1999

Technical Efficiency in Stochastic Production Frontier: A Simultaneous Equation Approach

Kalyan Chakraborty
Utah State University

Basudeb Biswas
Utah State University

W. Cris Lewis
Utah State University

Economic Research Institute Study Paper

ERI #99-10

TECHNICAL EFFICIENCY IN STOCHASTIC PRODUCTION FRONTIER: A SIMULTANEOUS EQUATION APPROACH

by

KALYAN CHAKRABORTY

School of Business
Emporia State University
Emporia, KS 66801-5087

BASUDEB BISWAS

W. CRIS LEWIS

Department of Economics
Utah State University
3530 Old Main Hill
Logan, UT 84322-3530

September 1998

TECHNICAL EFFICIENCY IN STOCHASTIC PRODUCTION FRONTIER: A SIMULTANEOUS EQUATION APPROACH

Kalyan Chakraborty, Assistant Professor

School of Business
Emporia State University
Emporia, KS 66801-5087

Basudeb Biswas, Professor
W. Cris Lewis, Professor

Department of Economics
Utah State University
3530 Old Main Hill
Logan, UT 84322-3530

Information on other titles in this series may be obtained from: Department of Economics, Utah State University, 3530 Old Main Hill, Logan, Utah 84322-3530.

TECHNICAL EFFICIENCY IN STOCHASTIC PRODUCTION FRONTIER: A SIMULTANEOUS EQUATION APPROACH

Kalyan Chakraborty, Basudeb Biswas, and W. Cris Lewis

ABSTRACT

This paper uses both the stochastic and nonstochastic production function approach to measure technical efficiency in public education in Utah. The stochastic specification estimates technical efficiency assuming half normal and exponential distributions. The nonstochastic specification uses two-stage DEA to separate the effects of fixed inputs on the measure of technical efficiency. The empirical analysis shows substantial variation in efficiency among school districts. While these measures are insensitive to the specific distributional assumptions about the one-sided component of the error term in the stochastic specification, they are sensitive to the treatment of fixed socioeconomic inputs in the two-stage DEA.

JEL codes: D20, R30, C14

TECHNICAL EFFICIENCY IN STOCHASTIC PRODUCTION FRONTIER: A SIMULTANEOUS EQUATION APPROACH

1. Introduction

Efficiency in the public education system is a significant issue in the United States. Nationwide, real expenditure per student in public education increased over 3 percent per year between 1960 and 1998, but output as generally measured by standardized test scores has not increased and in some cases (e.g., the verbal SAT score) has declined.[1] One explanation is that resources are not being utilized efficiently. There may be productive or technical inefficiency and/or allocative or price inefficiency (i.e., given the relative prices of inputs, the cost minimizing input combination is not used). This paper focuses on the former by evaluating technical inefficiency in public education using data from Utah school districts.

The pioneering work by Farrell in 1957 provided the definition and conceptual framework for both technical and allocative efficiency. While technical efficiency refers to failure to operate on the production frontier, allocative efficiency generally refers to the failure to meet the marginal conditions for profit maximization. Considerable effort has been made in refining the measurement of technical efficiency. The literature is broadly divided into deterministic and stochastic frontier methodologies.[2] The deterministic nonparametric approach that developed out of mathematical programming is commonly known as data envelopment analysis (DEA), and the parametric approach that estimates technical efficiency within a stochastic production, cost, or profit function model is called the stochastic frontier method.

Both approaches have advantages and disadvantages as discussed in Forsund et al. (1980). DEA has been used extensively in measuring efficiency in the public sector, including education, where market prices for output generally are not available. For example, Levin (1974), Bessent and Bessent (1980), Bessent et al. (1982), and Fare et al. (1989) used this method to estimate efficiency in public education. The stochastic frontier methodology was used by Barrow (1991) to estimate a stochastic cost frontier using data from schools in England. Wyckoff and Lavinge (1991) and Cooper and Cohn (1997) estimated technical efficiency using school district data from New York and South Carolina, respectively. Grosskopf et al. (1991) used the parametric approach to estimate allocative and technical efficiency in Texas school districts.

The recent literature has seen a convergence of the two approaches and their complementarity is being recognized.[3] However, there is a lack of empirical evidence in the literature about the proximity of these two approaches in measuring technical efficiency. Policy formulations based on only one of these efficiency estimates may not be accurate because of the inherent limitations of each. Before any correctional measures are taken, the stability of the technical efficiency estimates obtained from a parametric method should be evaluated by comparing them against those found when using the nonparametric method.

In this study the technical efficiency estimates for each school district using the stochastic frontier method and Tobit residuals from the two-stage DEA model are compared. In the two-stage DEA model, technical efficiency scores obtained from DEA using controllable inputs are regressed on student socioeconomic status and other environmental factors. The residuals of this regression measure pure technical efficiency after accounting for fixed socioeconomic and environmental factors.

The empirical analysis uses data from the 40 school districts in Utah for the academic year 1992-93. The standardized test score for 11th grade students is used as a measure of school output, and two classes of inputs are included. The first class is considered to be subject to control by school administrators and includes the student-teacher ratio, the percentage of teachers having an advanced degree, and the percentage of teachers with more than 15 years of experience. The second class includes such uncontrollable factors as socioeconomic status, education level of the local population, and net assessed real property value per student.

This paper is organized as follows. First, the relevant literature is reviewed and then a definition of the educational production function is provided. Next, the stochastic and DEA specifications of technical inefficiency are reviewed. Finally, the data set is discussed and the empirical results are presented.

