district facilities capital is important but should not significantly bias, and therefore not hinder, the application of frontier cost estimation to other states.

C. Comparison to a Non-frontier Model

The third column of results in Table 2 is for non-frontier estimates of the cost function. The purpose in this case is to assess the effect of a failure to account for inefficiency on the parameter estimates of the core cost function. Three differences in the estimates stand out. First, a failure to control for inefficiency leads to the mistaken conclusion that one of the quality measures — the average NCE score — is not significantly related to cost. Second, the non-frontier model yields strikingly different estimates of the economies of scale. Finally, as expected, the non-frontier model generates cost predictions that greatly exceed those produced by the baseline model. As Table 4 illustrates, the non-frontier model indicates that the cost of providing an average quality education is $1,000 per pupil higher than the cost projections from the baseline model.

D. Comparison to a Restricted Cobb-Douglas Cost Function Model

It is common in the school district cost literature to assume a Cobb-Douglas functional form for the cost function. Given the appropriate concerns about the translog approach that we mentioned above, it is worthwhile to estimate a Cobb-Douglas cost model for our data set. In addition to our direct interest in comparing findings from the two alternative specifications, we can also test the parametric restrictions on the translog function which yield the Cobb–Douglas function as a nested functional form.

The fourth column of Table 2 presents the marginal effects from a modified Cobb-Douglas version of the baseline model — one that, as is common in the literature, includes the quadratic terms but not the interaction terms.

An F-test of the Cobb-Douglas restrictions easily rejects the assumption of a Cobb-Douglas functional form. A comparison between the first and fourth columns of Table 2 illustrates some of the potential impacts of mis-specifying the model as being Cobb-Douglas. Economies of scale are exaggerated under the Cobb-Douglas specification, while the marginal effects of a higher teacher wage are greatly understated. Intriguingly, the estimated marginal effect of an increase in the capital stock is very similar across the two specifications, as is the predicted cost of an average quality education.

IV. CONCLUSIONS AND POLICY IMPLICATIONS

In an era of increasing accountability for the quality of performance of public schools, the importance of understanding the relationships among the cost of education, quality, quantity, and input market conditions is heightened. Econometric exploration of these relationships is particularly challenging due to a number of features of the public school decision environments; these include the inherent difficulties in measuring quality, the importance of student and family inputs in the production of quality, the absence of measures of public school capital, and weak incentives for cost efficient behavior. We offer a framework for education cost function estimation that attempts to address these challenges with credible best-practice methods.

In addition to identifying a basic education cost framework, we provide initial model estimates that focus on the potential importance of including capital stock measures in the empirical specification. Our analysis of a cross-section of Texas school districts suggests that capital has a significant direct effect on operating cost. We find that the cost of education is higher in school districts with more capital per pupil, a pattern that suggests either that there is gross overcapitalization by school districts or that nicer facilities reflect an important dimension of school quality that is not well correlated with traditional measures of student achievement. We find no evidence that differences in student achievement across school districts in Texas can be attributed to inadequate facilities.

We also find that cost function estimates for Texas are surprisingly insensitive to the inclusion or exclusion of capital stock measures, suggesting that cost function estimates that do not include a capital stock measure are unlikely to be biased. Indeed, our analysis suggests that the bias introduced by failing to model school district inefficiency is more problematic than any bias arising from a lack of controls for school district capital.

The policy implications of our analysis are substantial. Our results suggest that policies that foster increased investment in school facilities could increase the cost of education, while policies that help school districts in declining markets reduce their capital stocks could lower the cost of education. Under current policy, capital funding is targeted toward districts in which enrollments are growing rapidly and the need for new facilities is obvious. Our analysis suggests that it could be at least as cost effective to target some funding toward districts with declining enrollments so that they could trade down to facilities that better match their current and projected enrollment levels.

Our analysis strongly suggests that equalizing access to educational facilities is not a critical margin for equalizing test score measures of student achievement. For a given level of educational spending, there is no reason to believe that student achievement gains would be systematically greater in school districts with large capital stocks per pupil than in school districts with more modest capital stocks, all other things being equal. Instead, the evidence suggests that performance gains would be lower in the heavily capitalized district because resources would be diverted to maintenance. Although students and teachers would undoubtedly be happier spending their days in a modern school building with a multitude of amenities, there is no evidence that students would learn more in such surroundings, or that students whose schools are less attractive are *academically* disadvantaged by the differential.

