ASSESSING THE EFFICIENCY OF OKLAHOMA PUBLIC SCHOOLS: A DATA ENVELOPMENT ANALYSIS

Susanne Rassouli-Currier, University of Central Oklahoma

ABSTRACT

In this paper, the efficiency of the Oklahoma school districts using two different specifications is measured by the Data Envelopment Analysis (DEA) method. To determine the possible sources of inefficiency, a second stage Tobit regression was employed. Here, the specification of the inefficiency models includes (1) environmental variables that school districts have no control over (e.g., the percentage of students in special education and the poverty rate in the district) and (2) non-traditional inputs that school districts do have control over (e.g., teachers' salaries) but were not included in the first stage DEA. The findings of the models are compared and both suggest that the key factors affecting efficiency measures among the Oklahoma school districts are primarily the students' characteristics and family environment.

INTRODUCTION

Given that tax revenue is the direct source of the operational funding for public schools in the United States, taxpayers expect a certain level of quality in the provision of educational services. While schools' real expenditures have been increasing, standardized test scores—often used indicators of school quality—have shown little if any improvement. This unsatisfactory outcome has raised serious questions about the management and efficiency of public schools in the U.S.

The general findings of the 1991 International Assessment of Education Progress (IAEP) show that American students 9 and 13 years old are generally behind their peers from other countries, particularly in science and mathematics. Hanushek (1996) suggests that U.S. schools have large increases in resources with very little, if any, improvement in outcomes. These findings confirmed his earlier statement in (Hanushek, 1986, p. 1162) that "*there appears to be no strong or systematic relationship between school expenditures and student performance*".

After reviewing a number of educational production frontiers, Taylor (1994) provides even more evidence of inefficiency. According to her findings, the United States' public schools are on average 15 percent inefficient. This has significant economic consequences, especially its effect on gross domestic product (GDP). Conversely, Bishop (1989) suggested that if the test scores had been rising during the 1970s, labor quality-measured by productivity growth- would have had increased by at least 2.9 percent and thereby led to an increase of 86 billion dollars in GNP.

Many studies suggest that perhaps the problem is not the level of school funding, but that reallocation of existing expenditures in ways that improve performance should be considered. The studies suggest that inefficiencies could be due to exogenous factors such as the breakdown of the family, poverty, increased immigration, a misallocation of resources within the schools themselves, or the adoption of inferior pedagogy.[1]

This paper attempts to extend previous efficiency studies, particularly Rassouli-Currier (forthcoming 2007). Given the nature of the production frontier function in this study, i.e., multiple outputs, multiple inputs, DEA is better suited for estimation of efficiencies than its parametric counterpart (Coelli et al., 1998; Kirjavainen, et al., 1998). The DEA will yield estimates of district efficiencies that can subsequently be modeled as functions of other district characteristics.[2]

REVIEW OF LITERATURE

To stress the maximal property of the production function in the context of efficiency measurement, the literature refers to the function as the production frontier (Coelli et al.1998, p.12). Firms operating on the frontier (production or cost) are considered to be technically efficient. Any deviation from the frontier is an indication of technically inefficiency.

According to Farrell (1957), in efficiency studies a frontier (extreme point) estimator rather than an estimator of central tendency (average) is prefered.[3]The author suggests that the technical efficiency (TE) equals one minus the maximum equal-proportional reduction of the input vector, given an output vector. Thus TE can take values between 0 and 1, and the degree of inefficiency of the firm can be measured.

Charnes, Cooper, and Rhodes (1978) pioneered the concept of data envelopment analysis (DEA), which is the parallel approach to the econometric estimation of frontier models and involves the use of linear programming. One of the serious shortcomings of the DEA approach is that the sampling distributions of DEA estimators are unknown and it does not account for one of the most significant and robust results of schools' input-output studies, the effect of the socioeconomic factors not under the control of the school.

To correct this, following McCarty and Yaisawarng (1993), Kirjavainer and Loikkaner (1998) employ a two-stage model. In the first stage they use DEA and calculate the efficiency scores using traditional inputs i.e., variables that are controlled by schools. The second stage involves the use of ML estimation of the Tobit regression model and provides efficiency measures based on variables that are not included in the DEA and are not under the control of the school districts. The Tobit model has well known and desirable statistical properties.