2. Background

For a given technology and a set of input prices, the production frontier defines the maximum output forthcoming from a given combination of inputs. Similarly, the cost frontier defines the minimum cost for providing a specified output rate given input prices, and the profit frontier defines the maximum profit attainable given input and output prices. Inefficiency is measured by the extent that a firm lies below its production and profit frontier and above its cost frontier. Koopmans (1951) defines a technically efficient producer as one that cannot increase the production of any one output without decreasing the output of another product or without increasing some input. Debreu (1951) and Farrell (1957) offer a measure of technical efficiency as one minus the maximum equiproportionate reduction in all inputs that still allows continuous production of a given output rate (Lovell 1993).

An early study that measured technical inefficiency in education production is that by Levin (1974, 1976). He used the Aigner and Chu (1968) parametric nonstochastic linear programming model to estimate the coefficients of the production frontier and found that parameter estimation by ordinary least squares (OLS) does not provide correct estimates of the relationship between inputs and output for technically efficient schools; that technique only determines an average relationship. Klitgaard and Hall (1975) used OLS techniques to conclude that the schools with smaller classes and better paid and more experienced teachers produce higher achievement scores. Their study also estimates an average relationship rather than an individual school specific relationship between inputs and output.

Among the studies on technical efficiency in public schools using the DEA method, one of the earliest was done by Charnes, Cooper, and Rhodes (1978), who evaluated the efficiency of individual schools relative to a production frontier. Bessent and Bessent (1980) and Bessent et al. (1982), made further refinements by incorporating a nonparametric form of the production function, introducing multiple outputs, and identifying sources of inefficiency for an individual school. Further extensions were made by Ray (1991) and McCarty and Yaisawarng (1993), who considered controllable inputs in the first stage of the DEA model to measure technical efficiency. Then the environmental (i.e., noncontrollable) inputs were used as regressors in the second stage using OLS or a Tobit model, and the residuals were analyzed to determine the performance of each school district.

In these studies, it is postulated that all firms have an identical production frontier that is deterministic, and any deviation from that frontier is attributable to differences in efficiency. The concept of a deterministic frontier ignores the possibility that a firm's performance may be affected by factors both within and outside its control. That is, combining the effects of any measurement error with other sources of stochastic variation in the dependent variable in the single one-sided error term may lead to biased estimation of technical inefficiency. In response to this, the concept of a stochastic production frontier was developed and extended by Aigner, Lovell, and Schmidt (1977), Meeusen and van den Broeck (1977), Battese and Corra (1977), Battese and Coelli (1988), Lee and Tyler (1978), Pitt and Lee (1981), Jondrow et al. (1982), Kalirajan and Flinn (1983), Bagi and Huang (1983), Schmidt and Sickles (1984), and Waldman (1984). The basic idea behind the stochastic frontier model as stated by Forsund, Lovell, and Schmidt (1980) is that the error term is composed of two parts: (1) the systematic component (i.e., a traditional random error) that captures the effect of measurement error, other statistical noise, and random shocks; and (2) the one-sided component that captures the effects of inefficiency.

Frontier production models have been analyzed either within the framework of the production function or by using duality in the form of a cost minimizing or profit maximizing framework. Barrow's (1991) study of schools in England tested various forms of the cost frontier and found that the level of efficiency was sensitive to the method of estimation. In their study of technical inefficiency in elementary schools in New York, Wyckoff and Lavinge (1991) estimated the production function directly and found that the index of technical inefficiency depends on the definition of educational output. For example, if output is measured by the level of cognitive skill of students rather than their college entrance test score, the index of technical inefficiency based on each output measure will be different. Grosskopf et al. (1991) used a stochastic frontier and distance function to measure technical and allocative efficiency in Texas school districts and concluded that they were technically efficient but allocatively inefficient.

3. Defining an Educational Production Function

In the production of education, school districts use various school and nonschool inputs to produce multiple outputs that are assumed to be measurable by achievement test scores. As one purpose of education is to develop the student's basic cognitive skills, these abilities often are measured by the scores in reading, writing, and mathematics tests. However, there are references in the literature where output is measured either by the number of students graduating per year, student success in gaining admission to institutions of higher education, or a student's future earning potential. In most of the studies of the education production function, the measure of output is limited by the availability of data. School inputs that are associated with achievement scores are typically measured by the student-teacher ratio, the educational qualifications of teachers, teaching experience, and various instructional and noninstructional expenditures per student. Nonschool inputs include socioeconomic status of the students and other environmental factors that influence student productivity. While family income, number of parents in the home, parental education, and ethnic background measure the socioeconomic status of the students; geographic location (e.g., rural vs. urban) and net assessed value per student often are used to capture the environmental factors.

School inputs that are basically associated with the instructional and non-instructional activities are under the control of the school management. Most studies in educational production find an insignificant relationship between most of the school inputs and outputs. In contrast, Walberg et al., (1987), Hanushek (1971; 1986), Deller and Rudnicki (1993), Cooper and Cohn (1997), and Grosskopf et al. (1989) find that socioeconomic and environmental factors significantly affect achievement scores.

A school district is technically efficient if it is observed to produce the maximum level of output from a given bundle of resources used or, conversely, uses minimum resources to produce a given level of output. In this study, output of the educational production function is measured by the average test score of the 11th graders on a standardized battery test. The use of a single output production technology to estimate efficiency in stochastic frontier models is somewhat restrictive in the sense that measures of efficiency are sensitive to the selection of output (Wyckoff and Lavinge 1991). An indirect approach to compare the performances of the school districts would be to estimate a cost frontier; however, that requires data on input prices that generally are not available for production in education.

