THE DETERMINANTS OF SCHOOL EFFICIENCY IN OKLAHOMA: RESULTS FROM STOCHASTIC PRODUCTION FRONTIER AND DATA ENVELOPMENT ANALYSIS

By

SUSANNE RASSOULI-CURRIER

Bachelor of Science
National Iranian Oil Company College of
Accounting and Finance
Tehran, Iran
1980

Master of Business Administration
University of Central Oklahoma
Edmond, Oklahoma
1994

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2002

THE DETERMINANTS OF SCHOOL EFFICIENCY IN OKLAHOMA: RESULTS FROM STOCHASTIC PRODUCTION FRONTIER AND DATA ENVELOPMENT ANALYSIS

Thesis Approved:

Thesis Advisor

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to express my appreciation to my major advisor, Dr. Lee Adkins, for the guidance and support that made the completion of this dissertation possible. I also thank the members of my committee, Dr. Mary Gade, Dr. Janice Jadlow, and Dr. Ronald Moomaw for their suggestions and support.

Special thanks go to my good friend and colleague, Steven Miller, for his assistance with technical support and various computing aspects of this research; and more importantly, for his continuous encouragement and moral support. Also, I thank Cindy Conway for her diligence and patience with my work.

Special thanks go to my husband, Dr. Kevin Currier, and my sons Keia and Kaivon for their encouragement and many sacrifices.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The operational funding for public schools in the United States comes directly from tax revenue, and consequently taxpayers expect public schools to attain 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. Serious questions have been raised about the management and efficiency of public schools in the U.S.

The demand for accountability has spawned literally hundreds of studies that attempt to determine factors upon which school performance depends. Most studies focus on the "money matters" question: "Do increases in per-pupil spending improve student performance?" Hanushek (1994) found that the performance of U.S. students ranks below that of many other countries. 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 had large increases in resources with very little, if any, improvement in outcomes. These findings confirmed his earlier statement in (1986, p. 1162) that "there is no strong or systematic relationship between school expenditures and student performance."

In *Assessing Education Practices,* Becker and Baumol (1995) state that real expenditures on secondary and elementary education more than tripled between 1960 and 1990, resulting in a lower student-teacher ratio and a rise in the average age, education, and specialization level of teachers. However, performance continues to decline, leading them to refer to the educational expenditures as *deceptive indicators*.

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 would have had increased by at least 2.9 percent and thereby led to an increase of 86 billion dollars in GDP.

Efficiency studies suggest a different question about the link between school funding and performance: "Can schools reallocate existing expenditures in ways that improve performance?" 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 (e.g., the Becker and Baumol (1995) criticism that the poor quality of learning accomplishment signals the lack of rigorous curriculum and lack of sufficient rewards for learning.)

Deller and Rudnicki (1993), observed a positive relationship between instructional expenditure per student and, students' test scores while non-instructional (e.g. administration, operation and busing) spending has a negative effect on test scores.

In any case, the authors identify the existence of production inefficiencies in the educational production process. They suggest that, if schools are not utilizing their resources optimally, then additional resources may not produce improved outcomes.

Like other states in the U.S., Oklahoma seeks improvements in students' performance. A major educational reform law, House Bill 1017 (HB1017), was passed in early 1990 in an attempt to improve the quality of education in the state. Because the funding came from a tax increase as well as reallocation of state funds toward common education and away from other popular programs (Moomaw and Yusof, 1995, p. 1), the law created much controversy concerning its effectiveness.

Abdul Rahman (1996) estimated a simple model of the determinants of school inefficiency. Her study included various causes of inefficiencies (e.g., inputs which school administrators have control over such as expenditures and factors that are beyond their control, such as socioeconomic variables). Based on her analysis, she concluded that better quality teachers and smaller class size are relevant to better student performance. This suggests that the measures in HB1017 were a move in the right direction. However, further examination suggested that socioeconomic factors play an important role in the districts' inefficiencies. Schools with low socioeconomic status (regardless of how it is measured) are generally less efficient.

Jacques and Brorsen (1997) estimate the effects of school spending (several categories) on school performance as measured by test scores. They conclude that higher levels of instructional expenditures per student are in fact associated with higher test scores. There is no significant evidence that higher levels of instructional support lead to improved performance and in fact find a negative relationship between test scores and spending on student support (which included expenditures on attendance and social work services, guidance services, health services, individual psychological services, speech pathology, and audiological services) holding other factors considered constant.

In a report to the Oklahoma State Senate staff, Michael Metzger (1999) concluded that a 10 percent increase in total expenditures would result in a little over 1 percent increase in student performance and suggested that a 15 percent increase in expenditures would be required to raise Oklahoma's average ACT test score from 20.6 in 1997 to the national average of 21. Abdul Rahman (1996) also found the elasticity of test scores with respect to instructional spending to be very small for the years 1991-1995. Adkins and Moomaw (1997) obtained similar results to those of Abdul Rahman using a maximum likelihood estimator.

This dissertation uses data provided by the Oklahoma Office of Accountability to estimate the stochastic production frontier for Oklahoma school districts and the determinants of district inefficiency. One new element in this research includes the use of DEA to explore the determinants of inefficiency in Oklahoma school districts. More explicitly, the objectives are:

1. examine the relationship between school district inputs and educational outcomes for Oklahoma.
2. specify and estimate a stochastic production frontier.
3. determine causes of inefficiency based on the most recently available data.
4. compare these results to those estimates from Data Envelopment Analysis (DEA).
5. use an estimator that permits the random errors of the stochastic frontier to be heteroscedastic.

1.1 Significance of the Study

The study provides a thorough and up-to-date account of schools production in Oklahoma using the latest available data and means of analysis.

1.2 Organization of the Study

This study contains six more chapters organized as follows. Chapter II is a literature review, which is in two sections. Section 1 is concerned with the existing literature on education in general and Oklahoma's education system in particular, and section 2 focuses on the production models. The data description and sources for the stochastic production frontier model are presented in Chapter III. Chapter IV develops the model and discusses the econometric issues surrounding it. Chapter V presents the estimates, results, and discussions regarding the study's ability to improve the analytical tools for Oklahoma school performance as well as other areas of interest. Chapter VI includes the estimates, results and discussions based on a different specification than in Chapter V and using Data Envelopment Analysis technique. Chapter VII contains the summary and conclusion.

CHAPTER II

REVIEW OF THE LITERATURE

2.1 Educational Production

Many economic studies of educational production, efficiency, and cost structure have been inspired by the Coleman report (Coleman, T. et al., 1966). The Coleman report was influential in that the analysis covered approximately 3,000 elementary and secondary schools with approximately half a million students. The report suggested an input-output relationship between administrative resource allocation and students' achievements. In addition, the report introduced policymakers to analytical issues such as production efficiency and the existence of multicollinearity among variables (Hanushek, 1979). In the report, the researchers found that students' performance was related largely to their socioeconomic background rather than the variation in schools (Hanushek, 1986, 1989).

Policy issues implied by the Coleman report generated significant interest in analysis of school performance. These studies differ in their focus and methodology; however, they provide some understanding of school efficiency.

Hanushek's survey of 147 studies (1986) suggests that in most studies expenditures per pupil, student/teacher ratio, teacher education and experience as well as family characteristics are used as the primary determinants of student achievement. The results of these studies are in many ways contradictory; however, they are consistent in that expenditure and student performance lack a strong or systematic relationship and family characteristics definitely have an effect on their achievement.

Deller and Rudnicki (1993) suggested that competitive pressure within highly concentrated counties results in better school performance and is responsible for the *school choice* argument, i.e., allowing parents to choose which public school their children attend.

Caroline Hoxby conducted two separate studies on the effect of school choice on school performance. The first (Hoxby, 1994a) examines the choice between private and public schools. More specifically, she investigates the effect of private school enrollment on public school performance, holding public school spending constant. Hoxby concludes that increased competition between private and public schools increased public schools' productivity without any increase in spending.

The second study Hoxby (1994b) examines the extent to which greater choice among public schools affects public schools' performance. The author suggests that public schools with the lower per pupil spending, lower teacher salaries, and larger class size in the areas that have choices among public schools tend to have better than average student performance.

The possible effect of school size and/or district size on student performance cited by researchers captures the effect of economies of scale on schools' productivity. The results from studies of the effects of the scale economies associated with public schools are inconsistent; some find evidence of economies of scale and others do not. These inconsistencies affect the confidence in the widely promoted and practiced consolidation policies based on economies of scale of the school districts.

Abdul Rahman (1996) studied the determinants of Oklahoma school efficiency. According to her study, the progress reports on educational performance in Oklahoma since HB 1017 suggest that schools' performance, measured by several standardized test scores, improved. The study's main purpose was to evaluate the potential effect of increases in spending in school districts and the effect of socioeconomic as well as other external factors on school district efficiency. The study finds:

- evidence of inefficiency in Oklahoma schools.
- that inefficiency, to a certain degree, is the consequence of the district's socioeconomic status. School districts with less favorable socioeconomic environments are generally less efficient.
- school districts with smaller class sizes perform better.
- evidence that economies of scale exists; i.e., size efficiency is beneficial to school districts performance. The evidence is stronger for upper grades.

Adkins and Moomaw (1997) studied the determinants of technical efficiency in Oklahoma schools. They find that:

- money matters, but not much; estimated elasticity of test score to spending are positive but very small.
- districts that have more experienced teachers are more efficient in all grades considered, except for grade 3. This may suggest that more experienced teachers may be more effective in higher grades but youthful ones may be more effective in lower grades.

- districts that pay higher salaries get better results.
- there also is evidence of possible efficiency gains from the size of the school district. Larger districts in Oklahoma tend to be more efficient than small districts. Thus districts might benefit by consolidation. According to Adkins and Moomaw's estimation, the optimal size for technical efficiency is between 18,000 and 22,000 students.

In a report to the Senate staff, Metzger (March 1999) suggests that the likely cause of the contradictory conclusions regarding expenditures and district structure may be because of errors in data and choice of model and model specification errors (i.e., econometric issues). Hanushek (1979, 1986) suggests that future research in this area should bear in mind the following. First, measuring and defining educational inputs and outputs can be problematical. Second, data availability may necessitate compromise regarding model selection. Finally, consideration must be given to the definition of efficiency and how it is being measured.

Some consistent findings have emerged:

- If money matters, it doesn't matter much.
- Schools tend to be inefficient and hence reallocation of resources within a district could improve performance.
- Socioeconomic factors are important.
- Few are willing to make sweeping changes in policy based on their results.

2.2 Production in Economics -- Production Function

A production function is a mathematical expression that relates inputs to outputs, given technology. The production function indicates the maximum output attainable with a given vector of inputs (Henderson and Quandt, p. 66).

Assume the following production function with two inputs and one output:

$$y = f(x_1, x_2) \tag{2.1}$$

where y is output; x_1 and x_2 are inputs; and $f(.)$ is a twice continuously differentiable function. In equation (2.1), y is the maximum quantity of output that can be obtained with different quantities of the inputs x_1, x_2.

A typical production function is represented below in Figure 2.1.

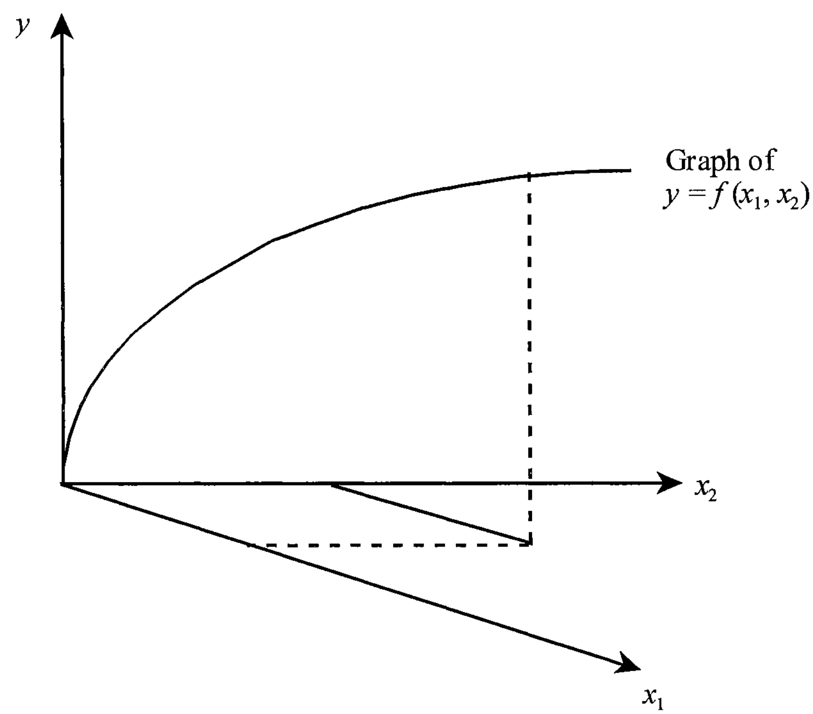

Figure 2.1
Production Function

To determine the output supply and factor demand equations for firms with optimizing behavior, i.e., profit maximizing or cost minimizing firms, two different, but equivalent approaches can be taken. These are the primal approach and the dual approach.

2.2.1 The Primal Approach

When the output supply and factor demands are derived from a *direct* objective function, it is referred to as the primal approach.

Profit Maximization

The profit of the firm is defined as total revenue (TR) minus total cost (TC). Assume a profit-maximizing firm with production function given by equation (2.1) and profit π as:

$$\pi = \mathrm{TR} - \mathrm{TC} \tag{2.2}$$

or

$$\pi = p.f(x_1, x_2) - (w_1 x_1 + w_1 x_2) \tag{2.3}$$

where p, w_1, w_2 are the prices of output, x_1 and x_2, respectively. The values of x_1, x_2 that maximize profit can be obtained by setting the first order partial derivatives of equation (2.3) with respect to x_1 and x_2 equal to zero and solving for x_1 and x_2 simultaneously. That is:

$$x_1^* = x_1^* (p, w_1, w_2) \tag{2.4}$$

$$x_2^* = x_2^* (p, w_1, w_2) \tag{2.5}$$

By substitution of (2.4) and (2.5) into the production function we obtain the output supply function:

$$y^* = \mathrm{y}^* (p, w_1, w_2) \tag{2.6}$$

Cost Minimization

For a cost-minimizing firm the goal is to produce a certain level of output at minimum cost, given input prices.

Consider the following direct cost function associated with production in equation (2.1):

$$c = w_1 x_1 + w_2 x_2 \tag{2.7}$$

The input level, that minimizes cost, is obtained by minimizing (2.7) subject to production technology described in (2.1).

This is a restricted minimization problem, which can be solved by setting up a Lagrangian function:

$$l = w_1 x_1 + w_2 x_2 + \lambda\,(y - f(x_1, x_2)) \tag{2.8}$$

setting the first partial derivatives with respect to x_1, x_2, and λ equal to zero and solving for x_1 and x_2 simultaneously, then

$$x_1^* = x_1\,(y,\ w_1,\ w_2) \tag{2.9}$$

$$x_2^* = x_2\,(y,\ w_1,\ w_2) \tag{2.10}$$

x_1^* and x_2^* are conditional input demands.

In empirical applications, the primal approach requires the knowledge of the production function $f(.)$. The parameters of the production function have to be estimated (e.g., using econometric methods) and only then the output supply and factor demand equations can be derived.