Educational Production

In education studies, data on input prices generally are not available. Therefore, in analyzing production frontier models in education, the use of the production function rather than the cost function may be more practical (Chakarborty et al., 2001).

The Coleman report (Coleman et al., 1966) has inspired several economic studies of educational production, efficiency, and cost structure. His report suggested an input-output relationship between administrative funding and students' achievements. Therefore, the production of education consists of school and non-school inputs which produce multiple outputs (performance test scores). In addition, the report pointed out the consideration of analytical issues such as production efficiency and the existence of multicollinearity among variables (Hanushek, 1979). The report suggests that students' achievements are largely related to their socioeconomic background rather than the differences in schools (Hanushek, 1986, 1989).

Subsequent analyses of school performance differ in their focus and methodology; however, they provide some understanding of school efficiency. Hanushek's 1986 survey of 147 studies suggests that most studies include expenditures per pupil, student/teacher ratio, teacher education and experience as well as family characteristics as the primary determinants of student achievement. These studies consistently conclude that (i) expenditure and student performance are not systematically related and (ii) family characteristics have an effect on the students' achievement. Adkins and Moomaw (1997) used a parametric method of estimation (Stochastic Frontier Regression) for grades 3,7,9 and 11 separately and found that, in the case of Oklahoma public schools, the relationship of test scores with respect to spending is positive but very small. Their findings also suggest that districts with more experienced teachers are generally more efficient (except for grade 3) and that districts that pay higher salaries get better results.

The effect of scale economies on schools' productivity, captured by the effect of the school size and/or district size on student performance, is inconsistent across studies. Some find evidence of economies of scale and others do not. Adkins and Moomaw (1997) suggest the presence of economies of scale in Oklahoma public schools, i.e., larger districts in Oklahoma tend to be more efficient. Therefore, districts might benefit by consolidation.

Data Envelopment Analysis (DEA) and Second-Stage Tobit

The basic idea of this approach is to view schools as productive units with multiple inputs and outputs. DEA assumes that all firms (schools) have the same deterministic production frontier and that any deviation from the frontier is due to inefficiency.

In this method the technical efficiency is identified as a proportional increase in the output vector with a given input vector. Therefore, the output-oriented measure of technical efficiency is the solution to the following constant returns to scale (CRS) DEA linear programming problem (Coelli et al., 1998):

$$\max_{\phi,\lambda} \phi,$$

$$\begin{aligned} s.t. \quad & -\phi y_i + Y\lambda \geq 0 \\ & x_i - X\lambda \geq 0 \\ & \lambda \geq 0 \end{aligned}$$

where ϕ is a scalar, and y_i and x_i are column vectors of outputs and inputs respectively for the ith school district. λ is an N x 1 vector of constants. The variable Y is an M x N output matrix and X is a K x N input matrix, and the proportional increase in outputs that could be achieved by the ith school district, holding inputs constant, is ϕ-1($1 \leq \phi < \infty$) and 1/ϕ is the school district's efficiency score which is between 0 and 1.

McCarty et al. (1993) suggest using efficiencies generated by DEA as dependent variables in a second stage with Tobit regression to assess the effects of variables not included in the first stage on technical efficiency. The efficiency estimates from the first stage are between 0 and 1, data is censored, and therefore Tobit regression, rather than OLS, is the appropriate method of estimation.

In order to obtain efficient parameter estimates, the possibility of the existence of heteroscedasticity in this stage should be considered and if in fact it exists, it should be incorporated into the model.

THE DATA SET

The data for the present study were obtained from the Oklahoma Department of Education, Office of Accountability for the academic years 1996-1997, 1997-1998, and 1998-1999. The data includes observations on several socioeconomic indicators (e.g., students eligible for the subsidized lunch program, parents' education level, family income, etc.). Students' performance measures are based on different standardized test scores appropriate for the different grades (e.g., ITBS, CRT, ACT scores) for over 600 school districts in Oklahoma.