4. Stochastic Specification of Technical Efficiency

In the stochastic frontier model, a nonnegative error term representing technical inefficiency is subtracted from the traditional random error in the classical linear model. The general formulation of the model is:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_k x_{ik} + \varepsilon_i,$$

where y_i is output and the x_j are inputs. It is postulated that $\varepsilon_i = v_i - u_i$ where $v_i \sim N(0,\sigma_v^2)$ and $u_i \sim \left| N(0,\sigma_u^2) \right|$, $u_i \geq 0$, and the u_i and v_i are assumed to be independent. The error term (ε_i) is the difference between the standard white-noise disturbance (v_i), and the one-sided component (u_i). The term v_i allows for randomness across firms and captures the effect of measurement error, other statistical noise, and random shocks outside the firm's control. The component u_i captures the effect of inefficiency (Forsund et al. 1980).

Most of the earlier stochastic production frontier studies only estimated mean technical inefficiency of firms because the residual for individual observations could not be decomposed into the two components. Jondrow et al. (1982) solved the problem by defining the functional form of the distribution of the one-sided inefficiency component and deriving the conditional distribution of $[u_i \mid v_i - u_i]$ for two popular distribution cases (i.e., the half normal and exponential) to estimate firm-specific technical inefficiency.[4]

For this study, let the production function for the *i*th school district be represented by:

$$y_i = A\prod_{j=1}^{k} x_j^{\alpha_j} e^{v}, \tag{1}$$

where y_i is output, and x_j are exogenous inputs. A is the efficiency parameter and v is the stochastic disturbance term. The production function in (1) is related to the stochastic frontier model by Aigner, Lovell, and Schmidt (1977) who specify A as:

$$A = a_0 e^{-u} \quad u \geq 0,$$

where a_0 is a parameter common to all districts and u is the degree of technical inefficiency that varies across school districts. Units for which $u = 0$ are most efficient. A district is said to be technically inefficient if output is less than the maximum possible rate defined by the frontier. The term v is the usual two-sided error term that represent shifts in the frontier due to favorable and unfavorable external factors and measurement error.

After including the component of inefficiency (i.e., e^{-u}), the actual production function is written as:

$$y_i = a_o \prod_{j=1}^{k} x_j^{\alpha_j} e^{(v-u)}. \tag{2}$$

If there is no inefficiency and potential output is denoted by Y, then the production function is written as:

$$Y_i = a_o \prod_{j=1}^{k} x_j^{\alpha_j} e^v .$$

Hence, the appropriate measure of technical efficiency is:

$$\frac{\text{actual output}}{\text{potential output}} = \frac{y_i}{Y_i} = \frac{a_o e^{-u} \prod_{j=1}^{k} x_j^{\alpha_j} e^v}{a_o \prod_{j=1}^{k} x_j^{\alpha_j} e^v} = e^{-u} .$$

Potential output is the maximum possible when $u = 0$ in equation (2). A technically efficient school district produces output (i.e., standardized test scores) that are on the stochastic production frontier that is subject to random fluctuations captured by v. However, because of differences in managerial efficiency, actual performance deviates from the frontier.

Since $u \geq 0$, $0 \leq e^{-u} \leq 1$, and e^{-u} is a measure of technical efficiency, the mean technical efficiency is E(e^{-u}). Thus, technical inefficiency is measured by 1-e^{-u} where e^{-u} is a measure of technical efficiency bounded by 0 and 1. That is, technical efficiency lies between 1 and 0. This study uses the method of estimation suggested by Jondrow et al. (1982) to estimate technical inefficiency in each school district.

5. DEA Specification of Technical Efficiency

The data envelopment analysis (DEA) approach constructs the best practice production frontier as a piecewise linear envelopment of the available data on all producers in such a manner that all observed points lie on or below the frontier. In this construct, the performance of a producer is evaluated in terms of his ability to either reduce an input vector or expand an output vector subject to the restrictions imposed by the best-observed practice. This measure of performance is relative in the sense that efficiency in each school district is evaluated against the most efficient district and measured by the ratio of maximal potential output to actual observed output. The major advantage of DEA is that it is capable of modeling multi-output multi-input technologies. It is assumed that a school district converts various instructional and noninstructional inputs into multiple learning outputs measured as students' achievement test scores in reading, writing, language, science, social science, and mathematics. Hence, measuring technical efficiency based on a single output production technology such as the stochastic frontier approach might be inadequate. A simple output-oriented DEA model is presented in this section; for a detailed methodological discussion readers are referred to Seiford and Thrall (1990), Lovell (1993), and Fare, Grosskopf, and Lovell (1994).

Assume there are K school districts using N inputs, i.e., $x = (x_1,\ldots,x_N) \in \Re_+^N$, and producing M outputs denoted by $y = (y_1,\ldots,y_M) \in \Re_+^M$. **N** is a (N, K) matrix of observed inputs; **M** is a (M, K) matrix of outputs of K different school districts; and (x^k, y^k) represents the input output vector or the activity of the *k*th district. Assuming inputs and outputs are nonnegative, the piecewise linear output reference set satisfying the properties of constant returns to scale and strong disposability of inputs and outputs (C, S), can be formed from **N** and **M** as:

$$P(x \mid C,S) = \{y : y \le z\mathbf{M}, z\mathbf{N} \le x, z \in \Re_+^K\}, x \in \Re_+^N$$

where z is the (1, K) vector of intensity variables identifying the extent that a particular activity (x^k, y^k) is utilized. The assumption of strong disposability of inputs and outputs as a feature of technology implies that the same input vector can produce a lower output rate and a higher input vector can produce the same rate of output. Given the (C, S) technology in the above specification, an output measure of technical efficiency for activity k is the solution to the linear programming problem:

$$F_o(x^k, y^k \mid C, S)^{-1} = \text{Max}_{\theta}\ z^{\theta}$$

$$\text{s.t.} \qquad \theta y^k \leq z\text{M}$$

$$z\text{N} \leq x^k$$

$$z \in \Re_+^K$$

or

$$F_o(x^k, y^k \mid c, s)^{-1} = \text{Max}_{\theta}\ z^{\theta}$$

$$\text{s.t.} \qquad \theta y_{km} \leq \sum_{k=1}^{K} z_k y_{km},\ m = 1,2,\ldots,M,$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{kn},\ \ n = 1,2,\ldots,N,$$