AKNOWLEDGEMENTS

The authors thank the State of Texas Education Research Center at Texas A&M University for research support on this project. Professors Gronberg and Jansen also thank the Private Enterprise Research Center at Texas A&M University for research support. All views expressed are those of the authors alone. The conclusions of this research do not necessarily reflect the opinions or official positions of the Texas Education Agency (TEA), the Texas Higher Education Coordinating Board (THECB), or the State of Texas.

REFERENCES

Abdulkadiroglu, Atila, John Angrist, Sarah Cohodes, Susan Dynarski, Jon Fullerton, Thomas Kane, and Parag Pathak, 2009. "Informing the Debate: Comparing Boston's Charter, Pilot and Traditional Schools." Boston Foundation, Boston, MA, http://www.masscharterschools.org/advocacy/HarvardStudy_54pg.pdf.

Alexander, Celeste D., Timothy J. Gronberg, Dennis W. Jansen, Harrison Keller, Lori L. Taylor, and Philip U. Treisman, 2000. "A Study of Uncontrollable Variations in the Costs of Texas Public Education." Report prepared for the 77th Texas Legislature. Charles A. Dana Center, University of Texas at Austin, Austin, Texas, http://www.utdanacenter.org.

Callan, Scott J., and Rexford E. Santerre, 1990. "The Production Characteristics of Local Public Education: A Multiple Product and Input Analysis." *Southern Economic Journal* 57 (2), 468–480.

Cellini, Stephanie R., Fernando Ferreira, and Jesse Rothstein, 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics* 125 (1), 215–261.

Cowing, Thomas, and Alphonse G. Holtmann, 1983. "The Multiproduct Short-run Hospital Cost Function: Empirical Evidence and Policy Implications from Cross-section Data." *Southern Economic Journal* 49 (2), 637–653.

Duncombe, William, Anna Lukemeyer, and John Yinger, 2008. "The No Child Left Behind Act: Have Federal Funds Been Left Behind?" *Public Finance Review* 36 (4), 381–407.

Duncombe, William, and John Yinger, 2007. "Does School District Consolidation Cut Costs?" *Education Finance and Policy* 2 (4), 341–375.

Duncombe, William, and John Yinger, 2005. "How Much More Does a Disadvantaged Student Cost?" *Economics of Education Review* 24 (5), 513–532.

Eakin, B. Kelly, and Thomas J. Kniesner, 1992. "Estimating a Non-minimum Cost Function for Hospitals." *Southern Economic Journal* 54 (3), 583–597.

Gronberg, Timothy J., and Dennis W. Jansen, 2009. "Charter Schools: Rationale and Research." Working Paper. Show-Me Institute, University of Missouri, Columbia, MI.

Gronberg, Timothy J., Dennis W. Jansen and Lori L. Taylor, forthcoming. "The Adequacy of Educational Cost Functions: Lessons from Texas." *The Peabody Journal of Education*.

Gronberg, Timothy J., Dennis W. Jansen, Lori L. Taylor, and T. Kevin Booker, 2004. "School Outcomes and School Costs: The Cost Function Approach." Joint Select Committee on Public School Finance, Austin, TX, http://bush.tamu.edu/research/faculty/TXSchoolFinance/.

Gronberg, Timothy J., Dennis W. Jansen, Lori L. Taylor, and T. Kevin Booker, 2005. "School Outcomes and School Costs: A Technical Supplement." Joint Select Committee on Public School Finance, Austin, TX, http://bush.tamu.edu/research/faculty/TXSchoolFinance/.

Hoxby, Caroline M., and Sonali Murarka, 2009. "Charter Schools in New York City: Who Enrolls and How They Affect Their Student's Achievement." NBER Working Paper No. 14852. National Bureau of Economic Research, Cambridge, MA.