Among the numerous measures of performance available, the Iowa Test of Basic Skills (ITBS) and Criterion Reference Test (CRT) are probably most reliable. Hanushek (1986) acknowledges that test scores are imperfect measures of educational output. However, test performance is used to allocate funds and evaluate educational programs. Test scores are also commonly available and appear to be valued by educators, as well as parents and decision makers as a measure of education efficacy. Here, ITBS for grades 3 (IT3) and 7 (IT7) and CRT for grades 5(CRT 5), 8 (CRT 8) and 11 (CRT 11) as well as ACT scores are used as measures of educational output. Other measures of performance for high school may be of interest. However, the available data on these measures are not consistently measured. For example, graduation rate is not measured adequately. According to Profiles (1998, p. xxvi) District Report, since Oklahoma does not have a statewide student identification system to monitor student migration, the graduation rate could be understated or overstated for all districts in the state. The average GPA of high school seniors has no uniform measure of grading; also, advanced placement (AP) participation rate and AP tests scoring college credit, suffer from an inadequate number of observations. Another interesting measure of performance is Oklahoma college freshmen taking at least one remedial course. However, observations are not consistently measured for the years under this study (1996-1999).

Elimination of *dependent* districts along with availability of data results in 354 observations for each of the years in the study. Summary statistics of the variables of interest (and their definitions) for the panel can be found in Table 1.

MODEL I SPECIFICATION

Model I is specified as:

$$\text{Score}_{it} = f\,(\text{YRSEXP}_{it}, \text{DEG}_{it}, \text{I}_{it}, \text{O}_{it}) \tag{1}$$

where output Score includes: IT3, IT7, CRT5, CRT8, CRT11and ACT. Following Kirjavainen (1998), teacher's education (DEG) and experience (YRSEXP) are included in the model as inputs. According to Kirjavainen, in statistical analysis teacher's education and experience are rarely found to have an impact on student achievement. However, they could affect efficiency distribution and efficiency ranking even though they are not traditional inputs. Instructional and non-instructional expenditures, I and O respectively, are the traditional inputs in the model.

TABLE 1
SUMMARY STATISTICS FOR OKLAHOMA SCHOOL DISTRICTS 1996-1999 (PANEL)

Variable	Mean	Std. Dev.	Minimum	Maximum
MIN	0.27	0.16	0.00	1.00
HHINCOME	21313.03	5753.18	10833.00	45790.00
PVALUATION	19982.70	14535.03	3639.00	172102.58
POVERTY	0.19	0.07	0.03	0.41
SED	0.12	0.03	0.04	0.26
LUNCH	.51	0.16	.04	1.00
STR	16.00	2.17	8.16	21.97
SALARY	30101.04	1081.49	27119.53	35332.96
DEG	0.33	0.13	0.04	0.81
YRSEXP	15.28	4.51	5.83	30.67
I	3010.35	519.93	2090.59	6310.28
O	1914.50	465.25	1099.58	5242.42
IT3	62.27	9.77	33.00	93.00
CRT5	65.36	11.97	30.67	100.00
IT7	55.70	8.19	30.00	85.00
CRT8	61.33	10.24	34.67	93.00
CRT11	70.05	9.87	29.00	94.25
ACT	19.88	1.39	15.40	23.70
ADM	1590.80	3826.49	148.53	41471.46

Where:

MIN	Percentage of minority students LUNCH
HHINCOME	Average household income (1990) ($)
PVALUATION	Assessed value of property within the boundaries of the district per student ($)
POVERTY	Poverty rate (1990)
DEGADULTS	Percentage of adults age 20+ with education beyond high school diploma (1990)
SED	Percentage of Students in Special education
LUNCH:	Percentage of students eligible for reduced cost or free lunch
STR	Student/Teacher ratio
SALARY	Average salary per full-time equivalent teacher ($)
DEG	Percentage of teachers with advanced degree
YRSEXP	Average experience of teachers (year)
I	Instructional expenditure per student ($)
O	Noninstructional expenditure per student ($), i.e., administrative and other expenses that are not directly used for instructional purposes.
IT3	ITBS for grade 3 (composite scores)
CRT5	CRT for grade 5 (average scores)
IT7	ITBS for grade 7 (composite scores)
CRT8	CRT for grade 8 (average scores)
CRT11	CRT for grade 11 (average scores)
ACT	Average ACT scores for all seniors in the district
ADM	Average daily membership (number of students)