$$z_k \geq 0,\ k = 1,2,\ldots,K.$$

In an output-oriented DEA model, technical efficiency is measured by the reciprocal of the output distance function, which is obtained by maximizing θ subject to the restriction imposed by the assumptions of input and output disposability and returns to scale. Hence, $F_o(x^k, y^k \mid C, S)^{-1} = 1$ implies the district k is the most efficient and lies on the frontier, and any value less than unity implies that a district is operating below the frontier. The technical efficiency score measures the extent that the output vector may be increased given the combination of input vectors. The assumption of constant returns to scale is replaced by variable returns to scale (V, S) with the following restrictions on the intensity vector, as $\sum_{k=1}^{K} z_k = 1$.

The output-oriented DEA measure of technical efficiency seeks the maximum proportionate increase in output given inputs while remaining on the same production frontier. Hence, this method assumes that outputs are capable of expansion. For the educational production function, inputs measuring student socioeconomic status and environmental factors are fixed and beyond the control of the school. Hence, technical efficiency estimates from DEA using these inputs along with other controllable inputs will lead to specification error. One solution to this problem is to use a subvector efficiency model (Fare et al. 1994) that specifically treats the environmental factors as fixed;[5] an alternative is to use the conventional two-stage DEA model. Following McCarty and Yaisawarng (1993), a two-stage DEA model is used where efficiency scores from output-oriented DEA using controllable school inputs only are regressed on the nonschool inputs (i.e., socioeconomic and environmental factors) using a Tobit regression model. The residuals of the Tobit model separate the effects of these fixed factors and measure pure technical efficiency that is bounded between $-\infty$ and 1. Hence, the higher the value of the residual, the better is the performance of the school district.

6. The Data Set

Relevant data for the 40 school districts in Utah were collected from reports prepared by the Utah State Office of Education (1992-93) and the Utah Education Association (1993). The single output of the educational production function (y) is the battery test score in the 11th grade, a composite of reading, writing, and mathematics skills.[6] The average district level data are aggregated over schools and over students. The school inputs used in this study are: the student-teacher ratio (x_1), percentage of teachers with an advanced degree (x_2), and the percentage of teachers with over 15 years of experience (x_3). Nonschool inputs consist of the percentage of students who qualify for Aid to Families with Dependent Children (AFDC) subsidized lunch (x_4), percentage of district population having completed high school (x_5), and net assessed value per student (x_6). While x_1 is a proxy for the level of instructional input, x_2 and x_3 measure quality of teaching inputs, and x_4, x_5, and x_6 measure the socioeconomic status and the environmental factors. In the single equation model, the first three inputs (x_1, x_2, x_3) are subject to control by management whereas inputs x_4 through x_6 are exogenous. Summary statistics for both inputs and output are reported in Table 1.

Measuring technical efficiency in the output-oriented DEA model uses the same inputs as in the stochastic frontier model; however, the first stage of the DEA model uses only controllable inputs (x_1, x_2, and x_3), while the second stage Tobit regression uses the uncontrollable inputs (x_3, x_4, and x_5). Following Schmidt and Lovell (1979) and Battese and Coelli (1988), a Cobb-Douglas functional form of the production function is postulated.[7] This function in log linear form is:

$$\ln y_i = \alpha_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_5 \ln x_5 + \beta_6 \ln x_6 + v - u$$

where y_i is the educational output (i.e., average test score), the x_j are the inputs described above, and $v_i \sim N(0,\sigma_v^2)$ and $u_i \sim |N(0,\sigma_u^2)|$. The condition that $u_i \geq 0$ allows production to occur below the stochastic production frontier.

The following relationships between output and each explanatory variable are hypothesized:

Variable	Coefficient	Hypothesized Sign
Student-teacher ratio	β_1	< 0
Percentage of teachers with advanced degree	β_2	> 0
Percentage of teachers with over 15 years of experience	β_3	> 0
Percentage of students receiving subsidized lunch	β_4	< 0
Percentage of population with high school education	β_5	> 0

Net assessed value per student	β_6	> 0

7. Empirical Results

Maximum-likelihood estimates[8] of the parameters based on half normal and exponential distributions of u are reported in Table 2. Except for the net assessed value per student, all the coefficients have the correct sign, but only the coefficient of the percentage of population with a high school education is significant at the 0.05 or lower level. One possible reason for a negative sign on net assessed value per student input is multicollinearity with other socioeconomic inputs. The highly significant coefficient on the education level of the district population implies a 1 percent change in population with a high school diploma is associated with a 0.91 to 0.96 percent change in test score. This indicates the importance of the environment for learning provided in the home. The negative sign on the student-teacher ratio is as expected and confirms the conventional wisdom that smaller classes are more conducive to better learning. Positive coefficients on the advanced degree and experience variables indicate positive contributions of these inputs in the learning process. Finally, the welfare variable has the expected negative sign, but the coefficient is not statistically significant.

These results are consistent with those obtained by Walberg and Fowler (1987) and Cooper and Cohn (1997) who found a positive relationship between the quality of instructional staff and a weak and negative relationship of the student-teacher ratio with achievement test scores. The coefficient of the parameter λ (= σ_u/σ_v) in the half normal specification indicates the presence of inefficiency in the production process. (See Deller and Rudnicki 1993; and Cooper and Cohn 1997.) A highly significant coefficient on λ implies $\sigma_u > \sigma_v$ and means that there is a high degree of inefficiency. The insignificant coefficient on λ means that, on average, these school districts are utilizing their resources efficiently.