2.2.2 The Dual Approach

The dual approach is the alternative to the primal approach. Its primary advantage is that it avoids the extensive computation involved in the primal approach. In this approach, the output supply and factor demands are derived from an *indirect* objective function.

Profit Maximization

Assume the profit-maximizing firm in equation (2.1). The *maximum* profit can be obtained by substituting equations (2.4), (2.5), and (2.6) into equation (2.3):

$$\pi^* = p[y(p, w_1, w_2)] - \{w_1[x_1(p, w_1, w_2)] + w_2[x_2(p, w_1, w_2)]\} \tag{2.11}$$

$$= \pi^*(p, w_1, w_2). \tag{2.12}$$

In the primal approach, profit is solely a function of input and output prices. In practice, the profit function is specified with appropriate properties (e.g., monotonicity, homogeneity, symmetry, etc.) and is estimated from observations from a sample data.

The profit function in equation (2.12) is the *indirect* profit function. According to Hotelling's Lemma, the first partial derivative of this profit function with respect to input prices is the negative of the input demands then:

$$\frac{\partial \pi^*}{\partial w_1} = -x_1^*(p, w_1, w_2)$$

and

$$\frac{\partial \pi^*}{\partial w_2} = -x_2^*(p, w_1, w_2)$$

And the first partial derivative of the profit function with respect to output price is the output supply equation:

$$\frac{\partial \pi^*}{\partial p} = y(p, w_1, w_2)$$

Young's Theorem from calculus imposes certain *symmetry* restrictions between the cross partial derivatives of input demand and output supply functions. Young's Theorem states that the order of differentiation does not affect a second partial derivative for any twice continuously differentiable function:

$$-\frac{\partial x_1^*(p, w_1, w_2)}{\partial w_2} = \frac{\partial^2 \pi^*(p, w_1, w_2)}{\partial w_1 \partial w_2} = \frac{\partial^2 \pi^*(p, w_1, w_2)}{\partial w_2 \partial w_1} = -\frac{\partial x_2^*(p, w_1, w_2)}{\partial w_1}$$

and also:

$$-\frac{\partial x_1^*(p, w_1, w_2)}{\partial p} = \frac{\partial^2 \pi^*(p, w_1, w_2)}{\partial w_1 \partial p} = \frac{\partial^2 \pi^*(p, w_1, w_2)}{\partial p \partial w_1} = \frac{\partial y^*(p, w_1, w_2)}{\partial w_1}$$

These symmetry conditions must be imposed on the profit function (2.12) when using the dual approach.

Cost Minimization

Assuming that the cost-minimizing firm whose behavior is defined by (2.7) and (2.8), the *indirect* cost function of the cost minimizing firm can be obtained by substituting the cost minimizing input demand equations (2.9) and (2.10) into equation (2.7):

$$\begin{aligned} c^* &= w_1\left[x_1^*(y, w_1, w_2)\right] + w_2\left[x_2^*(y, w_1, w_2)\right] \\ &= c^*(y, w_1, w_2) \end{aligned} \quad (2.13)$$

Equation (2.13) is the *indirect* cost function. In practice, this function is also specified with appropriate properties (e.g., monotonicity, homogeniety, etc.). The function can then estimated and Shephard's Lemma is used to derive the input demand equations.

Shephard's Lemma states that the input demand equations can be derived by the first partial derivative of the cost function with respect to input prices. These input demand functions are conditional upon the output level y. Then:

$$\frac{\partial c^*}{\partial w_1} = x_1^*(y, w_1, w_2)$$

and

$$\frac{\partial c^*}{\partial w_2} = x_2^*(y, w_1, w_2)$$

Young's Theorm implies the symmetry condition:

$$\frac{\partial x_1^*(y, w_1, w_2)}{\partial w_2} = \frac{\partial^2 c^*(y, w_1, w_2)}{\partial w_1 \partial w_2} = \frac{\partial^2 c^*(y, w_1, w_2)}{\partial w_2 \partial w_1} = \frac{\partial x_2^*(y, w_1, w_2)}{\partial w_1}$$

2.2.3 Advantages of the Dual Approach

The major advantage of the dual approach is that it does not require specific functional knowledge of the production function in order to derive the output supply and input demand equations. These equations can be derived directly from the cost function or the profit function. The dual approach, with the *appropriate* cost or profit function, avoids the computational difficulties of the primal approach. Other advantages of the dual approach will be discussed in Section 2.2.5.

2.2.4 Concepts of Efficiency in Production

Efficiency in production can be defined in several ways including technical efficiency, allocative efficiency, and total efficiency.

Technical Efficiency

Without loss of generality and to facilitate graphical description, assume a production process with a single input (x) and a single output (y). The production function which explains the relationship between input and output is $y = f(x)$. This function is represented in Figure 2.2.

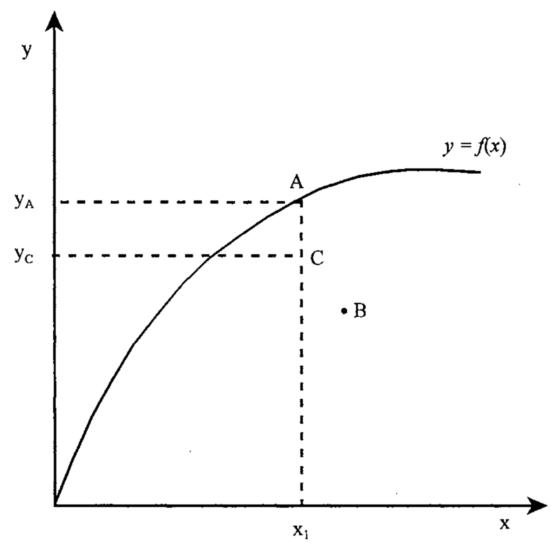

Figure 2.2
Production Function and Technical Efficiency

In the context of efficiency measurement, the literature tends to refer to the production function as the production frontier to stress the maximal property of the function (Coelli, 1998, p. 12). Thus, the production frontier represents the maximum output attainable from each level of the input, with the current state of technology in the industry.

If firms in the industry operate *on* the frontier (e.g., point A), then they are technically efficient. If they operate *below* the frontier (e.g., points B and C), then they are technically inefficient.

A firm at point C is technically inefficient because it can operate at point A which produces a higher output level ($y_A > y_C$) with the same amount of input (x_1). The same type of argument explains the inefficiency resulting from operation at point B.

Allocative Efficiency

Allocative efficiency in input selection is the mix of inputs that produces a given quantity of output at minimum cost. When input prices and output are known and when certain behavioral assumption such as profit maximization or cost minimization is appropriate; allocative efficiency can be measured.

Total efficiency: Total efficiency is the sum of technical efficiency and allocative efficiency. If a firm is technically and allocatively efficient, then the firm is said to be economically efficient (totally efficient).

2.2.5 Production Function vs. Cost Function

Depending on the objectives of estimation, a firm's production function or cost function can be considered.

Production Function

The *production function* should be estimated if:

1. the only known objective of the firm is to operate on its frontier as opposed to operating below it, i.e., obtaining maximum output from any given combinations of inputs. Estimation of the production function does not

require a behavioral assumption such as profit maximization or cost minimization.

2. there is no information available on input or output prices. The production function is purely a technical relationship.

The downsides of using the production function in estimation are:

1. if firms are profit maximizers or cost minimizers, estimation can suffer from simultaneous equation bias. This happens because the input levels and error terms are not independent of each other (Coelli, 1995, p. 226).
2. the production function captures *only* the *technical* inefficiency. This is a major drawback if the analyst is concerned with allocative inefficiency as well as technical inefficiency.
3. modeling multiple output production can be difficult.

Cost Function

The *cost function* should be estimated:

1. if the firm desires to produce a certain output level at least cost and, if input price information is available. In this case no knowledge of the production function is required.
2. because the *only* algebraic manipulation to obtain the factor demands is the partial differentiation of the indirect profit function This is another advantage of the dual approach (Coelli, 1998).
3. if the firm produces multiple outputs.
4. if the researcher is concerned with the firm's allocative as well as technical efficiency, i.e., total efficiency.

5. because it is easier to obtain information on cost and input prices than obtaining information on input quantities.
6. the symmetry property of the cost function is of use in reducing the number of parameters being estimated, i.e., conserving the degrees of freedom and possibly elimination of multicolinearity problems (Coelli, 1998).

The downsides of using the cost function in estimation are:

1. input price information is required.
2. the hypothesis of cost minimization or profit maximization is required. These *maintained* hypotheses could be false in reality if the firm chooses to pursue other goals.
3. the total efficiency can be decomposed into its technical and allocative components *only* if the production function implied by the estimated cost function can be explicitly derived. This class of functions is referred to as self-dual functions, e.g., Cobb-Douglas technology (Coelli, 1998).

Multiple Output Production and Distance Functions

As discussed earlier, the direct estimation of a production function does not allow for multiple output production technologies. In the past, researchers faced with this situation estimated the production function using a single aggregate output measure (Coelli, 1998).

In recent years some researchers dealt with this problem by using *distance functions*. Distance functions allow one to describe a multi-input, multi-output production process without specifying a behavioral assumption such as profit maximization or cost minimization.

An input distance function is concerned with a minimal proportional decrease of input vector, given an output vector. Alternatively, an output distance function is concerned with a maximal proportional increase of the output vector, given an input vector.

The focus of this section is to discuss the notion of output distance function. Following Coelli (1998), assume a single input x_1 which produces two outputs y_1, y_2. The input requirement function can be defined as:

$$x_1 = g(y_1, y_2) \tag{2.14}$$

The function defined in equation (2.14) can be illustrated by a production possibility curve (PPC). PPC represents the different combinations of output that can be produced with a given level of input. PPC is the output counterpart of an isoquant and its properties are similar to the properties of isoquant. The production possibility curve in equation (2.14) is represented in Figure 2.3.

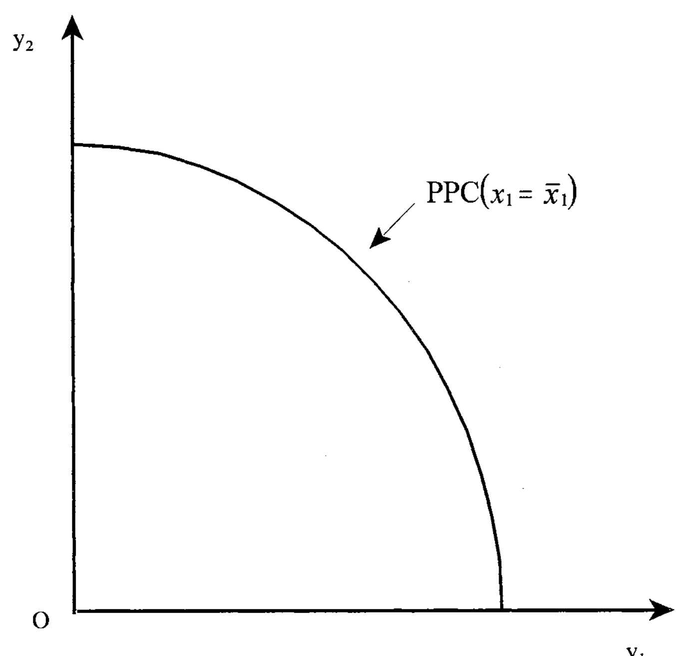

Figure 2.3
Production Possibility Curve

Now assume that the production technology defined by output sets, P(x), represents the set of all output vectors, y, that can be produced using input vector x then:

$$P(x) = \{y : x \text{ can produce } y\}$$

Coelli (1998, p. 62) summarizes the properties of this set. The output distance function on the output set ρ(x) can be defined as:

$$d_O(x, y) = \min\{\rho : (y/\rho) \in P(x)\}$$

Following the axioms on the technology set, properties of $d_O(x, y)$ are:

1) $d_O(x, y)$ is increasing in x and non-decreasing in y;

2) $d_O(x, y)$ is linearly homogeneous in y;

3) $d_O(x, y) \leq 1$ if y belongs to the production possibility set of x, i.e., $y \in P(x)$; and

4) $d_O(x, y) = 1$ if y is on the PPC curve of x.

The concept of output distance function can be illustrated by the following example. Figure 2.4 represents the production technology where the outputs y_1, y_2 are produced using the input vector x.

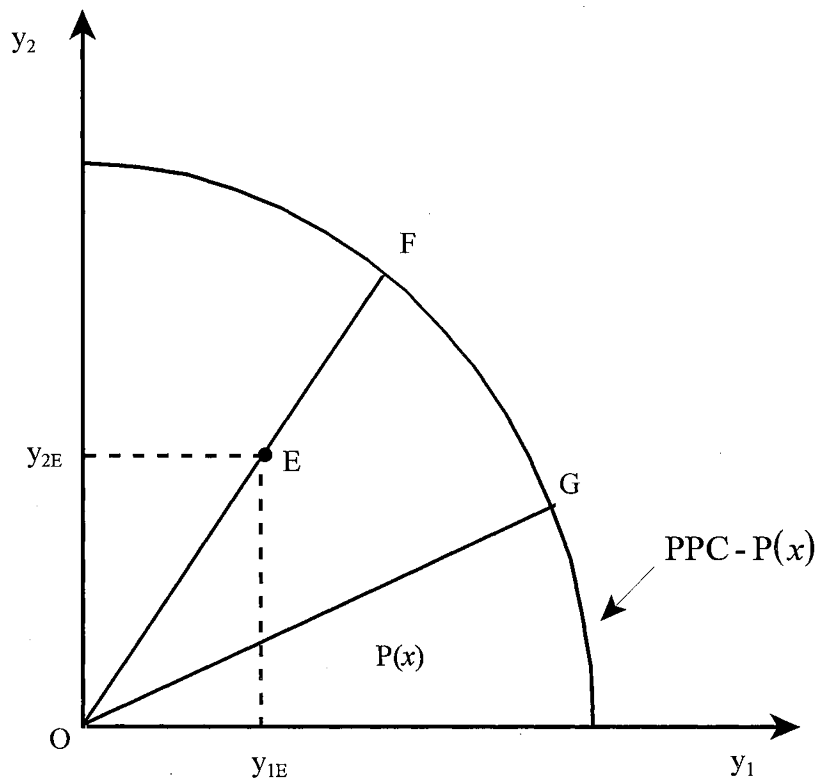

Figure 2.4
Output Distance Function and Production Possibility Set

The production possibility set in Figure 2.4, P(x), is the area bounded by the production possibility curve (or frontier), PPC-P(x), and the y_1 and y_2 axes. For a firm operating at point E in the production possibility set, P(x), using the given input level x_1 to produce the outputs y_1, y_2, the *value* of the distance function for the firm is equal to the ratio:

$$\rho = OE/OF \tag{2.15}$$

The firm can operate at point F, which is on the frontier with the given input level x and produce more of both outputs. Hence the value of distance function, ρ, for operating at point F (and G) is equal to 1. Therefore, the firm that operates at point E can increase both its output quantities by moving to point F and still remain within the feasible production possibility surface, i.e., PPC – P(x). Note, the output distance function in equation (2.15) is exactly the inverse of Farrell's (1957) output oriented technical efficiency measure.

The input distance function is defined in a similar manner. The value of the distance function is exactly the inverse of Farrell's (1957) input oriented technical efficiency measure, which can be explained using the isoquants. Distance functions can be estimated directly by econometric methods or mathematical programming methods.