Estimation

The model in equation (1) is estimated using the DEA method for the panel data (1996-1999). Thus the number of observations for the panel is N = 1062. The differences in efficiency scores of school districts generated by DEA could be explained by some variables not included in the DEA analysis (e.g., environmental variables). Efficiency may also be affected by the scale of operation (e.g., district size). In general, exogenous factors that affect output are built into the measure of technical efficiency (Kumbhakar, et al., 1991).

A linear model that accounts for these nontraditional inputs can be written as:

$$EFF_{it} = \alpha_0 + MIN_{it}\alpha_1 + LUNCH_{it}\alpha_2 + HHINCOME_{it}\alpha_3 + PVALUATION_{it}\alpha_4 + POVERTY_{it}\alpha_5 + DEGADULTS_{it}\alpha_6 + SED_{it}\alpha_7 + SALARY_{it}\alpha_8 + ADM_{it}\alpha_9 + ADM^2_{it}\alpha_{10} + STR_{it}\alpha_{11} + e_{it} \quad (2)$$

where EFF is the efficiency score generated by DEA and e is a random error term.

With the exception of SALARY, ADM and STR that are exogenous factors which affect output, the remainder of the variables in the efficiency equation are socioeconomic variables and are outside the control of the school districts. These variables are proxies for family influences.[4]

Results

DEA

The results of the DEA estimation are obtained using DEAP (2.1) software developed by T. J. Coelli and are presented in Table 2. Table 2 contains basic information on the distribution of efficiency scores generated by DEA under constant returns to scale (CRS) and variable returns to scale (VRS) assumptions. In DEA, under the VRS assumption, the possibility of scale of operation is considered and the efficiency measures are affected by it.

TABLE 2
SUMMARY STATISTICS FOR DEA EFFICIENCY SCORES, MODEL I (PANEL)

	CRS	VRS
Mean	.82	.91
SD	.11	.06
Minimum	.44	.71
Maximum	1	1

Efficiency differences among school districts under both CRS and VRS assumptions are quite considerable. The mean efficiency of 82 percent under the CRS assumption suggests an average inefficiency of 18 percent.

To investigate the number of school districts that fall within certain efficiency intervals, frequencies of school districts are grouped based on their efficiency scores. These frequencies are presented in Table 3.

Table 3 suggests that school districts have become more efficient each year under both CRS and VRS assumptions. Even so, in the 1998-1999 school year, only 102 districts have efficiency estimates of .9 and above under the CRS assumption.

TABLE 3
FREQUENCIES OF SCHOOL DISTRICTS IN CLASSES BASED ON EFFICIENCY SCORES OF THE DEA MODEL I

	1996-1997		1997-1998		1998-1999	
Efficiency Class (Range)	**CRS**	**VRS**	**CRS**	**VRS**	**CRS**	**VRS**
<.5	2	0	0	0	1	0
.5 - <.7	55	0	49	0	48	0
.7 - <.9	224	186	214	161	202	131
.9 - 1	73	168	90	192	102	222

Tobit Regression

In the second stage, the efficiency scores generated from CRS DEA for 1996-1999 are regressed on the right-hand side variables in equation (2) by the Tobit regression method, using LIMDEP (7.0) software.[5] The assumption of CRS does not seem to be a great draw back in this application since the results of Tables 2 and 3 suggest very little differences among efficiency scores obtained under the CRS and VRS. However, the CRS assumption in this stage is more appropriate since VRS could bias the efficiency scores upward (Coelli et al. 1998). In equation (2) school size and student/teacher ratio are explanatory variables that explain the effect of non-optimal scale of operation, if any, on the efficiency differences obtained under the CRS assumption (Kirjavainen et al., 1998, p. 388).