The technical efficiency (e^{-u}) estimates based on half normal and exponential distributions of the one-sided component of the disturbance are compared and contrasted in Table 3. While there are differences in the measures of technical efficiency between these distributions, the rankings are very similar. The correlation coefficient for the two rankings is 0.976. The mean efficiency is 0.858 for the half normal estimates and 0.897 for the exponential function. The size of the district (i. e., number of students) also is shown in column 3 of Table 3. There is no obvious relationship between size and efficiency discernible in the results from the stochastic frontier model.

For the half normal distribution, the most and the least efficient school districts are Grand and North Sanpete whose technical efficiency scores are 0.991 and 0.625, respectively. Depending on the measure used, 18 to 27 of the school districts have efficiency scores of 0.90 or more. This should be interpreted as being good performance given the nature of the production system and the constraints on resource allocation decisions, especially with regard to personnel, many of whom have rather strong employment security.

Table 4 presents the efficiency estimates obtained from a simple DEA model (under VRS) using controllable and uncontrollable inputs and Tobit residuals from the two-stage DEA model. The simple DEA model addresses a somewhat different research question; i.e., given the factors both within and beyond management control, how efficient is the district? In this case it is appropriate that the exogenous factors that affect output be built into the measure of technical efficiency.[9] Most of the school districts are found to be more efficient under the simple DEA model than for the two-stage DEA model. Ordering these districts from the most efficient to the least efficient, we found that the rank orderings are quite different. Most of the school districts found to be efficient in the simple DEA model became inefficient in the two-stage DEA model. This implies that when the effects of uncontrollable factors were separated in the measure of pure technical efficiency, districts such as Juab, San Juan, and Weber became less efficient. However, controllable inputs did not have any effect on the performance of the least efficient school districts such as North Sanpete, South Summit, and Ogden.

Table 5 presents efficiency estimates from the stochastic frontier model (half normal) and Tobit residuals from the two-stage DEA model. Orderings of the districts from the most efficient to the least efficient for half normal are quite similar to those from Tobit residuals. Ranks for the most and the least efficient school districts remain the same in both models. The data in Table 5 indicate that districts that are technically less efficient in the stochastic estimation (e.g., Daggett, Kane, Rich, and Tintic) are more efficient when compared using Tobit residuals. The opposite is true for Washington and Salt Lake districts, which are less efficient based on the two-stage DEA model estimates.

The reason for minor differences in the orderings of efficiency scores between two models is due to the basic assumption about the random disturbance term. In the stochastic specification, a deviation of the production function from the frontier is the sum of a random component (v_i) and the inefficiency component (u_i). Nonstochastic DEA specification does not allow for such randomness where any deviation of the production function from the maximal is regarded as inefficient. School districts that appear highly efficient under stochastic specification contain a relatively larger random component of the error term (v_i) than the inefficiency component (u_i). Hence, in two-stage DEA, Tobit residuals for these school districts reflect less efficiency. However, districts that are found to be inefficient in the stochastic estimation and contain a relatively small random component of the error term (v_i) show better performance in the two-stage DEA. Examples are the Rich, Millard, Ogden, and Juab districts.

8. Summary

This study measures technical efficiency in each of the 40 school districts in Utah using both stochastic and nonstochastic estimation methods. In the stochastic estimation, substantial variation of technical efficiency among school districts is observed and it is invariant as to the distributional assumption of the one-sided component of the error term ε. The results of this study suggest that most of the school districts in Utah are technically efficient with mean efficiency scores 85.8 and 89.7 percent for the half normal and exponential distributions, respectively. The empirical results also indicate that the single most important factor explaining student performance is the level of parental education. The two-stage DEA model also indicates that socioeconomic and environmental factors have a strong influence on student success. There does not appear to be systematic variation among the groups of most efficient and least efficient school districts. In terms of size, ten districts at each end of the efficiency scale include both large and small districts and they are geographically dispersed. [10] There also is no apparent correlation between efficiency and the local economic base. Both efficient and inefficient districts are located in areas where agriculture, mineral extraction, or tourism is the predominant economic activity.

These results have several important policy implications. For example, districts with high socioeconomic status students might improve efficiency by better management of controllable inputs (i.e., teaching and other staff, student workload, etc.) and/or adoption of programs that link part of teacher compensation to student performance. Districts with a large number of low status students face a more difficult challenge as they deal with students who have less intellectual support at home. In such districts, efficiency could be enhanced by some resource reallocation to pre-kindergarten programs to better prepare young children for entering school, to adult education, and/or to greater teacher-parent interaction designed to encourage parental support of the student's educational activity.

The major limitation of this study is the use of aggregated data. Though there are other studies that used district level data (e.g., Levin et al. 1988; Fare et al. 1989; and Bessent et al. 1982), it is recognized that the same decisions regarding controllable inputs are often made at the school rather than the district level. Hence, aggregation of inputs and outputs at the district level may have caused some specification error that could have been transmitted to the estimation of the final efficiency score in both models. However, given these observations and the similarities of the results from parametric and nonparametric methods, it appears that researchers can safely select any of these methods without great concern for that choice having a large influence upon the empirical results.

References

Aigner, D. J., C. A. K. Lovell, and P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6:21-37.

Aigner, D. J., and S-F. Chu. 1968. On estimating industry production functions. *American Economic Review* 58:826-839.

Ali, M., and D. Byerlee. 1991. Economic efficiency of small farmers in a changing world: A survey of recent evidence. *Journal of International Development* 3:1-27.

Bagi, F. S., and C. J. Huang. 1983. Estimating production technical efficiency for industrial farms in Tennessee. *Canadian Journal of Agricultural Economics* 31:249-256.

Barrow, M. M. 1991. Measuring local education authority performance: A frontier approach. *Economics of Education Review* 10:19-27.