2.3 Modeling Production

This section contains a discussion of issues related to modeling the underlying relationships between inputs and output in a production process.

2.3.1 Average Response vs. Frontier Functions

The application of empirical estimation techniques for production and cost dates back to the work of Cobb - Douglas in 1928. Since then many others have attempted to elaborate on the subject of production (e.g., Dean, 1951; Johnson, 1960) and cost structure (e.g., Nerlove, 1963). There is also literature on the potential use of the duality between production and cost functions (e.g., Cornes, 1992).

Before Farrell's (1957) introduction of the *frontier* approach in estimating the dual functions' efficiency, the linear average mean response functions (production or cost) were estimated using least squares (OLS) or some variant thereof. The *average* functions do not necessarily represent the *best technology*, and therefore, there is an explicit conceptual link missing between microeconomic definitions of production or cost functions and what is being estimated. The *average* function assumes that *all firms are* *efficient*, which is generally not a reasonable assumption. A *frontier* function is a bounding function against which inefficiency or the relative size of one-sided deviations from the maximum output or the minimum cost can be estimated.

If the relationship between inputs and output is estimated using OLS, a line satisfying the least squares assumptions is fitted through the data. The estimated line represents the *average function* (Figure 3.1) and does not necessarily catch the best technology.

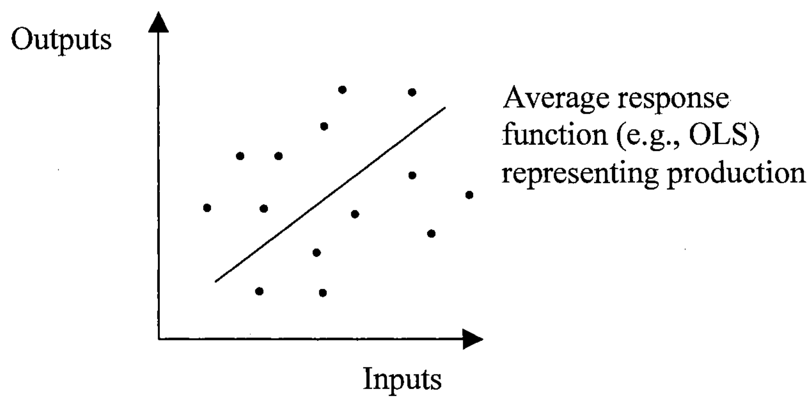

Figure 3.1
Fitting an Average Function

Farrell (1957) introduced the *frontier* approach in estimating the dual functions' efficiency. The concepts of frontier cost (Figure 3.2) and production (Figure 3.3) are shown below. The idea is to fit a line so that all the observed points are above the line in the case of cost or below the line in the case of production. The vertical distance from each point to the cost or production frontier represents inefficiency.

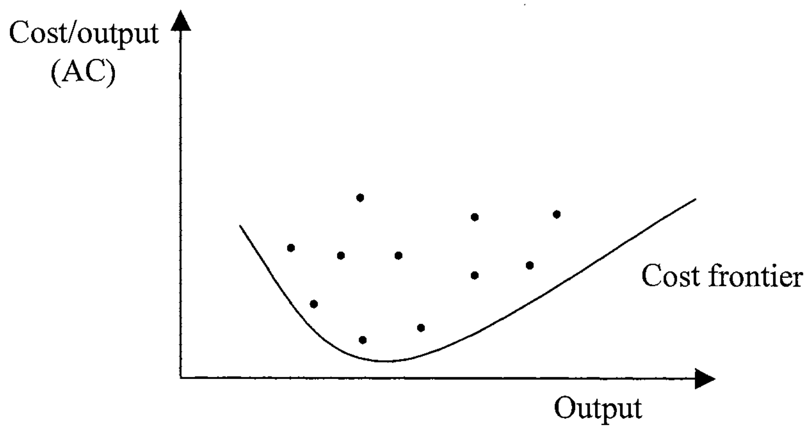

Figure 3.2
Cost Frontier with all the Observations on or Above the Frontier Boundary

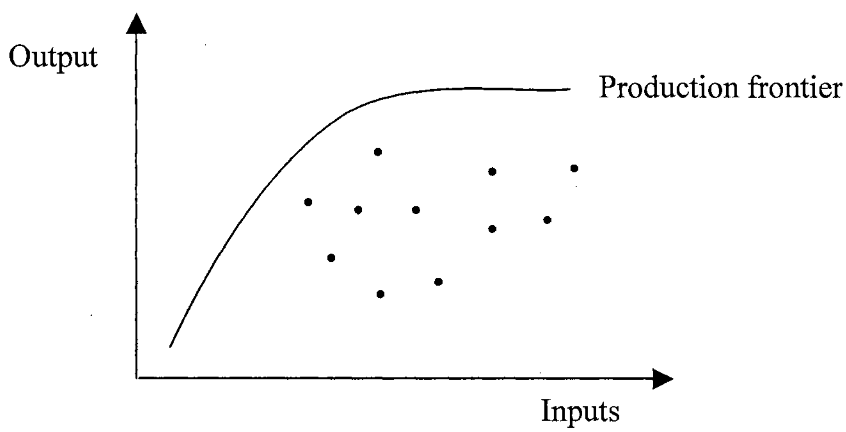

Figure 3.3
Production Frontier with all the Observations on or Below the Frontier Line

Examining Farrell's frontier concept (1957), and efficiency measurement via a geometric presentation is revealing. Farrell argues that *average* function does not accord with the standard definition of the production function. This argument led to the use of unit isoquant and isocost lines to explain his idea of the frontier production function and to show how it helps to arrive at an efficiency measurement.

Farrell (1957) considered a model with two inputs (x_1 and x_2) which produce a single output (y). Under the assumption of constant returns to scale, the technology is represented by unit isoquant and isocost.

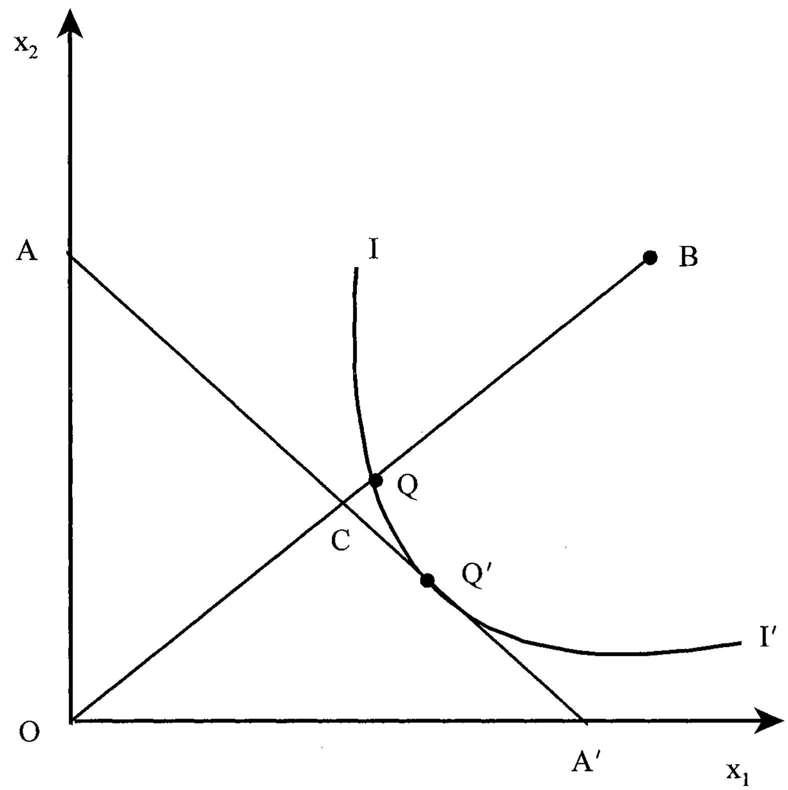

Figure 3.4
Farrell's Presentation of Technical and Allocative Efficiency

The unit isoquant of a fully efficient firm is represented by II′ in Figure 3.4. If a firm is operating at point B then the firm operates inefficiently. This follows because it could produce at Q, which yields the same output with fewer of both inputs. This technical efficiency (TE) can be measured by a distance OQ/OB.

$$TE = OQ/OB = 1 - QB/OB$$

TE can take values between 0 and 1, and hence the degree of inefficiency of the firm can be measured. This would imply that as B approaches Q, technical efficiency approaches one.

To explain allocative efficiency, Farrell adds an input price ratio (slope of the isocost line) represented by AA′ in the figure. The allocative efficiency (AE) when the firm operates at B, can be measured by distance OC/OQ.

The distance CQ represents the reduction in cost if the firm operates at Q′ which is a technically and allocatively efficient point. Then, the total economic efficiency (EE) is then:

$$EE = OC/OB$$

The product of TE and AE is total efficiency.

$$TE \text{ x } AE = (OQ/OB)(OC/OQ) = OC/OB = EE$$

The above efficiency measures are in the context of an input-oriented measure, i.e., given a desired output quantities, how much can input quantities be reduced without changing the output quantities? Alternatively, one can approach the question with an output-oriented (rather than input) measure in mind, i.e., how much output quantities can be expanded with a desired level of input quantities?

If the firm's production is governed by constant returns to scale (CRS) the two measures are equal, i.e., input-oriented measures = output oriented measures (Färe and Lovell, 1978). However, with varying returns to scale the two measures are not equal. To illustrate, assume output y is produced with a *single* input x with $y = f(x)$.

The following graphs in Figures 3.5 and 3.6 represent the input- and output-oriented technical efficiency measures with decreasing returns to scale (DRS) and constant returns to scale (CRS) technologies.

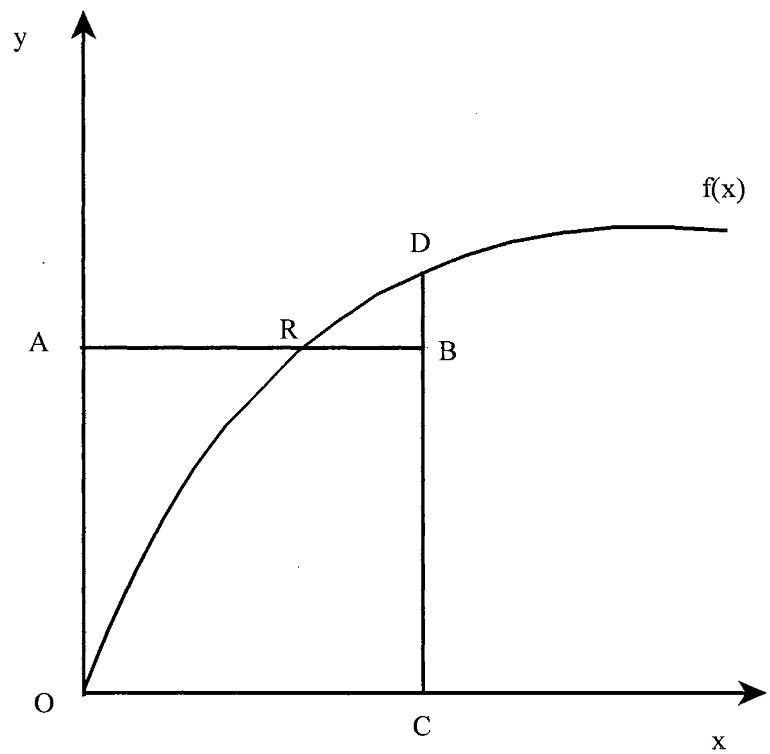

Figure 3.5
Input and Output-Oriented Technical Efficiency Measures with DRS

In Figure 3.5, the inefficient firm is operating at a point below the PPC, point B. Farrell's input-oriented technical efficiency (TE) is:

$$TE_I = AR/AB$$

At point B the firm is producing y = OA and fewer inputs are required by the firm to produce here if, it is technically efficient.

The output-oriented technical efficiency:

$$TE_O = CB/CD$$

and is based on the fact that with input usage at x = OC , the efficient firm can produce more output (BD). In general for DRS technology, $TE_I < TE_O$. Alternatively, when production is governed by constant returns to scale (CRS), the two measures are equivalent, i.e., AR/AB = CB/CD then $TE_I = TE_O$ (Färe and Lovell, 1978).

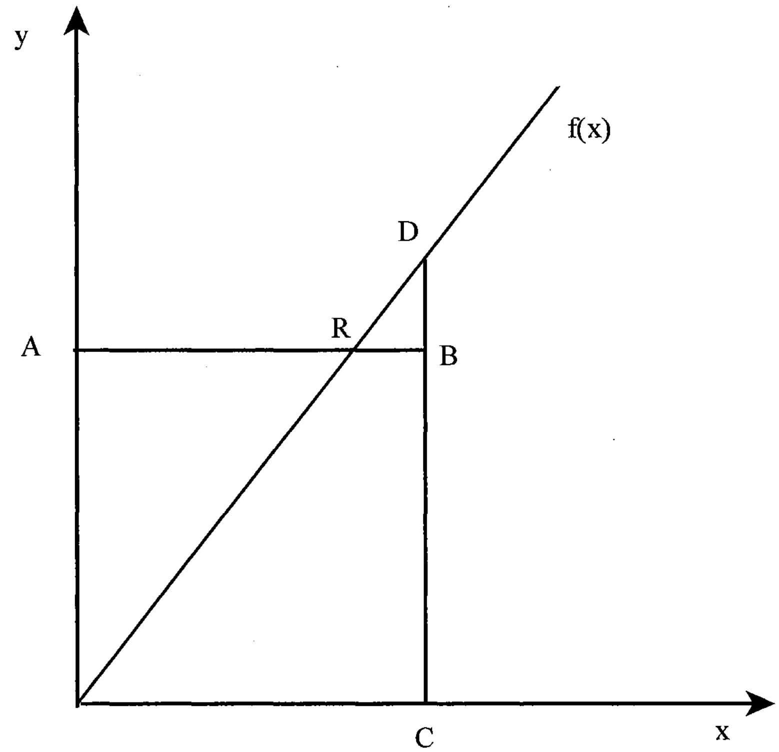

Figure 3.6
Input- and Output-Technical Efficiency Measures with CRS

If output-oriented measures of technical, allocative or economic efficiency (*TE, AE, EE*) are of interest, one can account for them by using production possibility curves and isorevenue lines (see Coelli, 1998 for details). All of these measures are between zero and one like the input-oriented measures

Battese (1992) proposed a more general case of Farrell's frontier concept. The focus of the study was on the technical efficiency rather than Farrell's total efficiency.

He used the production function explicitly in an output-input space rather than, inputs per unit of outputs space.

2.3.2 Functional Forms in Economic Analysis of Production

In this section the objectives of production analysis are discussed. The objectives are the underlying motivation for different functional forms. If one wishes to empirically measure certain quantities of interest in production, then the researcher must choose a specific parametric functional form to estimate. The choice of functional form is governed both by practical considerations and, by the specific objectives of the research. It is generally believed that the form chosen should be consistent with the underlying data generation process of the system under study. The principle objectives of production analysis are discussed in Fuss, McFadden, and Mundalk (1978). These objectives include determining:

1. distribution parameters, e.g., income share of factors of production. These parameters are important in determining the incidence of tax and subsidy programs as well as evaluating the economic growth.
2. the returns to scale properties of technologies, e.g., the supply and financing of public services often center on the question of the existence of increasing returns to scale.
3. the degree of substitutability between factors of production. Substitutability is critical in many areas including determination of tax incidences.
4. whether production relationships may be decomposed into nested or additive components, i.e., separability. This separability allows econometricians to carry out their analysis in terms of subsets of variables instead of a total set of

possible variables. Since separability influences generality and simplicity of functional forms, it is crucial for empirical testing.