Imazeki, Jennifer, and Andrew Reschovsky, 2004a. "Financing Education so that No Child is Left Behind: Determining the Costs of Improving Student Performance." *Developments in School Finance 2003, 33–52.* National Center for Education Statistics, Washington, DC.

Imazeki, Jennifer, and Andrew Reschovsky, 2004b. "Estimating the Cost of Meeting the Texas Educational Accountability Standards." Unpublished manuscript. San Diego State University, San Diego, CA, http://aasa.files.cms-plus.com/PDFs/TXAdeq2004.pdf.

Imazeki, Jennifer, and Andrew Reschovsky, 2005. "Assessing the Use of Econometric Analysis in Estimating the Costs of Meeting State Education Accountability Standards: Lessons from Texas" *Peabody Journal of Education* 80, 96–125.

Imazeki, Jennifer, and Andrew Reschovsky, 2006. "Does No Child Left Behind Place a Fiscal Burden on States? Evidence from Texas," *Education Finance and Policy* 1 (2), 217–246.

Jensen, Uwe, 2005. "Misspecification Preferred: The Sensitivity of Inefficiency Rankings." *Journal of Productivity Analysis* 23 (2), 223–244.

Jones, John T., and Ron W. Zimmer, 2001. "Examining the Impact of Capital on Academic Achievement." *Economics of Education Review* 20 (6), 577–588.

Millimet, Daniel L., and Trevor C. Collier, 2008. "Efficiency in Public Schools: Does Competition Matter?" *Journal of Econometrics* 145 (1–2), 134–157.

Reback, Randall, 2008. "Teaching to the Rating: School Accountability and the Distribution of Student Achievement." *Journal of Public Economics* 92 (5–6), 1394–1415.

Taylor, Lori L., 2008. "Washington Wages: An Analysis of Educator and Comparable Non-educator Wages in the State of Washington." Washington State Institute for Public Policy and the Joint Task Force on Basic Education Finance, Olympia, WA.

Taylor, Lori L., Sara Barrineau, Leslie Barron, Matthew Fiebig, Sarah Forbey, Jennifer Gray, Joshua Hodges, Jeff Jewell, Ashley Kelm, Marcia Larson, Erin Lesczynski, Zach May, Steve Murello, Jennifer Myers, Megan Paul, Manal Shehabi, and Megan Stubbs, 2005. "Meeting Needs? A Survey of School Facilities in the State of Texas." Joint Select Committee on Public School Finance, Austin, TX, http://bush.tamu.edu/research/faculty/TXSchoolFinance/.

Taylor, Lori L., and William J. Fowler Jr., 2006. "A Comparable Wage Approach to Geographic Cost Adjustment." U.S. Department of Education, National Center for Education Statistics, Washington, DC.

Taylor, Lori L., Shawna Grosskopf, and Kathy J. Hayes, 2009. "Is a Low Instructional Share an Indicator of School District Inefficiency? Exploring the 65-Percent Solution." Unpublished manuscript. Texas A&M University, College Station, TX.

APPENDIX: ESTIMATING THE TEACHER SALARY INDEX

The teacher salary index for each district is based on the predicted wage for a teacher with five years of experience and a bachelor's degree, holding all other observable teacher characteristics constant at the statewide mean. The predicted wage is then divided by the average predicted wage to yield the salary index. The wage predictions come from a hedonic model of teacher salaries, in which salaries are a function of individual teacher characteristics, school characteristics, and local labor market conditions. Interaction terms allow all of the model coefficients to differ between metropolitan and non-metropolitan areas. Table A.1 presents the coefficient estimates and robust standard errors for the variables used in the estimation of the predicted wage.

The data on teacher salaries and individual teacher characteristics come from the Texas Education Agency (TEA) and Texas' State Board for Educator Certification (SBEC). The measure of teacher salaries that is used in this analysis is the total full-time equivalent monthly salary, excluding supplements for athletics coaching.[10] The hedonic model includes controls for teacher experience (the log of years of experience, the square of log experience and an indicator that takes on the value of one if teacher experience is missing)[11] and indicators for the teacher's gender,

[10] Just over 8 percent of the classroom teachers in Texas receive supplemental pay for coaching duties. The average coaching supplement is $5,300 per year, but each year during the analysis period at least 20 classroom teachers received a coaching stipend in excess of $20,000. More than one-quarter of the male classroom teachers in Texas received a coaching supplement during the analysis period.