Battese, G. E., and G. S. Corra. 1977. Estimation of the production frontier model: With application to the pastoral zone of eastern Australia. *Australian Journal of Agricultural Economics* 21:169-179.

Battese, G. E., and T. J. Coelli. 1988. Prediction of farm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* 38:387-399.

Bessent, A., and W. Bessent. 1980. Determining the comparative efficiency of schools through data envelopment analysis. *Educational Administration Quarterly* 16:57-75.

Bessent, A., W. Bessent, J. Kennington, and B. Reagan. 1982. An application of mathematical programming to assess productivity in the Houston Independent School District. *Management Science* 28:1355-1367.

Charnes, A., W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2:429-444.

Coelli, T. J. 1995. Recent developments in frontier modelling and efficiency measurement. *Australian Journal of Agricultural Economics* 39:219-45.

Coelli, T., and S. Perelman. 1998. A comparison of parametric and non-parametric distance functions: with application to European railways. *European Journal of Operations Research* 117:326-339.

Cooper, S. T., and E. Cohn. 1997. Estimation of a frontier production function for South Carolina educational process. *Economics of Education Review* 16:313-327.

Debreu, G. 1951. The coefficient of resource utilization. *Econometrica* 19:273-292.

Deller, S. C., and E. Rudnicki. 1993. Production efficiency in elementary education: The case for Maine public schools. *Economics of Education Review* 12(1):45-57.

Fare, R., S. Grosskopf, and W. Weber. 1989. Measuring school district performance. *Public Finance Quarterly* 17:409-429.

Fare, R., S. Grosskopf, and C. A. K. Lovell. 1994. *Production frontiers.* Cambridge, MA: Cambridge University Press.

Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of Royal Statistical Society* A.120:253-281.

Forsund, E. R., C. A. K. Lovell, and P. Schmidt. 1980. A survey of frontier production functions and their relationship to efficiency measurement. *Journal of Econometrics* 13:5-25.

Frontier analysis parametric and nonparametric approaches. Annals 1990-4, A Supplement. 1990. *Journal of Econometrics*. Amsterdam, North Holland.

Greene, W. H. 1993. The econometric approach to efficiency analysis. In *The measurement of production efficiency: Techniques and applications*, edited by H. O. Fried, C.A. K. Lovell, and S. S. Schmidt. New York: Oxford University Press.

Grosskopf, S., K. J. Hayes, L. Taylor, and W. L. Weber. 1997. Budget constrained frontier measures of fiscal equality and efficiency in schooling. *Review of Economics and Statistics* 117-124.

Grosskopf, S., K. J. Hayes, L. L. Taylor, and W. L. Weber. 1991. Teacher salaries and teacher productivity. Paper presented at Western Economic Association Meetings.

Hanushek, E. A. 1971. Teacher characteristics and gains in student achievement: Estimation using micro data. *American Economic Review* 61(2):280-288.

Hanushek, E. A. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 24:1141-1177.

Hanushek, E. A., and L. L. Taylor. 1990. Alternative assessments of the performance of schools: Measurement of state variations in achievement. *The Journal of Human Resources* 25:179-201.

Horrace, W. C., and P. Schmidt. 1996. Confidence statements for efficiency estimates from stochastic frontier models. *The Journal of Productivity Analysis* 7:257-282.

Jondrow, J., C. A. K. Lovell, I. S. Materov, and P. Schmidt. 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19:233-238.

Kalirajan, K. P., and J. C. Flinn. 1983. The measurement of farm specific technical efficiency. *Pakistan Journal of Applied Economics* 2:167-180.

Klitgaard, R., and G. Hall. 1975. Are there unusually effective schools? *Journal of Human Resources* 10:90-106.

Koop, G., J. Osiewalski, and M. F. J. Steel. 1997. Bayesian efficiency analysis through individual effects: Hospital cost frontiers. *Journal of Econometrics* 76:77-105.

Koopmans, T. C. 1951. *Activity analysis of production and allocation.* New York: Wiley.

Kumbhakar, S. C., S. Ghosh, and J. T. McGuckin. 1991. A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business and Economic Statistics* 9:279-286.

Lee, L. F., and W. G. Tyler. 1978. A stochastic frontier production function and average efficiency: An empirical analysis. *Journal of Econometrics* 7:385-390.

Levin, H. 1976. Concepts of economic efficiency and education production. In *Education as an industry*, edited by J. T. Froomkun, D.T. Jamison, and R. Renders. Cambridge, MA: Ballinger Publishing Co. for National Bureau of Economic Research.

Levin, H. 1974. Measuring efficiency in educational production. *Public Finance Quarterly* 2:3-24.

Lovell, C. A. K. 1993. Production frontiers and productive efficiency. In *The measurement of productive efficiency: Techniques and applications*, Chapter 1, edited by H. O. Fried, C.A. K. Lovell, and S. S. Schmidt. New York: Oxford University Press.

McCarty, T. A., and S. Yaisawarng. 1993. Technical efficiency in New Jersey school. In *The measurement of production efficiency: Techniques and applications*, Chapter 10, edited by H. O. Fried, C.A. K. Lovell and S. S. Schmidt. New York: Oxford University Press.

Meeusen, W., and J. van den Broeck. 1977. Efficiency estimation from Cobb-Douglas production frontiers with composed error. *International Economic Review* 18:435-444.

Pitt, M. M., and L. F. Lee. 1981. Measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics* 9:43-64.

Ray, S. C. 1991. Resource-use efficiency in public schools: A study of Connecticut data. *Management Science* 37:1621-1628.

Schmidt, P. S., and R. C. Sickles. 1984. Production frontiers and panel data. *Journal of Business and Economic Statistics* 2:367-374.