5. the existence of technical change.
6. the relative efficiency, i.e., relative efficiency of firms to the frontier technology.
7. whether a production function is homothetic, i.e., is the expansion path linear through the origin?

There are several important criteria for designing functional forms that are to be estimated. A suitable functional form should obey or satisfy:

1. maintained hypotheses of basic axioms on the nature of technology that are widely accepted (e.g., monotenicity, concavity, symmetry, etc.).
2. technological and behavioral assumptions, which may be relevant to the particular problem at hand. For example *a priori* knowledge of constant returns to scale.
3. some innocuous simplifying assumptions that may facilitate the analysis, such as the independence of error terms.

Also, the functional form should be a *parametric* (as opposed to non-parametric) functional form for the sake of convenience and tractability.

Other practical considerations come into play when working with parametric production functions.

1. The functional form should contain only the parameters that are necessary (parsimony in parameters). Excess parameters can create severe multicollinearity. In addition, when small samples are considered, excess

parameters are associated with a loss of degrees of freedom. Parameter necessity is usually a matter of judgment and differences in opinions may lead to model uncertainty.

2. The functional forms should be clear and easy to interpret. When dealing with complex functional forms, some economic interpretations can be difficult.
3. Functional forms should be chosen with computational ease in mind. Many empirical studies have used statistical models, which are linear in parameters because of their computational ease. Imposing linearity on an inherently nonlinear process can be very misleading; especially in predicting none sample events.

In an attempt to minimize the effect of rigid structure, the so-called *flexible* functional forms have gained widespread use.

2.3.3 Alternative Functional Forms

In this section we discuss the properties of some common functional forms used in econometric estimation.

Cobb-Douglas Functional Forms

The Cobb-Douglas functional form has been popular in the empirical estimation of the frontier model. This is due to the fact that the Cobb-Douglas function is easy to estimate and a logarithmic transformation makes the model linear in logarithm of the inputs. However, this attractive feature imposes a number of restrictions. For example, the elasticities of substitution are constant and equal to one, i.e., inputs are perfect substitutes. In addition, returns to scale properties for all the firms in the sample are identical.

Translog Functional Forms

The translog functional form was first introduced by Christensen, Jorgenson, and Lau (1973). This function is a direct generalization of the Cobb-Douglas function. However, unlike the Cobb-Douglas function, elasticities of substitution need not equal to one and no (sample wide) restriction upon returns to scale is imposed. Translog forms do not *necessarily* satisfy concavity, monotonicity, or other important axioms of production economics. According to Terrell (1996), concavity and monotonicity can be imposed at the cost of lost flexibility. In fact, classical econometric methods have no elegant way of imposing these types of restrictions; however, this is fairly easily accomplished through Bayesian analysis. Therefore, the benefit and cost of imposing these restrictions has to be carefully considered.

Generalized Leontief and Generalized McFadden Functional Forms

Diewert (1971) introduced the generalized Leontief system, which is a *flexible* functional form that satisfies basic axioms and other maintained hypotheses. Diewert (1974, p. 113) defines a *flexible* functional form for a cost function. If a cost function provides second order differential approximation to an arbitrarily twice continuously differentiable cost function that satisfies the *linear homogeneity in prices* at any point in the domain, then it is a *flexible* functional form.

Diewert (1974) utilized the Shephard duality theorem in order to obtain derived demand equations, linear in technology parameters, for ease of econometric estimations.

He chose the quadratic form in the square roots of the input prices, thus a generalization of the Leontief cost function. One feature of this function is that any set of partial elasticities of substitution can be obtained with a minimal number of parameters. A generalized Leontief production function can attain an arbitrary set of shadow elasticities of substitution where a generalized linear function attains an arbitrary set of direct elasticities of substitution at a given set of inputs and input prices. Furthermore, Diewert constructs a functional form for production function using the Shephard duality theorem. The significance of this theorem is that the cost function can be interpreted as the total cost function of some underlying production function keeping in mind that the production function may not always be expressed explicitly. By estimating the parameters of the cost function, one can be assured that these parameters are equivalent to estimating the parameters of the underlying production function, assuming the firm is operating in a competitive market.

Another generalization of a functional form due to McFadden (1978), *generalized McFadden*, also received attention by researchers. Like generalized Leontief, generalized McFadden is a cost function; therefore, it uses duality theory to obtain derived demand equations.

Diewert and Wales (1987) compared translog, generalized Leontief and generalized McFadden cost functions. They established that the generalized Leontief cost function and generalized McFadden cost function are equivalent in terms of flexibility, ease of estimation, and hypothesis testing.

The forms and some of the properties of these functional forms are summarized in Table 2.1.

Table 2.1
Forms and Some Properties of Alternative Functional Forms

Production Function	Homogeneity	Elasticity of Sub.
Cobb-Douglas $\log y = a_0 + \sum_{i=1}^{n} a_i \log x_i$	Homogeneous	$\sigma = 1$
Translog* $\log y = a_0 + \sum_{i=1}^{n} a_i \log x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} (\log x_i)(\log x_j)$	Not Homogeneous	σ not a constant
Generalized Leontief* $y = a_0 + a_1 x_1 + a_2 x_2 + b_1\sqrt{x_1} + b_2\sqrt{x_2} + b_3\sqrt{x_1 x_2}$	Not homogeneous	σ not a constant[1]
Generalized McFadden* $y = \sum\sum x_j \phi^{ij}(x_i \mid x_j) a_{ij}$	Not homogeneous	σ not a constant[1]

* *Flexible* functions
[1] For a general computation of elasticity of substitution, see Fuss et al. (1978), p. 231.

2.4 Estimation

In this section, several issues related to estimation of the underlying relationships are discussed. Following Farrell's introduction of the production frontier, a number of authors have discussed the process by which the frontier may be estimated.

Let's begin with a discussion related to estimation of a deterministic production function. Following Aigner and Chu (1968) let output be determined as a function of inputs for a given firm:

$$y_i^* = f(x_i; \beta)$$

where

y_i^* is the maximum output attainable for the *i*th firm

x_i is the (fixed) vector of inputs

β is a vector of unknown parameter to be estimated.

The parameters of the model can be estimated substituting the observed output, y_i, for y_i^* and minimizing:

$$\sum_{i=1}^{N} \left| y_i - f(x_i;\beta) \right|$$

subject to:

$$y_i \le f(x_i;\beta).$$

By doing so an implicit disturbance term has been assumed.

Aigner and Chu estimate a Cobb-Douglas *parametric* frontier production function, using data on a sample of N firms. Their model is defined as:

$$\ln(y_i) = x_i\beta - u_i \qquad i = 1,2,\ldots,\mathrm{N}$$

where:

$\ln(y_i)$ is the logarithm of the (scalar) output for the *i*th firm.

x_i *is* a $(k + 1)$ row vector, with first element being "1" and the remaining elements being the logarithm of the *k* input quantities for the *i*th firm.

$\beta = (\beta_0, \beta_1, .., \beta_k)'$ *is* a $(k+1)$ column vector of unknown parameters to be estimated.

u_i is a non-negative random variable associated with technical inefficiency in production of the *i*th firm.

The technical efficiency is defined by the ratio of actual output of firm *i* and its potential output, i.e., the output frontier conditional on the input vector x_i is:

$$\text{TE}_i = \frac{y_i}{\exp(x_i\beta)} = \frac{\exp(x_i\beta - u_i)}{\exp(x_i\beta)} = \exp(-u_i)$$

TE_i is the output-oriented Farrell technical efficiency measure with a value between zero and one. In their model the disturbance term, u_i, measures technical inefficiency due to factors that are under the firm's control, i.e., disturbances that are associated with the producer's and employees' effort and damaged products, etc.

The shortcoming of the deterministic model is that it does not account for other sources of disturbance *noise*, such as measurement error, luck, climate, etc., which are not under a firm's control and can be favorable or unfavorable to the firm's output. This suggests that frontiers themselves can vary for different firms or over time for the same firm. Then the productive efficiency should be measured by:

$$y_i / (f(x_i, \beta) + v_i)$$

where v_i is the disturbance associated with noise. Aigner, Lovell, and Schmidt (1977) suggested a new model that takes v_i into consideration. Since the economic logic suggests the specification of the disturbance term, then:

$$\ln(y_i) = x_i\beta + \varepsilon_i \qquad i = 1, 2, \ldots, N \qquad (2.16)$$

where $\varepsilon_i = v_i - u_i$

substituting for ε_i in equation (16):

$$\ln(y_i) = (x_i\beta + v_i) - u_i$$

where:

v_i is the symmetric disturbance term, $iid, \text{N}(0, \sigma_v^2)$,

u_i is truncated above at 0, and distributed as $N(0, \sigma_u^2)$,

u_i's and v_i's are independent of each other.

The model defined by equation (2.16) is a stochastic frontier production function, i.e., the output is bounded above by the stochastic frontier exp $(x_i\beta + v_i)$, rather than $\exp(x_i\beta)$ the deterministic frontier.

Using the stochastic frontier approach allows one to estimate the variances of v_i and u_i, and their relative sizes can be measured. Based on these results, firms can make more accurate decisions concerning their production process.

A stochastic frontier model can be illustrated in Figure 4.1.

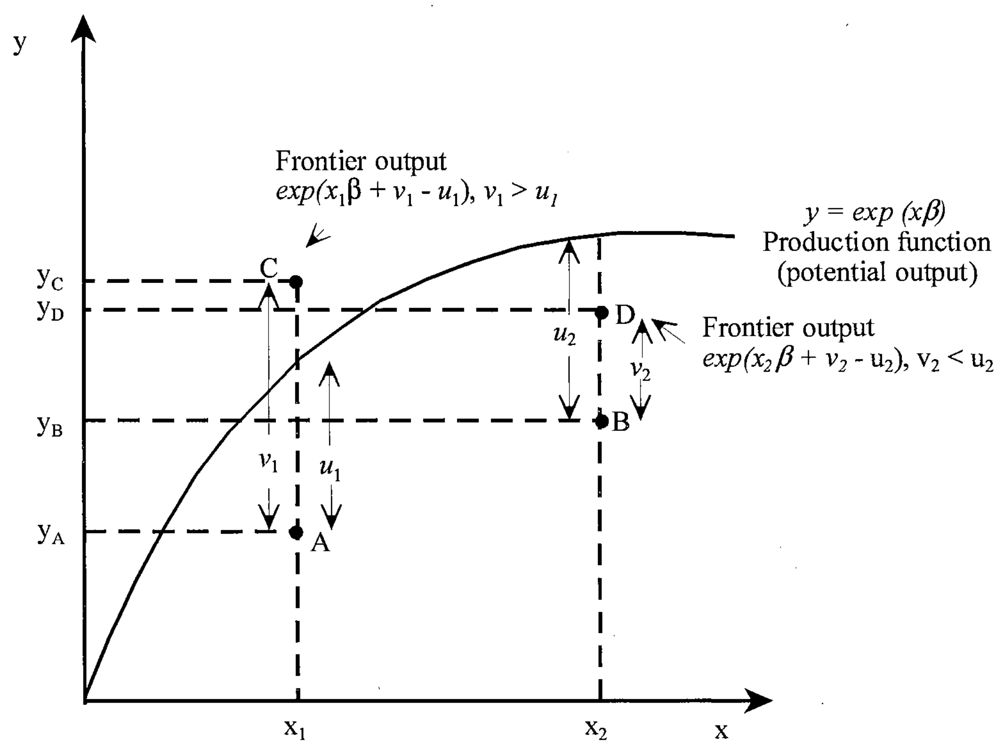

Figure 4.1
The Stochastic Frontier Production Function

where: y is output

x is input.

For a deterministic production function, $y = \exp(x\beta\text{-}u)$, the observed input-output values for firms 1 and 2 are denoted by points A and B, respectively.

Assuming $v_1 > u_1$ (since $u > 0$), the value of stochastic frontier output for firm 1 is shown by point C. In addition, if $v_2 < u_2$, the value of stochastic frontier output is shown by point D. The observed outputs will be above the deterministic frontier if $v_i > u_i$ and below the frontier if $v_i < u_i$, e.g.,

$$y_i > exp(x_i\beta) \text{ if } v_i > u_i$$

$$y_i < exp(x_i\beta) \text{ if } v_i < u_i$$

The use of the maximum-likelihood (ML) method to estimate, standard errors and perform hypotheses tests is possible only with a stochastic frontier model because, the deterministic models violate certain maximum likelihood regularity conditions, i.e., the range of random variables cannot depend on unknown parameters (Coelli, 1998). Greene (1980a) suggested a particular class of distributions for u_i's,, which could eliminate the irregularity of the likelihood function associated with deterministic frontier models. However, Greene's model does not account for random errors.

The stochastic frontier models suffer from the common criticism that there is no *a priori* justification for selection of any particular distributional form for the u_i's (Coelli, 1998). Stevenson (1980) assumed truncated normal distribution for u_i's. Greene (1990) assumed two parameter gamma distribution for u_i's. Although their assumptions eliminate regularity problems, the efficiency measures may be sensitive to the distributional assumptions. However, there is no hard evidence that the choice of distributional assumptions of u_i's have a significant effect on predicted technical efficiencies (C.A.K Lovell, 1995 in Coelli, 1998 p.187). The specific distributional assumptions of disturbance will be discussed later in this chapter.

2.4.1 Modeling Stochastic Frontiers

Typically, least squares or some variant of it such as two-stage or generalized least squares have been used to estimate parameters of interest. However, it has been argued that the distributional assumptions of the models' errors and the estimation techniques used to estimate the models' parameters are not consistent with the microeconomic definition of production and cost functions (Greene, 1980). The parameter estimates can be made consistent with a simple modification of ordinary least squares; however, since the distribution of the error term is asymmetric then OLS estimates are not efficient. A maximum likelihood estimator with asymmetric error is, asymptotically, more efficient.

Aigner and Chu (1968) pioneered the estimation of deterministic frontier functions models. Since a stochastic frontier model contains deterministic as well as stochastic disturbance terms, their results are of great importance. Following Aigner and Chu (1968), assume the following frontier model where all residuals are negative

$$y = ax_1^{\alpha_1} x_2^{\alpha_2} u,$$

where:

y is output,

x_1, x_2 are inputs,

u is a random disturbance,

$\alpha_1, \alpha_2 > 0$.

They suggest minimizing the sum of absolute residuals subject to the constraint that residuals have to be negative. In this case a non-parametric functional form is assumed. This is a linear programming problem, which will be discussed later in this chapter. Schmidt (1975) estimated parameters of Aigner and Chu's model using maximum likelihood under the assumption that the disturbances have an exponential distribution and showed that the two estimates are equivalent.

Alternatively, Aigner and Chu (1968) minimized the sum of squared residuals rather than the absolute residuals, but with the same constraint. This is a quadratic programming problem. Schmidt (1975) estimated the parameters of this model using ML under the assumption that the errors are half normal and showed that the two estimates are equivalent. Unfortunately, the asymptotic properties of these maximum likelihood estimators are not well established and provide no guidance in formulating or estimating the standard errors (Greene, 1980a). This problem arises because of the fact that the regularity conditions of maximum likelihood are violated.