[11] Teaching experience is unknown for approximately ten percent of the observations. This missing-data indicator takes on the value of one if the teacher's experience is unknown, and zero otherwise. Years of experience was set to zero for all observations where the missing data indicator was set equal to one.

Table A1

Selected Coefficients from the Hedonic Model of Teacher Salary

	Metropolitan Areas	Non–metropolitan Areas
Percent of Students who are		
Economically disadvantaged	–0.052 (0.002)**	–0.048 (0.004)**
Limited English proficient	0.004 (0.005)	0.021 (0.010)*
Hispanic	0.080 (0.003)**	0.034 (0.005)**
African American	0.080 (0.003)**	–0.008 (0.006)
School size (log)	0.064 (0.001)**	0.032 (0.001)**
Comparable Wage Index	–0.173 (0.007)**	0.001 (0.012)
Fair market rent (log)	0.005 (0.001)**	–0.001 (0.002)
Border county	0.015 (0.002)**	0.027 (0.003)**
County population declining	0.032 (0.003)**	–0.011 (0.002)**
Miles to metro center	–0.001 (0.000)**	–0.000 (0.000)**
Very small district	0.005 (0.001)**	–0.009 (0.001)**
Small district	–0.002 (0.001)**	–0.009 (0.001)**
Midsized district	–0.004 (0.000)**	–0.010 (0.001)**
Large district	0.008 (0.001)**	0.000 (0.000)
Micropolitan Area		–0.001 (0.002)
Observations	1,433,349	
Number of teachers	412,571	
Adjusted R-squared	0.81	

Notes: The model was estimated in a single equation using interaction terms to allow all of the coefficients to differ between metropolitan and non–metropolitan areas, and with individual teacher fixed effects. The model also includes year indicators, metropolitan area indicators, and separate metropolitan and non–metropolitan indicators for teacher demographics certification status, coaching status, and whether or not the teacher is new to the district. Robust standard errors are in parentheses. Asterisks denote significance at the 1% (***), 5% (**), and 10% (*) levels.

race (black, Hispanic, or Asian/Indian), educational attainment (no degree, master's degree or doctorate), teaching assignment (math, science, special education, health and physical education, or language arts) and certification status (certified in any subject, and specifically certified in mathematics, science, special education, or bilingual education). Only teachers who worked at least half-time for a traditional Texas school district during the analysis period are included in the analysis. For purposes of this analysis, a traditional Texas school district collects taxes within a designated geographic boundary and serves the full spectrum of grade levels. Charter schools, K–8 school districts, school districts on military bases, and special purpose residential school districts have been excluded. The analysis covers the five-year period from 2003–04 through 2007–08.

The student demographics used in this analysis are the percentage of students in the district who are economically disadvantaged, limited English proficient, black, and Hispanic. We measure school size as the log of average campus enrollment in the district. The analysis includes four indicators for school district size — one indicator variable for very small districts (those with less than 800 students in average daily attendance), one for small districts (those with at least 800, but less than 1,600 students), one for midsized school districts (those with at least 1,600, but less than 5,000 students) and one for very large school districts (those with more than 50,000 students in average daily attendance).[12]

We use a number of indicators for local labor market conditions outside of education. We use the National Center for Education Statistics' Comparable Wage Index to measure the prevailing wage for college graduates in each school district (Taylor and Fowler, 2006).[13] We include the Department of Housing and Urban Development's estimate of Fair Market Rents (in logs), and indicators for counties on the border with Mexico, each metropolitan area, and for whether or not the district was located in a micropolitan statistical area. To accommodate typical rent gradients and geographic isolation, the analyses also include the distance from the district to the center of the closest metropolitan area. To accommodate potential salary differentials arising from population flows, we also include an indicator for whether or not the county lost population over the previous five-year period.

[12] The median Texas district has less than 1,000 students in average daily attendance. Only nine Texas school districts have more than 50,000 students.

[13] We extended the Comparable Wage Index through 2008 following the methodology used by Taylor and Fowler (2006).