Schmidt, P. S., and C. A. K. Lovell. 1979. Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. *Journal of Econometrics* 9:343-366.

Seiford, L. M., and R. M. Thrall. 1990. Recent development in DEA: The mathematical programming approach to frontier analysis. *Journal of Econometrics* 46(1/2):7-38.

Utah Education Association. 1993. *Teacher salary schedules*. Research Bulletin Volume 38, No. 2, Salt Lake City, Utah, March.

Utah Office of Education. 1993. Summary of statistical and financial data of the Utah State Superintendent of Public Instruction, 1992-93. Salt Lake City, Utah, September.

U. S. Department of Commerce. 1999. *Statistical abstract of the United States, 1999*. Washington, DC: Bureau of Census, Economic and Statistics Administration.

van den Broeck, J., G. Koop, J. Osiewalski, and M. F. J. Steel. 1994. Stochastic frontier models: A Bayesian perspective. *Journal of Econometrics* 61:273-303.

Walberg, H., and W. Fowler. 1987. Expenditure and size efficiencies of public school districts. *Educational Researcher* 16:5-13.

Wyckoff, J. H., and J. Lavinge. 1991. The relative inefficiency of public elementary schools in New York. Working Paper, State University of New York, Albany.

Waldman, D. M. 1984. Properties of technical efficiency estimators in the stochastic frontier model. *Journal of Econometrics* 25:353-364.

Table 1. Summary Statistics for Utah School Districts, 1992-93

Variable	Mean	Standard Deviation	Minimum	Maximum
Average 11th grade test score	52.10	7.52	30.00	68.00
Student-teacher ratio	20.17	3.29	10.59	30.03
Percentage of teachers with advanced degree	26.04	10.02	2.78	43.59
Percentage of teachers with over 15 years experience	17.36	4.02	5.88	25.81
Percentage of population with high school diploma	82.82	6.14	59.70	91.60
Percentage of students receiving subsidized lunch	25.65	10.62	5.00	51.00
Net assessed value per student	$191,290	$162,970	$56,700	$702,800

Table 2. Stochastic Frontier Parameter Estimates
Dependent Variable: Ln(Test Score)

Variable	MLE (Half-Normal)	MLE (Exponential)
Constant	0.877	0.766
	(0.402)	(0.463)
Ln(Student-teacher ratio)	-0.289	-0.196
	(-1.546)	(-1.380)
Ln(Percentage of teachers with advanced degree)	0.024 (0.382)	0.057 (1.195)
Ln(Percentage of teachers with experience over 15 years)	0.032 (0.234)	-0.016 (-0.206)
Ln(Percentage of students receiving subsidized lunch)	-0.039 (-0.430)	-0.041 (-0.739)
Ln(Percentage of population with a high school diploma)	0.959* (2.215)	0.909* (3.029)
Ln(Net assessed value per student)	-0.015	-0.011
	(-0.329)	(-0.308)
λ	12.705	
	(0.468)	
θ		8.832*
		(3.247)
Log of the Likelihood Function	32.969	32.622

t-statistics are in parentheses.
*Indicates coefficient is significant at the 5% or lower probability level.

Table 3. Measuring Technical Efficiency Using Half-Normal and Exponential Distributions

School District	District Size	Half-Normal Efficiency	Rank	Exponential Efficiency	Rank
Alpine	40,322	0.963	8	0.967	6
Beaver	1,396	0.850	25	0.908	27
Box Elder	11,190	0.933	9	0.945	15
Cache	12,593	0.889	19	0.931	20
Carbon	5,150	0.880	21	0.935	17
Daggett	191	0.914	14	0.955	10
Davis	57,116	0.881	20	0.925	23
Duchesne	4,411	0.862	23	0.911	25
Emery	3,400	0.816	28	0.890	28
Garfield	1,097	0.746	35	0.797	34
Grand*	1,576	0.991	1	0.981	1
Granite	79,575	0.901	18	0.930	21
Iron	5,475	0.917	13	0.948	13
Jordan	68,843	0.909	16	0.942	16
Juab	1,644	0.770	33	0.829	33
Kane	1,415	0.904	17	0.960	9
Millard	3,861	0.724	36	0.776	36
Morgan	1,889	0.853	24	0.909	26
Nebo	17,161	0.829	26	0.876	29
No.Sanpete*	2,352	0.625	40	0.672	40
No.Summit	944	0.711	37	0.762	37
Park City	2,540	0.971	5	0.966	8
Piute	385	0.973	4	0.974	3
Rich	549	0.797	29	0.915	24
San Juan	3,400	0.640	39	0.686	39
Sevier	4,859	0.793	31	0.834	32
So.Sanpete	2,899	0.878	22	0.934	18
So.Summit	1,106	0.668	38	0.720	38
Tintic	241	0.774	32	0.855	30
Tooele	7,355	0.924	11	0.948	14
Uintah	6,795	0.970	7	0.974	4
Wasatch	3,137	0.979	3	0.975	2
Washington	14,596	0.982	2	0.966	7
Wayne	580	0.795	30	0.927	22
Weber	26,832	0.818	27	0.849	31
Salt Lake	25,538	0.921	12	0.948	12
Ogden	12,589	0.758	34	0.787	35
Provo	13,565	0.971	6	0.971	5
Logan	5,894	0.931	10	0.952	11
Murray	6,799	0.909	15	0.933	19
Mean	11,531	0.858		0.897	

*Indicates the most and the least efficient school districts.