Schmidt (1975) suggests that irregularity arises from the fact that the range of the observed random variable depends on the parameters being estimated. According to Green (1980a,p.36), aside from the range problem, if one is willing to assume that the density function of the observed random variable is continuous everywhere over its range and all the appropriate derivatives exist and are finite, the two likelihood estimators are well behaved. Large samples properties as well as standard analysis of better-behaved problems are not affected by the range problem. In these cases the information matrix can be computed to form the standard errors of the estimator. Therefore the maximum likelihood estimator (MLE) will be consistent, asymptotically efficient and asymptotically normally distributed.

It is also worth mentioning that the gamma distribution for the disturbance term, which is asymmetric, makes MLE of the parameters of the model more efficient than OLS. The gamma density has very useful features in specification and estimation of frontier production functions. The maximum likelihood estimator of a gamma distribution has all of the usual desirable properties MLE's. Therefore, the asymptotic distribution of the estimator can be derived and the asymptotic variance matrix is easily estimated.

The additional efficiency of the maximum likelihood estimator over OLS is not artificially built into the model. The efficiency depends on the degree of skewness of the error distribution away from the frontier, i.e., when accounting for asymmetry, the higher the degree of skewness, the higher the efficiency gain by using the maximum likelihood method.

In an application of different methods, Greene (1980a) uses OLS and maximum likelihood estimators to estimate the stochastic frontier of Aigner et al. (1977). In this model, the disturbance term ε is defined as $\varepsilon = v - u$ where v is $N(0, \sigma_V^2)$ while u has a half-normal distribution. Therefore, ε has an asymmetric distribution. For the deterministic frontier (where the disturbance u is due to inefficiency) Greene assumed gamma density and uses the maximum likelihood estimator. His results indicate that the maximum likelihood estimates of the stochastic frontier function's parameter are quite close to OLS estimates, while the estimates of the deterministic frontier using the same estimators are very different.

Furthermore, he employs the same estimators for a cost frontier. Following Nerlove (1963), who uses a generalized 3 input Cobb-Douglas cost function, Greene concludes that, in terms of symmetry of the error distribution, the skewness coefficient of the cost function is much smaller than in the case of production function, suggesting that the error distribution associated with a cost function is closer to being symmetric.

2.4.2 Data Envelopment Analysis (DEA)

There are two parallel approaches to the estimation of stochastic frontier models, the econometric (statistical) and mathematical programming (non-statistical) approach. In the previous sections the econometric approach which requires a parametric production function was considered.

In this section attention is turned to the mathematical programming approach and conclusions are drawn based on the comparison of the two methods. As mentioned earlier, the efficiency measures discussed assume that production function is known which in practice is not the case and the efficient isoquants must be estimated. Farrell (1957) suggested the use of a non-parametric piece-wise-linear convex isoquant, where no observed point lies below or to the left of it in the context of constant returns to scale (Figure 4.2).

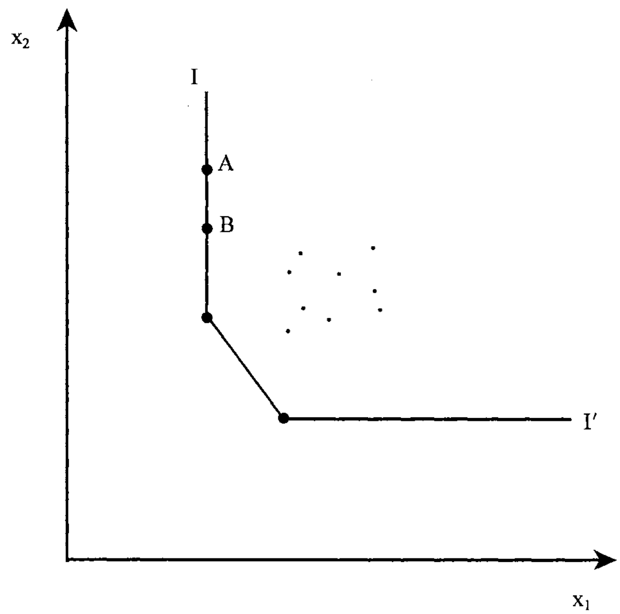

Figure 4.2
Piece-Wise Linear Convex Unit Isoquants

Following Farrell's input-oriented frontier model with constant returns to scale, Charnes, Cooper, and Rhodes (1978) pioneered the concept of data envelopment analysis (DEA), which involves the use of linear programming.

Coelli (1998) states that a natural measure of performance is the productivity ratio:

$$\text{productivity ratio} = \frac{\text{outputs}}{\text{inputs}}$$

Following Coelli's notations, if data are available on K inputs and M outputs for N firms the data for the *i*th firm is represented by column vector x_i of inputs and column vector y_i of outputs. X and Y represent the input and output matrix for the data for all firms, respectively. For each firm, the measure of the ratio of all outputs over all inputs is denoted as, $u' y_i / v' x_i$ where u is an Mx1 vector of output weights and v is a Kx1 vector of input weights. Solving the following linear programming problem for the *i*th firm yields the optimal weights:

$$\max_{u,v} \left(u' y_i\right)$$

subject to:

$$v' x_i = 1$$

$$u' y_j - v' x_j \leq 0 \qquad j = 1, 2, \ldots, N$$

$$u,\ v \geq 0$$

Note that the above maximization with $v' x_i = 1$ prevents u and v from having an infinite number of solutions. This is called the multiplier DEA linear programming problem.

Another approach involves the use of duality, which gives an equivalent result but with fewer constraints than the multiplier DEA. Hence is usually preferred and is called the envelopment form:

$$\min_{\theta,\lambda} \theta,$$

subject to:

$$-y_i + Y\lambda \geq 0,$$
$$\theta x_i - X\lambda \geq 0,$$
$$\lambda \geq 0,$$

where θ is a scalar of inefficiency score of the *i*th firm and λ is a Nx1 vector of constants. Note that $\theta \leq 1$. If $\theta = 1$, then the firm is operating on the frontier and hence is technically efficient in accordance with Farrell's definition of efficiency.

If all the firms are not operating at an optimal level, then variable returns to scale as opposed to constant returns to scale is the appropriate assumption. To solve the linear programming problem for variable returns to scale (VRS), a convexity constraint, $NI'\lambda = 1$,is added to the constant returns to scale (CRS) model, where NI is a Nx1 vector of ones. Then solve:

$$\min_{\theta,\lambda} \theta,$$

subject to:

$$-y_i + Y\lambda \geq 0,$$
$$\theta x_i - X\lambda \geq 0,$$
$$NI'\lambda = 1,$$
$$\lambda \geq 0,$$

When doing so, the technical efficiency scores of the VRS technology are obtained. These scores are greater than or equal to the scores of CRS technology.

Note that **VRS TE – CRS TE = scale inefficiency** and not technical inefficiency. If NI $'\lambda \leq 1$ then the ith firm is compared to firms that are smaller or equivalent in size and not firms that are substantially larger. Changing the convexity constraint of NI $'\lambda = 1$ to NI $'\lambda \leq 1$ allows one to calculate the scale economy. If VRS TE is equal to non-increasing returns to scale technology, then there are decreasing returns to scale. Otherwise increasing returns to scale are present.

One of the piece-wise linear forms of the non-parametric DEA problem is the input-slack problem. This slack is related to the parts of the frontier that are parallel to the axis (Figure 4.2). The problem arises from the fact that even though point A is on the frontier, we can move to B, which is a reduction of input x_2 and still be efficient. A number of studies suggest ways to deal with these slacks (e.g., Ali and Seiford, 1993; Coelli, 1997).

Environmental variables can influence efficiency. These variables are not traditional input variables and firms do not have control over them. To account for these variables, Coelli (1998) suggests that variables can be directly included into the linear programming problem as discretionary inputs.

If the direction of the influence of the environmental variable upon efficiency is known, then the parameters can be estimated in a single stage model. If the direction of the influence of the environmental variable upon efficiency is not known, then a two-stage procedure could be used to determine the direction of the influence upon efficiency. Once the direction of the influence is known, the procedure follows that of a single-stage model.

DEA is favored by Bessent et al. (1980) and others because:

- it makes no parametric assumptions about the relationship between a firm's inputs and output.
- it easily allows for multiple outputs, which can be related to inputs linearly or otherwise.

The disadvantages of DEA are:

- It assumes that all deviations of output from its potential are due to inefficiency. If any noise is present, it may influence the efficiency measurement of DEA more than the stochastic frontier approach.
- DEA does not provide the means for hypothesis testing regarding the existence of inefficiency or the structure of the production technology. This is because mathematical programming techniques have estimators with unknown statistical properties (Aigner et al., 1977).

2.4.3 Modeling Inefficiency

Assume a stochastic frontier production function in which the error term is specified as:

$$\varepsilon_i = u_i - v_i$$

where: ε_i is the disturbance term

u_i is the deterministic component of the error term with a suitable distributional assumption, (e.g., truncated normal, gamma, etc.) associated with technical inefficiency.

v_i is the random component of the error term

and u and v are independent of each other where $u_i > 0$

The inefficiency effects u_i in the stochastic frontier model can be explicitly modeled as:

$$u_i = z'_i\delta + w_i$$

where: z_i is a vector of explanatory variables

δ is a vector of unknown parameters to be estimated

w_i is a vector of random variables

Alternatively we can rewrite the efficiency model as:

$$\text{TE}_i = z'_i\delta + w_i$$

where TE_i represents the technical efficiency of firm i.

In this section, 3 alternative estimation procedures of inefficiencies will be discussed in the contexts of the econometric approach (two-stage, MLE) and the mathematical approach (DEA).

Two-Stage Approach (OLS)

To investigate the determinants of technical inefficiency, Pitt and Lee (1981) suggested a two-stage procedure. In the first stage inefficiencies are predicted from an estimated stochastic production function. In the second stage, the predicted inefficiencies are regressed over a vector of firm-specific variables (e.g., age, experience, and education of the firm's managers) using OLS.

According to Kumbhakar, Ghosh, and McGuckin (1991) there are two basic econometric problems associated with the two-stage procedure. First, technical inefficiency and inputs may be correlated. This correlation causes inconsistent estimates of the parameters as well as technical inefficiency in the first stage.

The use of OLS in the second-stage regression is inappropriate. OLS ignores the fact that technical inefficiency (the dependent variable in the inefficiency model) is inherently one-sided (non-negative).

Greene (1980a) considers the following production function:

$$y_t = \alpha + \beta' x_t + \varepsilon_t \qquad t = 1, \ldots, T$$

where y_t is output,

x_t is the corresponding vector of exogenous variables,

ε_t is a random disturbance,

α, β are fixed and unknown parameters to be estimated,

and T is the sample size.

The errors, ε_t differ from zero due to random shocks (Aigner and Chu, 1968). The disturbance in the model typically results from logarithmic transformation of

$exp(y_t) = f(x_t)\, u_t$ where

$\varepsilon_t = -\ln(u_t)$ and $0 < u_t \leq 1$.

The only parameter that is not consistently estimated by OLS is α. The OLS intercept estimator is consistent for $(\alpha + u^2)$. If the analysis of efficiency is desired, the consistent estimation of α is necessary. In some instances, the least squares residuals can provide a consistent *moment* estimator of α, i.e., $\hat{\alpha}$. However, even after the correction of $\hat{\alpha}$ for non-zero disturbance mean, some of the residuals may still be negative.

The computation of meaningful efficiency measures requires all of the residuals to be positive. Greene (1980a) proposes a biased but consistent estimator of α which forces the positive sign on the residuals that can be easily obtained. Greene (1980a, p. 32) offers a detailed explanation with regard to the estimation of α.

In an attempt to investigate the determinants of inefficiency in Oklahoma school districts, Abdul Rahman (1996) adopted a two-stage procedure. Following Schmidt and Sickles (1984), Abdul Rahman, defined the frontier model for panel data as

$$y_{it} = \alpha + x'_{it}\beta + v_{it} - u_i$$

where:

$i = 1, \ldots, N$	is school district,
$t = 1, \ldots, T$	is time period,
y_{it}	is the output of school district i at time t,
x_{it}	is a 1xK vector of inputs
u_{it}	is a one-sided disturbance term representing technical inefficiency of the school district,
v_{it}	is a two-sided random disturbance which is uncorrelated with the regressors.

and u_{it} is iid with mean u and variance σ^2. u_{it} and v_{it} are independent of each other.

To assure that u_{it}'s have positive sign, (Schmidt and Sickles, 1984), let $E(u_i) = u > 0$ and define α^* and u^* as

$$\alpha^* = \alpha - u, \qquad u^* = u_i - u$$

so that u_i^* are iid with mean 0. The model then becomes:

$$y_{it} = \alpha^* + x'_{it}\beta + v_{it} - u_i^* .$$

With the two-stage procedure inputs can be separated into school and nonschool inputs. In the first stage school inputs were used to estimate the education production and obtain the relative efficiency. In the second stage the effect of nonschool factors on the efficiency was estimated.

To estimate the model, Abdul Rahman (1996) used the Fixed Effects estimator (Least Squares Dummy Variables) as well as Random Effects model (Generalized Least Squares) and compared the two estimates.

Maximum Likelihood Approach (MLE)

Kumbhakar, Ghosh, and McGuckin (1991) assume that the technical inefficiency effects are the non-negative truncation of a normal distribution. The mean of the inefficiency effects is a linear function of exogenous variables with unknown coefficients and an unknown variance.

Kumbhakar et al. (1991) developed a single-step MLE technique and claimed that this technique avoids the usual criticism targeted at the two-stage procedure. Many researchers such as Reifschneider and Stevenson (1991), Huang and Liu (1992), Coelli (1995a), and Battese and Coelli (1995) have also used ML procedures, preferring them over the two-stage methods.

Coelli (1995a) conducted a Monte Carlo study to investigate the finite sample properties of corrected ordinary least squares (COLS) and the MLE of Aigner et al. (1977) half-normal stochastic frontier production. The COLS estimator is basically an OLS estimator but requires the shifting of the intercept term to reach the frontier.

Aigner et al. (1977) reparameterize the variances of the deterministic error component u and the stochastic error component using $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\lambda = \sigma_u / \sigma_v$.

Battese and Corra (1997) replaced λ with, $\gamma = \sigma_u^2 / \sigma^2$. The parameter γ is the contribution of inefficiency error to the total error and takes a value between zero and one. This reparameterization simplifies computation of the MLE. Coelli's (1995a) results suggest serious bias in both ML and COLS estimates when γ is small. However, if γ is greater than 50 percent the ML estimator has a smaller mean square error and should be used rather than the COLS estimator.

DEA with Second-Stage Tobit Approach

Schmidt (1975) showed that the Aigner and Chu (1968) linear programming estimator is the maximum likelihood if the disturbance term of the model has an exponential distribution. If the disturbance term has a half-normal distribution, then the quadratic programming estimator of Aigner and Chu is the maximum likelihood. As discussed earlier in this chapter, one of the more serious shortcomings of the DEA approach is that the sampling distribution of DEA estimators are unknown. Therefore, their statistical properties cannot be identified. This implies that the standard errors cannot be derived and thus, it is not possible to make any statistical inferences based on these estimates. Since DEA model estimates deterministic frontiers (i.e., no account for the random disturbance) and, the MLE of a deterministic frontier suffers from an irregularity problem, then, the fact that the programming estimators are MLE is not sufficient to establish their statistical properties.