The possibility of existence of heteroscedasticity in the second stage is considered. Using the "Tobit Heterscedasticity" option in LIMDEP allows one to consider variables that may be the source of this misspecification error.[6] All the explanatory variables as well as the dependent variable in equation (2) are considered. Except for DEGADULTS, LUNCH, and SALARY; the coefficients for all variables are statistically significant at the 5 percent level and thus these variables are likely sources of heteroscedasticity. To test this hypothesis, the likelihood ratio test for heteroscedasticity was performed. The results suggests that H_0 should be rejected (test statistic $\lambda = 197.108 \sim \chi^2_{11}$, critical χ^2_{11} = 19.675 at α = .05), therefore there is substantial evidence that at least one of the variables "explains" the existence of heteroscedasticity in the Tobit regression. The results of Tobit regression under the assumption of heteroscedasticity are presented in Table 4.

TABLE 4
TOBIT REGRESSION COEFFICIENT ESTIMATES OF THE EFFICIENCY MODEL I
DEPENDENT VARIABLE: EFFICIENCY ESTIMATES FROM THE FIRST-STAGE DEA MODEL UNDER CRS ASSUMPTION (PANEL)

Variable	Coefficient	t-statistic
Constant	1.35	15.55
MIN	-.15	-8.54*
LUNCH	-.14	-5.47*
HHINCOME	.02E-4	2.72*
PVALUATION	-.01E-4	-3.93*
POVERTY	-.10	-1.66
DEGADULTS	.22	4.36*
SED	-.36	-4.11*
SALARY	-.22E-4	-9.25*
ADM	-.02E-4	-2.15*
ADM^2	.00	1.76
STR	.01	10.38*

*Significant at the .05 level

Based on the results of Table 4, except for the assessed property value per student (PVALUATION), all of the coefficients of environmental variables over which school districts have no control, have the correct sign and, except for POVERTY, are statistically significant. One possible reason for the negative sign on PVALUATION is that it includes all types of commercial as well as residential properties in the school districts and therefore, districts with high property valuation could potentially have low income families.[7] These results are consistent with previous studies which suggest that school districts heavily populated by students from a less advantage family environment are more likely to be less efficient (Adkins and Moomaw, 1997). The effect of the remaining variables in the second stage on the efficiency is as follows:

First, the size of the school districts as measured by ADM has a negative effect on efficiency; second, the student/teacher ratio has a positive relationship with efficiency; and finally, the effect of teachers' salary on efficiency is negative.

To assess the magnitude of the effect of the explanatory variables on efficiency, the marginal effects of these variables under the assumption of heteroscedasticity is computed and presented in Table 5.

The results of Table 5 suggest that a one percent increase in MIN, LUNCH, and SED decreases the efficiency by almost .16, .15, and .41, respectively. A one unit increase in DEGADULTS and STR increases efficiency by almost .24 and .01, respectively. The effects of HHINCOME, POVERTY, and SALARY on efficiency are negligible. Also, school district size, as measured by ADM, does not seem to have a strong effect on efficiency, which is consistent with Kirjavainen, et al. (1998).

To determine the degree of robustness of this model a second model with a new specification was estimated, using the same estimation method as Model I. Then, the results of the two models are compared.[8]

TABLE 5
TOBIT SLOPE (MARGINAL EFFECT) ESTIMATES OF THE EFFICIENCY MODEL I (PANEL)

Variable	Slope	t-statistic
Constant	1.34	14.67
MIN	-.15	-9.25
LUNCH	-.14	-6.08*
HHINCOME	.02E-4	3.29*
PVALUATION	-.01E-4	-4.39*
POVERTY	-.08	-1.26
DEGADULTS	.23	4.63*
SED	-.41	-4.71*
SALARY	-.23E-4	-9.36*
ADM	-.02E-4	1.70
ADM^2	.00	1.75
STR	.02	10.89*

* Significant at .05 level

MODEL II SPECIFICATION

In the first stage, Model II includes the traditional inputs only:

$$Score_{it} = f\,(I_{it}, O_{it}) \qquad (3)$$

where outputs; Score, and inputs; I, O are as defined in equation (1). YRSEXP and DEG are included in the second stage. The model is estimated using DEA in the first stage and the Tobit regression method in the second stage.