Table 4. Technical Efficiency Estimates from Simple DEA Model and Tobit Residuals from Two-Stage DEA Models

School District	Simple DEA Model (VRS)	Rank	Tobit Model Residuals	Rank
Alpine	1.000	1	0.071	10
Beaver	0.863	10	-0.013	24
Box Elder	1.000	1	0.040	14
Cache	1.000	1	0.003	21
Carbon	1.000	1	-0.022	25
Daggett	1.000	1	0.216	1
Davis	1.000	1	0.016	17
Duchesne	0.963	4	-0.044	29
Emery	0.827	11	-0.034	28
Garfield	0.774	14	-0.105	33
Grand	1.000	1	0.179	2
Granite	0.910	6	-0.031	27
Iron	0.897	8	-0.006	23
Jordan	1.000	1	0.016	18
Juab*	1.000	1	-0.074	31
Kane	0.950	3	0.087	9
Millard	0.760	15	-0.113	35
Morgan	1.000	1	0.023	16
Nebo	0.885	9	-0.058	30
No. Sanpete**	0.673	18	-0.240	40
No. Summit	0.826	12	-0.132	36
Park City	1.000	1	0.105	6
Piute	1.000	1	0.178	3
Rich	1.000	1	0.097	7
San Juan*	1.000	1	-0.183	38
Sevier	0.819	13	-0.109	34
So. Sanpete	1.000	1	0.003	20
So. Summit**	0.756	17	-0.186	39
Tintic	1.000	1	0.050	13
Tooele	1.000	1	0.025	15
Uintah	1.000	1	0.064	11
Wasatch	1.000	1	0.107	5
Washington	0.918	5	-0.029	26
Wayne	1.000	1	0.163	4
Weber*	1.000	1	-0.082	32
Salt Lake	0.899	7	0.005	22
Ogden**	0.757	16	-0.145	37
Provo	1.000	1	0.087	8
Logan	1.000	1	0.054	12
Murray	0.952	2	0.006	19

*Indicates the most efficient district in the simple DEA model but least efficient in the two-stage DEA model.

**Indicates least efficient for both models.

Table 5. Technical Efficiency Estimates from Frontier Model (Half-Normal) and Tobit Residuals from Two-Stage DEA Models

School District	Half-Normal Efficiency	Rank	Tobit Model Residuals	Rank
Alpine	0.963	8	0.071	10
Beaver	0.850	25	-0.013	24
Box Elder	0.933	9	0.040	14
Cache	0.889	19	0.003	21
Carbon	0.880	21	-0.022	25
Daggett**	0.914	14	0.216	1
Davis	0.881	20	0.016	17
Duchesne	0.862	23	-0.044	29
Emery	0.816	28	-0.034	28
Garfield	0.746	35	-0.105	33
Grand*	0.991	1	0.179	2
Granite	0.901	18	-0.031	27
Iron	0.917	13	-0.006	23
Jordan	0.909	16	0.016	18
Juab	0.770	33	-0.074	31
Kane	0.904	17	0.087	9
Millard	0.724	36	-0.113	35
Morgan	0.853	24	0.023	16
Nebo	0.829	26	-0.058	30
No.Sanpete*	0.625	40	-0.240	40
No.Summit	0.711	37	-0.132	36
Park City	0.971	5	0.105	6
Piute	0.973	4	0.178	3
Rich	0.797	29	0.097	7
San Juan	0.640	39	-0.183	38
Sevier	0.793	31	-0.109	34
So.Sanpete	0.878	22	0.003	20
So.Summit*	0.668	38	-0.186	39
Tintic	0.774	32	0.050	13
Tooele	0.924	11	0.025	15
Uintah	0.970	7	0.064	11
Wasatch	0.979	3	0.107	5
Washington**	0.982	2	-0.029	26
Wayne	0.795	30	0.163	4
Weber	0.818	27	-0.082	32
Salt Lake	0.921	12	0.005	22
Ogden	0.758	34	-0.145	37
Provo	0.971	6	0.087	8
Logan	0.931	10	0.054	12
Murray	0.909	15	0.006	19

*Indicates the most and the least efficient districts.
**Indicates districts with a significant effect of uncontrollable factors.

Endnotes

[1]See U. S. Department of Commerce (1999), Tables 253, 254, and 296.

[2]See Ali and Byerlee (1991), Lovell (1993), Green (1993), and Coelli (1995) for a detailed discussion on the methods for analyzing technical efficiency.

[3]The *Journal of Econometrics* (1990) devoted an entire supplemental issue to parametric and nonparametric approaches to frontier analysis.

[4]A more sophisticated and satisfying approach uses the Bayesian paradigm for making inferences about firm-specific inefficiencies using both cross section and panel data (Koop et al. 1997; van den Broeck et al. 1994; and Horrace and Schmidt 1996).

[5]Hanushek and Taylor (1990) and Grosskopf et al. (1997) used a value added residual technique to measure educational output.

[6]Nonavailability of data on each component of the battery test precludes estimating a multioutput production function. Data on input prices for goods, such as education, often are not available; hence, the use of a production function instead of cost function is more convenient for measuring efficiency.

[7]We recognize that the Cobb-Douglas production function uses restrictive assumptions on the elasticity of substitution and scale properties. However, due to insufficient data, a more flexible form such as translog production function was not tested because of a limited number of degrees of freedom. Coelli and Perelman (1998) point out that if the production units do not behave as perfectly competitive firms in an industry, the use of a Cobb-Douglas function may be acceptable.

[8]Parameters of stochastic frontier production function and technical efficiency are estimated using LIMDEP, and the DEA model was estimated using DEAP (2.1) software developed by T. Coellie.

[9]Kumbhakar, Ghosh, and McGuckin (1991) incorporated factors that affect output in a stochastic production frontier model that specifies technical efficiency as a function of noncontrollable inputs.

[10]In order to check for systematic effect of in-migration and out-migration, if any, on the measure of efficiency scores in the stochastic frontier model, a dummy variable was added to identify districts that are adjacent to metropolitan areas. The coefficient on that variable was insignificant.