Kirjavainer and Loikkaner (1998) studied the efficiency of Finnish senior secondary schools with DEA. The authors acknowledge that the DEA model does not account for one of the most significant and robust results of schools' input-output studies, the effect of the socioeconomic factor, which are not controlled by schools, as well as the lack of identifiable statistical properties.

Following McCarty and Yaisawarng (1993), Kirjavainer and Loikkaner (1998), suggest the use of a two-stage model. The first stage uses DEA to calculate the efficiency scores using variables that are controlled by schools. The second stage involves the use of ML estimation of the Tobit regression model, which has well known and desirable statistical properties. The Tobit model provides efficiency measures based on variables that are not included in the DEA and are outside of the decision-making power of schools. Detailed discussion of this approach will be provided in Chapter IV.

Overall, this study suggests that inefficiencies in the education system exist. Modeling these inefficiencies requires careful consideration about the objectives of the research and availability and reliability of the data. This research continues to investigate the robustness of the results from two different estimation methods. In Chapter V the stochastic frontier regression results are compared with data envelopment results. Based on Chapter V, it is concluded that because of the existence of multiple output and the nature of the data, the Data Envelopment Analysis estimator may be a more appropriate technique. Chapter VI explores this technique on a new specification of the model.

CHAPTER III

DATA AND DATA DESCRIPTIONS

3.1 Panel Data

In recent years, the use of panel data in studying school efficiency, as well as agricultural efficiency, has gained popularity (e.g., Adkins and Moomaw, 1997). The use of panel data allows one to include firm-specific effects and time effects, i.e., pooling of cross-section and time-series data.

Panel data may reduce the bias due to the effect of omitted variables on the estimated relationship (Abdul Rahman, 1996). The principle advantage of panel data is that, for each school district, the data are observed over several periods of time. In general, the sample data are represented by observations on N cross-sectional units over T time periods hence the total sample size available to the researcher is potentially NxT.

3.2 Source of Data

The data for this 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. In order to maintain a balanced panel in this study, *dependent schools* which do not offer grades K-12 as well as those districts with incomplete data have been eliminated from the sample. The resulting sample includes observations on 366 school districts.

In order to preserve the consistency of the data for the years of interest, careful attention to the reliability of different measures is vital to the study. Through consultation with the staff of the Office of Accountability and close examination of the data themselves, the reliability of some of the measures is deemed questionable. Some data are logically inconsistent with their definition or with other data and some represent the subjectivity of individual respondents. In addition, the size and nature of the districts could result in different allocation of expenditures for different classifications.

3.3 Description of Data

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 as measures of output are imperfect. However, performance on tests are 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 measure of education. Here ITBS for grades 3 (IT3) and 7 (IT7) and CRT for grades 8 (CRT 8) and 11 (CRT 11) are used as measures of output.

Other measures descriptions are:

ADM: district enrollment is measured as the average daily membership in the school district rounded to the nearest whole number and is calculated by dividing the total days of membership throughout the year by the number of days taught.

I$: instructional expenditures per student is the total expenditure for instruction in each district divided by ADM.

O$: other expenditures per student is the total expenditure for administration and other related operations in each district divided by ADM.

SALARY: salary is the average teacher salary which is the gross salaries and fringe benefits of the district divided by the number of full-time equivalent (FTE) teachers for the school year.

DEG: percentage of teaching staff with an advanced degree is the number of FTE teaching staff with a master's degree or higher divided by the total number of FTE teachers.

YRSEXP: years of experience is the total number of years of experience in the district divided by the number of FTE teachers.

LUNCH: percentage of Oklahoma students eligible for federally funded or reduced payment lunch in school.

This is calculated by one time count of eligible students at the beginning of the school year, divided by ADM. In some districts, because of students' migration and /or dropout, ADM decreases while the count of eligible students for reduced lunch does not change. Thus LUNCH increases to more than 100 percent. In this study, for such districts, LUNCH is assumed to be 100 percent.

MIN: percentage of *noncaucasian* students in the district.

All the variables measured in dollars have been deflated by the Gross Domestic Product (GDP) deflator. Summary statistics of the variables for cross section and panel can be found in Tables 3.1 through 3. 4.

Table 3.1
Summary Statistics for Variables Used in Production Analysis of Oklahoma Schools (1996-1999)

	Mean	Maximum	Minimum	Std. Dev.
IT3	62.2377	93	26	9.949506
IT7	55.6949	85	30	8.192466
CRT8	74.18628	98.6	36.66667	9.39291
CRT11	69.06264	94.25	29	9.80891
O	1992.386	5468.443	1143.357	463.7583
I	2859.53	5767.599	1961.999	498.4934
ADM	1555.872	41471.46	143.91	3767.604
SALARY	29615	35334.72	26608	1190.961
DEG	32.49643	80.62	3.69	13.33091
YRSEXP	15.30133	30.6667	5.830112	4.562163
LUNCH	51.18705	100	4.477241	15.93732
MIN	27.36814	100	0	16.58325
STR	15.95792	21.97037	8.162783	2.200199

IT3: ITBS for grade 3 (composite scores)
IT7: ITBS for grade 7 (composite scores)
CRT8: CRT for grade 8 (average scores)
CRT11: CRT for grade 11 (average scores)
I: Instructional expenditure ($)
O: Noninstructional expenditure ($)
ADM: Average daily membership (number of students)
SALARY: Average salary per full-time equivalent teacher ($)
DEG: Percentage of teachers with advanced degree
YRSEXP: Average experience of teachers (year)
LUNCH: Percentage of students eligible for reduced payment lunch
MIN: Percentage of minority student
STR: Student/Teacher ratio

Table 3.2
Summary Statistics for Variables Used in Production Analysis of Oklahoma Schools (1996-1997)

	Mean	Maximum	Minimum	Std. Dev.
IT3	61.9235	92	26	9.75113
IT7	55.53	79	32	8.032327
CRT8	78.98361	98	52	8.051317
CRT11	69.83811	94.25	33	10.02964
O	1922.02	5468.443	1143.357	462.925
I	2782.808	5052.475	1961.999	456.5806
ADM	1547.944	41196.32	157.27	3771.071
SALARY	30463.23	35334.72	27820.5	992.2415
DEG	33.42098	78.47	4.89	13.58301
YRSEXP	20.50461	30.6667	8.5	3.644732
LUNCH	49.29037	100	4.477241	15.69812
MIN	26.98841	100	0	16.80441
STR	15.93757	21.97037	8.76589	2.248703

IT3: ITBS for grade 3 (composite scores)
IT7: ITBS for grade 7 (composite scores)
CRT8: CRT for grade 8 (average scores)
CRT11: CRT for grade 11 (average scores)
I: Instructional expenditure ($)
O: Noninstructional expenditure ($)
ADM: Average daily membership (number of students)
SALARY: Average salary per full-time equivalent teacher ($)
DEG: Percentage of teachers with advanced degree
YRSEXP: Average experience of teachers (year)
LUNCH: Percentage of students eligible for reduced payment lunch
MIN: Percentage of minority student
STR: Student/Teacher ratio

Table 3.3
Summary Statistics for Variables Used in Production Analysis of Oklahoma Schools (1997-1998)

	Mean	Maximum	Minimum	Std. Dev.
IT3	62.36066	91	35	10.07294
IT7	55.20	75	30	8.343586
CRT8	73.50738	98.6	47	8.881277
CRT11	71.66448	93.6	29	9.652471
O	1998.28	5093.221	1202.86	455.733
I	2857.977	5767.599	2033.183	488.0384
ADM	1556.476	41309.39	156.41	3769.034
SALARY	29153.09	33425.14	26608	1058.236
DEG	32.87582	78.57	3.97	13.27028
YRSEXP	12.67066	19.21429	5.830112	2.067773
LUNCH	51.64396	84.21338	4.668031	15.95561
MIN	27.17257	99.53	0	16.42288
STR	15.95705	21.44213	8.612003	2.229057

IT3: ITBS for grade 3 (composite scores)
IT7: ITBS for grade 7 (composite scores)
CRT8: CRT for grade 8 (average scores)
CRT11: CRT for grade 11 (average scores)
I: Instructional expenditure ($)
O: Noninstructional expenditure ($)
ADM: Average daily membership (number of students)
SALARY: Average salary per full-time equivalent teacher ($)
DEG: Percentage of teachers with advanced degree
YRSEXP: Average experience of teachers (year)
LUNCH: Percentage of students eligible for reduced payment lunch
MIN: Percentage of minority student
STR: Student/Teacher ratio

Table 3.4
Summary Statistics for Variables Used in Production Analysis of Oklahoma Schools (1998-1999)

	Mean	Maximum	Minimum	Std. Dev.
IT3	62.42896	93	34	10.04096
IT7	56.34973	85	31	8.17836
CRT8	70.06785	93	36.66667	8.995299
CRT11	65.68534	93	40.83333	8.750936
O	2056.858	4878.875	1292.437	463.9328
I	2937.804	5702.062	2099.873	536.7986
ADM	1563.197	41471.46	143.91	3773.009
SALARY	29228.68	32800.18	26781.6	1036.082
DEG	31.19249	80.62	3.69	13.06936
YRSEXP	12.72873	17.88462	5.899977	2.066434
LUNCH	52.62682	90.9613	4.543612	16.01495
MIN	27.94344	99.82	0	16.55008
STR	15.97912	20.79424	8.162783	2.126718

IT3: ITBS for grade 3 (composite scores)
IT7: ITBS for grade 7 (composite scores)
CRT8: CRT for grade 8 (average scores)
CRT11: CRT for grade 11 (average scores)
I: Instructional expenditure ($)
O: Noninstructional expenditure ($)
ADM: Average daily membership (number of students)
SALARY: Average salary per full-time equivalent teacher ($)
DEG: Percentage of teachers with advanced degree
YRSEXP: Average experience of teachers (year)
LUNCH: Percentage of students eligible for reduced payment lunch
MIN: Percentage of minority student
STR: Student/Teacher ratio

CHAPTER IV

THE MODEL SPECIFICATION

4.1 Choosing the Appropriate Function to Estimate Oklahoma Schools' Performance

In economic analysis firms are viewed as either profit maximizers or cost minimizers. A publicly funded educational institution may not have the same objective function as either a cost minimizing or a profit maximizing industrial firm. Hanushek (1986) suggests that, in the production function framework, school administrators may be better characterized as output maximizers. This behavioral assumption on the part of district administration has important implications for the choice of functional form and its econometric estimator.

This study proposes the estimation of a stochastic frontier production function associated with Oklahoma school districts' production process to investigate the efficiency of the school districts. The production function approach to efficiency study is appropriate for several reasons. First, the only necessary assumption is that Oklahoma school districts attempt to maximize output (maximizing test scores). Therefore, in view of the discussion in Chapter II, each district's objective is to operate at a point on its production function. Second, the available data do not contain information on input prices or output prices; hence, cost functions cannot be estimated. Finally, the production of public education is not a simple input-output relationship. Public education is affected by environmental variables that are outside the control of schools' administrators. This implies the existence of a random error as well as the deterministic error (stochastic vs. deterministic frontier).

Battese and Coelli (1993, p. 1) state that:

> the stochastic frontier production function postulates the existence of technical inefficiencies of production of firms involved in producing a particular output. For a given combination of input levels, it is assumed that the realized production of a firm is bounded above by the sum of a parametric function of known inputs, involving unknown parameters and a random error associated with the measurement error of the level of production or other factors such as the effects of weather, strikes, damaged product, etc. The greater the amount by which the realized production falls short of this stochastic frontier production, the greater the level of technical inefficiency.

A major advantage of stochastic frontier estimation is that it allows for the measurement of inefficiency, but the real advantage is that it allows for these inefficiencies to be firm-specific.

4.2 Econometric Specification

The econometric estimation of the production function requires the specification of a suitable functional form (e.g., Cobb-Douglas, translog, etc.) for the production function. The translog functional form is chosen because it is relatively well behaved in panel data studies and, although it is more complex than the Cobb-Douglas functional form, with the help of available computer software it is relatively easy to compute.

4.2.1 Two-equation Stochastic Frontier Production Model

Huang and Liu (1992) consider a two-equation stochastic frontier production model involving non-negative inefficiency effects which are a linear function of firm characteristics and a random error that can be positive or negative. The random error is assumed to have a truncated normal distribution. The truncation point is a function of firm characteristics. Huang and Liu use the maximum likelihood method (MLE) to estimate the parameters of the model using cross sectional data.

The advantage of a two-equation MLE, as opposed to a single-equation MLE, is that the inefficiency variables and the explanatory variables of the stochastic frontier can be estimated simultaneously, i.e., allowing interaction between firm-specific variables and the right-hand side variables of the frontier function. Allowing this interaction emphasizes the possibility of non-neutral shifting of *average response* functions, in which case OLS is not capable of determining the shape of the boundary function, which weakens its analytical ability even further. Battese and Coelli (1993) extended Huang and Liu's (1992) model to panel data, which allows the inclusion of both firm-specific effects and time effects in the inefficiency model, and applied it to farm-level data from an Indian village.

4.2.2 Battese and Coelli's Model of the Inefficiency Frontier for Panel Data (1993)

Consider the following stochastic frontier production model for panel data.

$$Y_{it} = exp(X_{it}\beta + V_{it} - U_{it}) \tag{4.1}$$

where Y_{it} denotes the production for the t-th observation (t = 1,2, . . ., T) for the i-th firm (i = 1,2, . . . , N); X_{it} is a (1xk) vector of inputs of production associated with the i-th firm at the t-th period of observation; β is a (kx1) vector of unknown parameters to be estimated; the V_{it}'s are assumed to be *iid* $N(0, \sigma_v^2)$ random errors, independently distributed of the U_{it}'s which are non-negative random variables, associated with technical inefficiency of production; the U_{it}'s are assumed to be independently distributed, such that U_{it} is obtained by truncation (at zero) of the normal distribution with mean, $Z_{it}\delta$, and variance σ^2; Z_{it} is a (1xm) vector of firm-specific variables which may vary over time; and δ is an (mx1) vector of unknown coefficients of the firm-specific inefficiency variables.

The authors assume that at least one of the N firms has observations available for the T time periods however, not all firms are required to have observations for all periods.

The inefficiency effect, U_{it} in the stochastic frontier model (1) is specified as:

$$U_{it} = Z_{it}\delta + W_{it} \tag{4.2}$$

where:

Z_{it}'s are a set of explanatory variables, which include any variable that explains the extent to which the production observations are below the stochastic frontier production value, $exp\ (X_{it}\beta + V_{it})$.

W_{it} is a random variable with $N(0, \sigma^2)$ truncated at $-Z_{it}\delta$, i.e., $W_{it} \geq -Z_{it}\delta$. These assumptions are consistent with the assumption of U_{it} in equation (4.1). The parameters of both models (4.1) and (4.2) are estimated by the Maximum Likelihood method. Derivation of the likelihood function and its partial derivatives with respect to the parameters of the model are presented in their Appendix.