Results

DEA

The results of the DEA estimation are presented in Table 6. The table contains the basic information on the distribution of the efficiency scores generated by DEA under CRS and VRS assumptions.

TABLE 6
SUMMARY STATISTICS FOR DEA EFFICIENCY SCORES, MODEL II (PANEL)

	CRS	VRS
Mean	.75	.89
SD	.12	.06
Minimum	.33	.68
Maximum	1	1

Efficiency differences among school districts under both CRS and VRS assumptions are quite considerable. The mean efficiency of 75 percent under the CRS assumption suggests an average inefficiency of 25 percent and under the VRS assumption the average efficiency of almost 89 percent suggests an average inefficiency of 11 percent.

Tobit Regression

The efficiency equation estimated in the second stage using the Tobit regression method is equation (2) including YRSEXP and DEG as explanatory variables:

$$EFF_{it} = \alpha_0 + MIN_{it}\alpha_1 + LUNCH_{it}\alpha_2 + HHINCOME_{it}\alpha_3 + PVALUATION_{it}\alpha_4 + POVERTY_{it}\alpha_5 + DEGADULTS_{it}\alpha_6 + SED_{it}\alpha_7 + SALARY_{it}\alpha_8 + ADM_{it}\alpha_9 + ADM^2_{it}\alpha_{10} + STR_{it}\alpha_{11} + YRSEXP_{it}\alpha_{12} + DEG_{it}\alpha_{13} + e_{it} \quad (4)$$

The possibility of the existence of heteroscedasticity in the second stage was also considered. The dependent variable, as well as all the explanatory variables in equation (4), is considered as the possible source of this misspecification. Except for SALARY, YRSEXP, and DEG all of the variables are likely sources of heteroscedasticity. This hypothesis is tested (likelihood ratio test) and the results suggests that H_0 should be rejected (test statistic $\lambda = 105.426 \sim \chi^2_{10}$, critical $\sim \chi^2_{10} = 18.307$ at $\alpha = .05$). Therefore, there is substantial evidence that at least one of the variables "explains" the existence of heteroscedasticity in the Tobit regression. The Tobit coefficient estimates computed under the assumption of heteroscedastic error terms in the model are computed and presented in Table 7.

TABLE 7
TOBIT REGRESSION COEFFICIENT ESTIMATES OF THE EFFICIENCY MODEL II
DEPENDENT VARIABLE: EFFICIENCY ESTIMATES FROM THE FIRST STAGE DEA MODEL UNDER CRS ASSUMPTION (PANEL)

Variable	Coefficient	t-statistic
Constant	.81	10.05
MIN	-.10	-5.75*
LUNCH	.22	10.04*
HHINCOME	.01E-4	1.75
PVALUATION	-.02E-4	-8.04*
POVERTY	-.13	-2.41*
DEGADULTS	.33	7.31*
SED	-.29	-3.73*
SALARY	-.06E-4	-2.35*
ADM	-.02E-4	-.59
ADM^2	.00	.31
STR	.02	14.22*
YRSEXP	-.01E-4	-3.10*
DEG	.03E-4	.26

* Significant at .05 level

The results in Table 7 suggest that, MIN, LUNCH, POVERTY, DEGADULTS, SED, and STR are the only variables whose effects on efficiency are significant in terms of their magnitude. To examine the magnitude of these effects, the marginal effects of these variables on efficiency, based on the heteroscedastic Tobit model, are computed and presented in Table 8.