Adkins and Moomaw (1997) employed Battese and Coelli's (1993) MLE to estimate a translog production function of Oklahoma school districts performance. In the first, the inputs to production are, instructional expenditure per student, ln(I/S), and other expenditures per student, ln(O/S). In the second, they introduce environmental factors LUNCH, MIN, and LEP in the production function where:

LUNCH: is the percent of students eligible for subsidized lunch as a measure of poverty;

MIN: is the percent of minority student; and

LEP: is the percent of students with limited English proficiency.

These factors enter not as production inputs, but as control variables. The sample they consider is the 1990-91 through 1994-95 academic years. The parameter estimates are computed under the assumption that both components of the random error are homoscedastic. However, if the error terms are heteroscedastic then the MLE is inconsistent and their conclusions may be incorrect.

"Heteroscedasticity is one specification error that researchers can reasonably expect to encounter in the estimation of stochastic frontier models" (Caudill et al., 1995, p. 106). The problem of heteroscedasticity in frontier models has been addressed in several studies (Reifschneider and Stevenson, 1991, Yuengert, 1993, and others). However, they all took different approaches to the incorporation of heteroscedasticity into the frontier models.

These studies suggest that in frontier models, residuals are more sensitive to specification errors than the average function models, and consequently, sensitivity to specification errors can be passed on to the inefficiency measures. Heteroscedasticity in the estimation of the average function does not affect the point estimation properties of the least squares estimator; it remains unbiased. However, least squares is no longer efficient and the validity of subsequent hypothesis tests is questionable. In a frontier model, the locus of the frontier is altered when the dispersion increases, indicating that the problem of heteroscedasticity is far more serious in these models.

The first study of this type, conducted by Reifschneider and Stevenson (1991), incorporated the heteroscedasticity into the composite error term, ε, by allowing the mean of the deterministic part of the error term, the *one-sided error term,* to change. Caudill, Ford, and Gropper (1995) account for heteroscedasticity in the *one-sided error term* of their banking model. Their findings suggest that heteroscedasticity leads to biased parameter estimates that overstate the intercept and understate the slope coefficients for a production frontier (the opposite for a cost frontier) when the model is estimated using the maximum likelihood method. Not surprisingly, the inefficiency measures are affected by heteroscedasticity as well, since most inefficiency measures are based on residuals.

This study was extended by Hardi (1999), who employed cross-section data and the assumption of heteroscedasticity in both random terms. Hardi's results confirm Caudill et al.'s (1995) claim that firm-specific inefficiency measures are highly sensitive to the possible existence of heteroscedasticity.

Following Adkins and Moomaw (1997), the basic translog model for Oklahoma schools for the periods of 1996-1999 considered in this dissertation is:

$$\begin{aligned} ln\ Score_{it} &= \beta_0 + ln(I_{it})\beta_1 + ln(O_{it})\beta_2 + [ln(I_{it})]^2\beta_3 \\ &+ [ln(O_{it})]^2\beta_4 + ln(I_{it})ln(O_{it})\beta_5 + (V_{it} - U_{it}) \end{aligned} \tag{4.3}$$

where output:

Score: is a measure of an average districts' performance on one of the several standardized tests, which are:

IT3	ITBS scores for grade 3
IT7	ITBS scores for grade 7
CRT8	Average CRT scores for grade 8
CRT11	Average CRT scores for grade 11

Inputs:

I: Instructional expenditure per student

O: Noninstructional expenses per student, i.e., administrative and any other expenses that are not directly used for instructional purposes.

Incorporating environmental variables in equation (4.3), the model can be rewritten as:

$$\begin{aligned} ln\ Score_{it} = {} & \beta_0 + MIN_{it}\beta_1 + LUNCH_{it}\beta_2 + ln(I_{it})\beta_3 + ln(O_{it})\beta_4 + [ln(I_{it})]^2\beta_5 \\ & + [ln(O_{it})]^2\beta_6 + ln(I_{it})ln(O_{it})\beta_7 + (V_{it} - U_{it}) \end{aligned} \tag{4.4}$$

Equation (4.4) is essentially Adkins and Moomaw's (1997) model except for some changes in the way some of the data are measured. The adjustments stem from changes in computation methods of some variables and/or elimination of some others by the Oklahoma Office of Accountability. In Adkins and Moomaw's study for the periods 1990-1991 through 1994-1995, the standardized test for 9th and 11th grades was TAP and data on LEP was available. For the period of this study (1996-1997 through 1998-1999), LEP is eliminated. The standardized test for 9th grade is replaced by 8th grade tests. For both 8th and 11th grade, TAP is replaced by CRT.

Following Adkins and Moomaw (1997) environmental and non-environmental variables in (4.4) are a part of the production function; however, the school administrators have no control over the environmental variables and, a change in these variables shifts the frontier.

The model of inefficiency includes other exogenous variables, which may measure differences in input quality, they are:

SALARY: is the average teacher salary;

YRSEXP: is the average years of experience for teachers; and

DEG: is the proportion of teachers with an advanced degree.

The model also includes variables that measure the various quantity adjustments available to district administrators, i.e., (ADM) total enrollment, enrollment squared, and student/teacher ratio (STR). The inefficiency model is written as:

$$TE_{it} = \delta_0 + \delta_1 SALARY_{it} + \delta_2 YRSEXP_{it} + \delta_3 DEG_{it} + \delta_4 ADM_{it} + \delta_5 ADM_{it}^2 + \delta_6 STR_{it} + e_{it} \quad (4.5)$$

where TE is the technical efficiency.

4.2.3 Incorporating Heteroscedasticity in the Present Study

According to Caudill et al. (1995, p. 106), "in many econometrics textbooks readers are advised to expect heteroscedasticity when the observations are of different size." Hardi (1999) suggests that the vector of exogenous variables that determine heteroscedasticity are related generally to characteristics of firm size (Hardi, 1999 p. 360). Given the differences in sizes of the schools in this application, it is perhaps reasonable to allow for heterogeneity in the variances and covariances, and then test for constant variance instead of imposing it a priori (Kumbhakar, 1997).

In this study, heteroscedasticity is incorporated in the stochastic part of the error term, i.e., V_{it}. To incorporate multiplicative heteroscedasticity (Greene, 1990), consider:

$$\sigma_{vi}^2 = \exp\{S_{it}\alpha\} \qquad \sigma_{vi} > 0 \tag{4.6}$$

where:

S_{it} is a vector of exogenous variables that determine heteroscedasticity.

α is a vector of the parameters to be estimated.

The multiplicative functional form is easily constrained to be homoscedastic to make a likelihood ratio test possible.

In Battese and Coelli's (1993 Appendix, equation A-12) model, the logarithm of the likelihood function is written as:

$$\begin{aligned} L^*(\theta^*; y) = &-\frac{1}{2}\left(\sum_{i=1}^{N} T_i\right)\left\{\ln 2\pi + \ln\left(\sigma_U^2 + \sigma_V^2\right)\right\} \\ &-\frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\left(y_{it} - x_{it}\beta + z_{it}\delta\right)^2 / \left(\sigma_V^2 + \sigma_U^2\right)\right] \\ &-\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\ln\phi(d_{it}) - \ln\Phi(d_{it}^*)\right] \end{aligned} \tag{4.7}$$

where $\theta = (\beta', \delta', \sigma_v^2, \sigma_u^2)$.

$$d_{it} = Z_{it}\delta / \sigma_u$$

$$d_{it}^* = \frac{u_{it}^*}{\sigma_*}$$

$$u_{it}^* = \left[\sigma_v^2 z_{it}\delta - \sigma_u^2 (y_{it} - x_{it}\beta)\right] / \left(\sigma_v^2 + \sigma_u^2\right)$$

$$\sigma_* = \left(\sigma_u^2 \sigma_v^2\right)^{1/2} / \left(\sigma_u^2 + \sigma_v^2\right)^{1/2}$$

Incorporating equation (4.6) in equation (4.7):

$$L^*(\theta^*;y) = -\frac{1}{2}\left(\sum_{i=1}^{N} T_i\right)\left\{\ln 2\pi + \ln\left(\sigma_u^2 + e^{S_{it}\alpha}\right)\right\} \tag{4.8}$$

$$-\frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\left(y_{it} - x_{it}\beta + z_{it}\delta\right)^2 / \left(e^{S_{it}\alpha} + \sigma_u^2\right)\right]$$

$$-\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\ln\Phi(d_{it}) - \ln\Phi(d_{it}^*)\right]$$

where: d_{it} and d_{it}^* are as before

$$u_{it}^* = \left[e^{S_{it}\alpha} z_{it}\delta - \sigma_u^2\left(y_{it} - x_{it}\beta\right)\right] \Big/ \left(e^{S_{it}\alpha} + \sigma_u^2\right).$$

$$\sigma_* = \left(\frac{e^{S_{it}\alpha}\sigma_u^2}{e^{S_{it}\alpha} + \sigma_u^2}\right)^{1/2}$$

then: $$d_{it}^* = \left\{\left[e^{S_{it}\alpha} Z_{it}\delta - \sigma_u^2\left(y_{it} - x_{it}\beta\right)\right] \Big/ \left(e^{S_{it}\alpha} + \sigma_u^2\right)\right\}$$

$$\bullet\left\{\left(e^{S_{it}\alpha} + \sigma_u^2\right)^{1/2} \Big/ \left(e^{S_{it}\alpha}\sigma_u^2\right)^{1/2}\right\}$$

and $\theta = (\beta', \delta', \alpha')$.

The partial derivatives of equation (4.8) with respect to the parameters, β, δ, and α are given by

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^{N}\sum_{t=1}^{T_i}\left\{\frac{\left(y_{it} - x_{it}\beta + Z_{it}\delta\right)}{e^{S_{it}\alpha} + \sigma_u^2} + \frac{\phi(d_{it}^*)}{\Phi(d_{it}^*)}\cdot\left(\frac{e^{S_{it}\alpha}}{\sigma_*\left(e^{S_{it}\alpha} + \sigma_u^2\right)}\right)\right\}x'_{it} \tag{4.9A}$$

where $\phi(.)$ represents the density function for the standard normal random variable.

$$\frac{\partial L}{\partial \delta}=\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left\{\frac{\left(y_{it}-x_{it}\beta+z_{it}\delta\right)}{e^{S_{it}\alpha}+\sigma_u^2}+\left[\frac{\phi\left(d_{it}\right)}{\Phi\left(d_{it}\right)}\cdot\frac{1}{\sigma_u}-\frac{\phi\left(d_{it}^*\right)}{\Phi\left(d_{it}^*\right)}\cdot\left(\frac{e^{S_{it}\alpha}}{\sigma_*\left(e^{S_{it}\alpha}+\sigma_u^2\right)}\right)\right]\right\}z'_{it} \quad (4.9B)$$

$$\begin{aligned}\frac{\partial L^*}{\partial \alpha}=&-1/2\left\{\sum_{i=1}^{N}T_i\left[\frac{e^{S_{it}\alpha}}{\sigma_u^2+e^{S_{it}\alpha}}\right]+\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\frac{e^{S_{it}\alpha}\left(y_{it}-x_{it}\beta+z_{it}\delta\right)^2}{\left(e^{S_{it}\alpha}+\sigma_u^2\right)^2}\right]\right\}s'_{it}\\&+\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left\{\frac{\phi\left(d_{it}^*\right)}{\Phi\left(d_{it}^*\right)}\left\{\left[\begin{array}{l}1/2\left[\dfrac{e^{S_{it}\alpha}z_{it}\delta-\sigma_u^2\left(y_{it}-x_{it}\beta\right)}{e^{S_{it}\alpha}+\sigma_u^2}\right]\cdot\left[\dfrac{\sigma_u^2+e^{S_{it}\alpha}}{e^{S_{it}\alpha}\cdot\sigma_u^2}\right]^{-1/2}\cdot\\\left[\dfrac{e^{S_{it}\alpha}\left(e^{S_{it}\alpha}\cdot\sigma_u^2\right)-e^{S_{it}\alpha}\cdot\sigma_u^2\left(e^{S_{it}\alpha}+\sigma_u^2\right)}{\left(e^{S_{it}\alpha}\cdot\sigma_u^2\right)^2}\right]\\+\left[\dfrac{\sigma_u^2+e^{S_{it}\alpha}}{e^{S_{it}\alpha}\cdot\sigma_u^2}\right]^{1/2}\cdot\\\left[\dfrac{\left(e^{S_{it}\alpha}+\sigma_u^2\right)z_{it}\delta e^{S_{it}\alpha}-\left[e^{S_{it}\alpha}z_{it}\delta-\sigma_u^2\left(y_{it}-x_{it}\beta\right)\right]e^{S_{it}\alpha}}{\left(e^{S_{it}\alpha}+\sigma_u^2\right)^2}\right]\end{array}\right]\right\}\right\}s'_{it}\end{aligned} \quad (4.9C)$$

This result could be used to generalize the Battese and Coelli estimator when the two-sided random errors, v_{it}, are heteroscedastic.

4.3 DEA and Second Stage Tobit Specification

In this section the specification of the non-parametric model, Data Envelopment Analysis, in the first stage and Tobit regression in the second stage are discussed.

4.3.1 DEA Model

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

To measure technical inefficiency, output-oriented DEA uses the same inputs and outputs as the stochastic frontier model. Therefore, for the production frontier in equation (4.4), the output-oriented measure of technical inefficiency is the solution to the following multiplier form of the DEA linear programming problem. Following Coelli (1998):

$$\max_{u,\, v} (u'y_i),$$

$$\begin{aligned} S.T \qquad & v'x_i = 1 \\ & u'y_j - v'x_j \leq 0 \qquad\qquad j = i, \ldots, N \\ & u,\ v \geq 0 \end{aligned}$$

where:

x_i is a column vector of inputs (MIN, LUNCH, I, O) for the ith school district

y_i is a column vector of outputs (IT3, IT7, CRT8, CRT11) for the ith school district

v is a K x 1 vector of input weights, $K = 4$, that maximizes efficiency

u is a M x 1 vector of output weights, $M = 4$, that maximizes efficiency

$u,\ v \geq 0$

4.3.2 Tobit Model

To assess the effects of variables not included in the first stage on technical efficiency, McCarty et al. (1993) suggest using efficiencies generated by DEA as dependent variables in a Tobit regression:

$$\begin{aligned} Y_{it}^{T} = \beta_0^T + \beta_1^T \mathrm{YRSEXP}_{it} + \beta_2^T \mathrm{DEG}_{it} + \beta_3^T \mathrm{SALARY}_{it} \\ + \beta_4^T \mathrm{ADM}_{it} + \beta_5^T \mathrm{ADM}_{it}^2 + \beta_6^T \mathrm{STR}_{it} + e_{it} \end{aligned} \qquad (4.10)$$

where the T superscript denotes Tobit and Y_{it}^{T} is the DEA efficiency estimates. Efficiency estimates from the first state are between 0 and 1, hence data is censored and thus Tobit regression is the appropriate method of estimation for equation 4.10. The explanatory variables in equation (4.10) are the variables of technical efficiency equation of the stochastic frontier model (equation 4.5). The possibility of existence of heteroscedasticty in this stage should be considered and, if in fact it exists, incorporated into the model to have efficient parameter estimates.

CHAPTER V

STOCHASTIC PRODUCTION FRONTIER

5.1 OLS Estimation of the Model

The first goal in specification of the stochastic frontier is to determine to what extent the data can be pooled. If there are significant structural changes in the school districts from year to year then aggregating them and estimating common coefficients for production function may be misleading. Also, an initial analysis for the existence of heteroscedasticity is of interest. Since test scores are measured as averages and given the wide variation in the number of pupils in each district, there is reason to suspect that the data are heteroscedastic. These issues of specification are investigated using least squares.