TABLE 8
TOBIT SLOPE (MARGINAL EFFECT) ESTIMATES OF THE EFFICIENCY MODEL II (PANEL)

Variable	Slope	t-statistic
Constant	.81	10.04
MIN	-.099	-5.77*
LUNCH	-.22	-10.04*
HHINCOME	.01E-4	1.76
PVALUATION	-.02E-4	-8.05*
POVERTY	-.13	-2.41*
DEGADULTS	.33	7.33*
SED	-.29	-3.74*
SALARY	-.06E-4	-2.35*
ADM	-.02E-4	-.59
ADM^2	.00	.31
STR	.02	14.24*
YRSEXP	-.01E-1	-3.10*
DEG	.03E-3	.26

*Significant at .05 level

Recall that minority students (MIN), students eligible for reduced or free lunch (LUNCH), poverty rate (POVERTY), students in special education (SED), and adults age 20+ with education beyond high school diploma (DEGADULTS) are measured in terms of percentages. Thus, the results of Table 8 suggest that a one percent increase in each of MIN, LUNCH, POVERTY, and SED decreases efficiency by .1, .23, .13, and .29, respectively; and a one percent increase in DEGADULTS increases efficiency by .33. In addition, for each unit increase in student/teacher ratio (STR), efficiency increases by .02. This is consistent with Kirjavainen, et al. (1998).

MODEL I VS. MODEL II

Comparison of the results of Model I and Model II suggests that the average efficiency scores in Model I are higher than that of Model II. This is expected, as Model I has more variables in the first stage (Kirjavainen et al., 1998).

As for the second stage Tobit regression results, both models suggest that the environmental variables which school districts have no control over, such as percentage of minority students (MIN), percentage of students eligible for reduced or free lunch (LUNCH), and percentage of students in special education (SED) have a strong negative effect and percentage of adults age 20+ with education beyond a high school diploma in the household (DEGADULTS) has a strong positive effect on efficiency of the school districts. Variables like teachers' salary (SALARY), teachers' years of experience (YRSEXP), teachers holding advanced degrees (DEG), and school size (ADM) which are under the control of school districts are clearly insignificant in explaining the variation in efficiencies among school districts. The student/teacher ratio (STR) affects efficiency positively; however, the relationship is not strong. The optimal school district size, as measured by ADM, is computed to be around 21,000 in both models. This result is consistent with Adkins, Moomaw (1997).

The efficiency rankings based on DEA CRS for Models I and II as well as the Spearman Rank Correlation coefficient between the two models are computed. The correlation coefficient is .81, which suggests that there are rather small differences in the efficiency ranking between the two models.[9]

CONCLUSIONS

This study uses two different models (empirical specification) and estimates the efficiency in the production of education in light of possible heteroscedasticity in the error term. The existence of heteroscedasticity in the data is supported based on hypothesis tests. In addition to the problem of heteroscedasticity, since the model consists of multiple outputs, the existing literature suggests the use of distance functions, which allow for multiple outputs, rather than parametric frontier functions. Thus, the non-parametric approach to the estimation of efficiency is employed, i.e., the DEA approach.

DEA suffers from a lack of well-known statistical properties and is not therefore very useful in answering questions regarding whether money matters. In addition, the production function is not parameterized and it yields no estimates of the various spending elasticities. To overcome these shortcomings a second-stage Tobit regression, which has well known statistical properties, was employed to explain the effects of variables such as teacher salary (SALARY), teacher years of experience (YRSEXP), teachers holding advanced degree (DEG), size of school district (ADM), student/teacher ratio (STR) etc. on the efficiency scores generated by the DEA model. Tobit regression is appropriate since the efficiency scores (dependent variable) are between 0 and 1. Heteroscedasticty is accounted for in the Tobit regression to ensure that the efficient coefficient estimates of the variables are obtained.

The data set includes observations on several input and output measures (e.g., teacher salary, the size of the district, standardized test scores etc.) for 354 independent (K-12) school districts in the state of Oklahoma. The time period under consideration is the 1996-97 through the 1998-99 academic school years.

The results of the two models are compared to examine their robustness. In general, the coefficient estimates in both models are consistent with the expected hypotheses. Therefore, for the most part, this study supports the results of past studies in that; socioeconomic factors are the primary reasons for the variation in the efficiency of the Oklahoma school districts. However, the estimates obtained in this study may be more reliable than those of past studies, which, generally, were based on the mean response function, single period data and/or the single output assumption.

ENDNOTES

1. The Baumol and Becker (1995) criticism that the poor quality of learning accomplishment signals the lack of rigorous curriculum and lack of sufficient rewards for learning.
2. For a detail discussion of advantage and disadvantage of DEA see Rassouli-Currier (forthcoming 2007).
3. Farrell (1957) offers a detail explanation of the shortcomings associated with the use of average functions.
4. The literature generally suggests the expectation of a positive effect on efficiency for favorable environmental variables and negative otherwise. As for teacher characteristics variables, the results in the literature are mixed.
5. Time dummies were statistically insignificant and thus were excluded from the model.
6. LIMDEP uses "White-Corrected" covariance matrix" (hc0) to correct for heteroscedasticity.
7. For example Frontier District with PVALUATION of $171,754 in the 1996-1997 school year had an average household income of $18,816 and a poverty rate of about 31%. Other examples with high PVALUATION, relatively low household income and relative high poverty rate include Laverne, Taloga, Medford and Timberlake school districts.
8. Two models are estimated to see if the sensitivity of DEA with respect to changes in inputs is small.
9. The efficiency rankings are provided by the author upon request.

REFERENCES

Adkins, L.C. and Moomaw, R.L. "On the Determinants of Technical Efficiency in Oklahoma Schools: A Stochastic Frontier Approach," Oklahoma State University Working Paper, 1997.

Baumol, W.J. and Becker, W.G. *Assessing Educational Practices*, The MIT Press, Cambridge, MA, 1995.

Bishop, J.H. "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review*, 1989, 79(1), pp. 178-197.

Chakraborty, K., Biswas, B. and Lewis W.C. "Measurement of Technical Efficiency in Public Education: A Stochastic and Nonstochastic Production Function Approach," *Southern Economic Journal*, 2001, 6 (4), pp. 889-905.

Charnes, A., Cooper W. and Rhodes E. "Measuring the Efficiency of decision Making Units," *European Journal of Operation Research*, 1978, 2, pp.429-444.

Coelli, T.J. "A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program," CEPA Working Paper 96/08, Department of Econometrics, University of New England, Armidale, 1996.

Coelli, T.J., Rao, D.S.P. and Battese G.E. *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, 1998.

Coleman, J.S. et al. *Equality of Educational Opportunity: Summary Report*, U.S. Government Printing Office, Washington, D.C. 1966.

Farrell, M.J "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, 1957, 120(3), pp. 253-281.

Hanushek, E.A. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *The Journal of Human Resources*, 1979, 14(3), pp. 351-388.

Hanushek, E.A. "The Economics of Schooling: Production and Efficiency in the Public Schools," *Journal of Economic Literature*, 1986, 24, pp. 1141-1177.

Hanushek, E.A. "Expenditures, Efficiency, and Equity in Education: The Federal Government's Role," *American Economic Association Papers and Proceedings*, 1989, 79(2), pp. 46-51.

Hanushek, E.A. "School Resources and Student Performance," in Gary Burtless, ed., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Washington, DC: Brookings Institution Press, 1996, pp. 43-73.

Kirjavainen, T. and Loikkanen, H.A. "Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit Analysis," *Economics of Education Review*, 1998, 17(4), pp. 377-393.

Kumbhakar, S.C., Ghosh, S and McGuckin J.T. "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms," *Journal of Business and Economic Statistics*, 1991, 9, pp. 379-386.

McCarthy, T.A. and Yaisawarng, S. "Technical Efficiency in New Jersey School Districts," in Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press: New York, 1993.

Office of Accountability. *Profiles 1998 District Report*, Oklahoma Educational Indicators Program, Oklahoma City, OK: State Department of Education Printing Services, April 1999.

Rassouli-Currier, S." The Choice of Estimation Method and Its effect on Efficiency Measurement in Public Education: Stochastic Frontier Regression vs. Data Envelopment Analysis" *The Journal of Economics*, 2007(forthcoming).

Taylor, L.L. "An Economy at Risk? The Social Cost of School Inefficiency," *American Economic Review*, 1994, pp.1-13.