The first concern with the use of panel data is whether there has been any major structural changes that have occurred over time.

H_0: No structural changes among cross-sections

H_1: H_0 not true

To test this hypothesis, a Chow test for each grade was performed. To compute the test statistic:

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(n-K)} \sim F_{(J,n-K)} \qquad (5.1)$$

where:

SSE_R is the sum of squares of error from the pooled regression.

SSE_U is the unrestricted sum of squares of error for each cross-section ($SSE_{96\text{-}97}$ + $SSE_{97\text{-}98}$ + $SSE_{98\text{-}99}$).

J is the number of restrictions, (2J = 16).

n is the panel number of observations, 1098.

K is the number of explanatory variables for each cross-section (3K = 3(8) = 24).

OLS parameter estimates of the model, under the null hypothesis, for each cross-section and for the panel data are obtained and presented in Tables 5.1 – 5.4. MIN and LUNCH have the expected sign and are consistent with Adkins and Moomaw's (1997) results.

Table 5.1
OLS Estimates of the Parameters of Production Function Year (1996-1997)
Dependent Variable: ln (test score)

	Grade 3		Grade 7		Grade 8		Grade 11	
Variable	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**
Constant	-6.0836	-0.3499	-6.0521	-.4275	14.2761	1.4885	-18.5929	-1.2826
MIN	-0.0576	-0.9556	-.1122	-2.2857	-0.1028	-3.0922	-0.1357	-2.7008
LUNCH	-0.3619	-4.3949	-.4787	-7.1387	-0.2907	-6.3999	-0.4814	-7.0102
ln(I)	2.4008	0.4924	.5963	.1502	-4.1302	-1.5354	3.6982	0.9096
ln(O)	0.0008	0.0004	1.7569	.9578	1.5368	1.2366	2.0134	1.0719
$[\ln(I)]^2$	-0.6153	-0.7501	.3155	.4723	0.817	1.805	-0.5184	-0.7577
$[\ln(0)]^2$	-0.3444	-0.8748	.1883	.5875	0.1061	0.4883	-0.3203	-0.9758
ln(I)ln(O)	0.3427	0.7442	-.3841	-1.0222	-0.289	-1.1375	0.0671	0.1747
Sum of Squares Errors	9.20211		6.102255		2.800867		6.398876	

Table 5.2
OLS Estimates of the Parameters of Production Function Year (1997-1998)
Dependent Variable: ln (test score)

	Grade 3		Grade 7		Grade 8		Grade 11	
Variable	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**
Constant	-33.6618	-2.1248	-9.9107	-.7431	-13.3632	-1.2659	2.1181	0.1744
MIN	-0.1009	-1.5975	-.1356	-2.5487	-0.1421	-3.3734	-0.2273	-4.6918
LUNCH	-0.0273	-3.3657	-.5128	-7.5032	-0.3694	-6.8288	-0.3517	-5.6521
ln(I)	7.7698	1.7922	2.1409	.5866	3.3849	1.1718	0.5513	0.1657
Ln(O)	1.6443	0.6654	1.2481	.5999	0.9465	0.5748	-0.0691	-0.0364
$[\ln(I)]^2$	-0.6449	-0.8428	-.2105	-.3268	0.0901	0.1768	-0.1541	-0.2628
$[\ln(0)]^2$	0.1262	-0.3247	-.1035	-.3162	0.4325	1.6694	-0.0874	-0.2935
ln(I)ln(O)	-0.3239	-0.729	-.0437	-.1169	-0.5223	-1.7644	0.0981	0.2882
Sum of Squares Errors	9.45593		6.701311		4.198008		5.554300	

Table 5.3
OLS Estimates of the Parameters of Production Function Year (1998-1999)
Dependent Variable: ln (test score)

	Grade 3		Grade 7		Grade 8		Grade 11	
Variable	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**
Constant	-39.1746	-2.7907	-8.3154	-.7.39	-13.6577	-1.3133	2.5475	0.2395
MIN	-0.0233	-0.3755	-.1461	-2.7989	-0.1601	-3.4824	-0.0766	-1.6297
LUNCH	-0.3847	-4.9276	-.4179	-6.3606	-0.3593	-6.2129	-0.4221	-7.1365
ln(I)	8.0683	2.3297	.1823	.0625	2.951	1.1502	1.8829	0.7177
ln(O)	2.6435	1.0815	2.8523	1.3866	1.4985	0.8275	-1.6226	-0.8763
$[\ln (I)]^2$	0.1049	0.1696	.3632	.6977	0.2309	0.5039	-0.4784	-1.0211
$[\ln (0)]^2$	0.8453	2.2145	.0424	.1322	0.4447	1.5728	-0.0659	-0.2282
ln(I)ln(O)	-1.1331	-2.8031	-.3846	-1.1305	-0.6101	-2.0374	0.2691	0.8788
Sum of Squares Errors	8.97321		6.35511		4.924563		5.149717	

Table 5.4
OLS Estimates of the Parameters of Production Function (Panel)
Dependent Variable: ln (test score)

	Grade 3		Grade 7		Grade 8		Grade 11	
Variable	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**
Constant	-26.3702	-3.0592	-9.3923	-1.3123	-3.6335	-0.5817	-2.5406	-0.3619
MIN	-0.0592	-1.6632	-0.1283	-4.3416	-0.1307	-5.0655	-0.1520	-5.2406
LUNCH	-0.3382	-7.3635	-0.4734	-12.4136	-0.3476	-10.4453	-0.4091	-10.9336
ln(I)	5.8497	2.5399	1.1379	0.5950	1.6940	1.0150	1.5517	0.8272
ln(O)	1.6906	1.2363	2.1120	1.9012	0.2211	0.2280	0.0887	0.0814
$[\ln (I)]^2$	-0.1633	-0.8024	0.0814	0.4817	0.1154	0.7824	-0.1699	-1.0249
$[\ln (0)]^2$	0.1044	0.9484	0.0252	0.2759	0.2227	2.7905	-0.0869	-0.9694
ln(I)ln(O)	-0.4035	-1.6214	-0.2982	-1.4431	-0.4495	-2.4927	0.1617	0.7979
Sum of Squares Errors	28.04635		19.33482		14.72695		18.60676	

The results of the Chow tests suggest that, for grades 3 and 7, no structural changes are detected over the years under study. However, the hypothesis for grades 8 and 11, is rejected at the 5% level. The results are presented in Table 5.5.

Table 5.5
Chow Test of Structural Changes among Cross-Sections in the OLS Model

Grade	F-statistic df = 15	Critical Value $\alpha = .05$	Decision Rule
3	1.008	1.67	Not Reject H_0
7	.617	1.67	Not Reject H_0
8	15.785	1.67	Reject H_0
11	5.870	1.67	Reject H_0

One possible reason for rejecting the null hypotheses for 8th and 11th grade could be that, CRTs are renormed each year. Another is that the technology used by the school districts for these grades may be different for each year under the study. To allow for structural changes for each year in the panel and to capture the effects of these changes on the parameter estimates, dummy variables for years 1996-1997 and 1997-1998 are introduced into the model. The OLS parameter estimates of this model are presented in Table 5.6.

Table 5.6
OLS Estimates of the Parameters of Production Function with Dummy Variables (Panel)
Dependent Variable: ln (test score)

	Grade 8		Grade 11	
Variable	**Coefficient**	**t-ratio**	**Coefficient**	**t-ratio**
Constant	-8.98612	-1.58407	-4.32687	-0.63780
MIN	-0.12999	-5.55676	-0.14655	-5.23848
LUNCH	-0.35233	-11.67508	-0.41673	-11.54713
ln(I)	1.85033	1.22303	1.55087	0.85718
ln(O)	1.38914	1.57481	0.52009	0.49303
$[\ln (I)]^2$	0.12726	0.95200	-0.16022	-1.00221
$[\ln (0)]^2$	0.16921	2.33677	-0.10531	-1.21612
ln(I)ln(O)	-0.49057	-3.00081	0.14388	0.73593
DUM1	0.12174	15.35997	0.05408	5.70517
DUM2	0.04964	6.35498	0.08468	9.06481
Sum of Squares Errors	12.08280		17.26241	

The results of Table 5.6 suggest that the effect of structural changes on the test scores are statistically significant. To test this hypothesis a Chow test based on the regression used for Table 5.6 which includes the two dummy variables; this regression becomes the restricted model against which the unrestricted is compared. The results are presented in Table 5.7.

Table 5.7
Chow Test of Structural Changes among Cross-Sections in the OLS Model with Dummy Variables

Grade	F-statistic df = 20	Critical Value $\alpha = .05$	Decision Rule
8	.708	1.57	Not Reject H_0
11	.501	1.57	Not Reject H_0

These results suggest that whatever structural changes occurred for grades 8 and 11, are adequately captured by the inclusion of yearly dummies. Furthermore, the coefficient estimates suggest downward secular trends in 8^{th} and 11th grade test scores, holding expenditures and the observed student characteristics constant. This may indicate an actual decline in technology employed by the school districts or merely indicate that the exams became more difficult over time.

The elasticities of output with respect to instructional (I) and noninstructional (O) expenditures based on the parameter estimates of the OLS model, using the panel data, for each grade are presented in Table 5.8. These elasticities are computed using the sample mean values of ln(I) and ln(O).

Table 5.8
Elasticities of Test Scores with Respect to Instructional and Noninstructional Expenditures Evaluated at their Asymptotic Means

	IT3	IT7	CRT8	CRT11
I	.197	.174	.159	.0918
Standard Error	.022	.016	.02	.019
O	.064	.123	.0515	.0683
Standard Error	.032	.0158	.0326	.008

The elasticities of test scores with respect to expenditures are positive and statistically significant for all grades at the 5 percent level. With respect to instructional expenses, the estimated effect is greater in grade 3 than in the other grades considered. However, the elasticities are fairly small even in grade 3. This result is consistent with findings of Adkins and Moomaw (1997). A one percent increase in instructional spending is expected to increase the 3rd grade ITBS scores by almost .2 percent, .17 percent for grade 7, .15 percent for grade 8, and.09 percent for grade 11.

Smaller elasticities with regard to noninstructional expenses, ranging from .05 to.12 suggest that reallocation of noninstructional spending to instructional spending may result in a small improvement of test scores for all grades. This result is consistent with the findings of Adkins and Moomaw (1997) except for grade 7.

To test for the existence of any form of heteroscedasticity, White's heteroscedasticity test on the residuals of panel data for each type of test scores was conducted. The results suggest the existence of heteroscedasticity in this model. The results are presented in Table 5.9.

Table 5.9
Results of White's Heteroscedasticity Test on the Residual of Panel Data

Output	χ^2-statistic	df	P-Value
IT3	86.34	25	.0000
IT7	87.84	25	.0000
CRT8	93.02	41	.0000
CRT11	63.45	41	.0137

Existence of heteroscedasticity in the average function model most likely suggests the presence of this specification error in the frontier model. For the moment this complication will be ignored and estimation of a stochastic frontier will proceed as if the model were homoscedastic. Later in the chapter, heteroscedasticity will be reconsidered using results from DEA and the efficiencies predicted by the two approaches will be compared.

5.2 The Frontier Approach (Stochastic Frontier Regression, SFR)

In the traditional average response function (OLS), school districts are assumed to be fully efficient, i.e., u_i's are not present in the model. This assumption should be tested in order to see whether we need to go beyond OLS and whether a stochastic frontier production is required at all. In order to test this, the maximum likelihood estimates of the parameters in the model were obtained using Frontier 4.1 (Coelli, 1995). The results are presented in Tables 5.10 through 5.13.

As expected, LUNCH and MIN each shift the frontier down for all grades. These results confirm the importance of students' socioeconomic variation on their educational performance and therefore, increasing spending per student in districts with a higher percentage of disadvantaged students could offset their disadvantages. These results are consistent with Adkins and Moomaw (1997).

Table 5.10
Maximum-Likelihood Estimates for Parameters of Stochastic Production Frontier and Inefficiency Models for Oklahoma School Districts (Panel Data Model)
Dependent Variable: ln (IT3)

Variable	Parameter	Coefficient	Standard Error	t-ratio
Stochastic Production Frontier				
Constant	β_0	-25.84899	1.10360	-23.42245
MIN	β_1	-0.09450	0.03377	-2.79807
LUNCH	β_2	-0.22242	0.03864	-5.75591
ln(I)	β_3	4.10110	0.72787	5.63440
ln(O)	β_4	3.35807	0.76602	4.38381
$[\ln (I)]^2$	β_5	0.09941	0.11820	0.84100
$[\ln (0)]^2$	β_6	0.16076	0.10202	1.57571
ln(I)ln(O)	β_7	-0.71907	0.20306	-3.54107
Inefficiency Equation				
Constant	δ_0	1.17627	0.46809	2.51292
*SALARY	δ_1	-0.36277	0.16257	-2.23143
YRSEXP	δ_2	-0.61496	0.14295	-4.30196
DEG	δ_3	0.00857	0.00402	2.13352
*ADM	δ_4	-0.01686	0.00538	-3.13307
*ADM^2	δ_5	0.00004	0.00001	3.20096
STR	δ_6	-0.00977	0.00818	-1.19415
Variance Parameters	$\sigma^2 = \sigma_U^2 + \sigma_V^2$	0.07761	0.01269	6.11962
	$\gamma = \frac{\sigma_U^2}{\sigma_U^2 + \sigma_V^2}$	0.86033	0.02692	31.95549
Log Likelihood Function		519.73200		
LR test of the one-sided error (H_0: $\gamma = 0$)		128.66600		
number of restrictions		8		

*Variable is scaled:
SALARY is SALARY/1,000
ADM is ADM/100
ADM^2 is ADM^2/10,000

Table 5.11
Maximum-Likelihood Estimates for Parameters of Stochastic Production Frontier and Inefficiency Models for Oklahoma School Districts (Panel Data Model)
Dependent Variable: ln (IT7)

Variable	Parameter	Coefficient	Standard Error	t-ratio
Stochastic Production Frontier				
Constant	β_0	-10.4653	1.2132	-8.6264
MIN	β_1	-0.1767	0.0262	-6.7391
LUNCH	β_2	-0.3354	0.0310	-10.8264
ln(I)	β_3	1.1231	0.7180	1.5643
ln(O)	β_4	2.4156	0.7306	3.3065
$[\ln (I)]^2$	β_5	0.1834	0.1054	1.7400
$[\ln (0)]^2$	β_6	0.1155	0.0806	1.4329
ln(I)ln(O)	β_7	-0.5087	0.1689	-3.0122
Inefficiency Equation				
Constant	δ_0	1.2290	0.4487	2.7393
*SALARY	δ_1	-0.4593	0.1604	-2.8644
YRSEXP	δ_2	-0.3958	0.1059	-3.7354
DEG	δ_3	0.0073	0.0037	1.9833
*ADM	δ_4	-0.0161	0.0046	-3.4802
*ADM^2	δ_5	0.0004	0.0001	3.5336
STR	δ_6	0.0019	0.0082	0.2335
Variance Parameters	$\sigma^2 = \sigma_U^2 + \sigma_V^2$	0.0600	0.0073	8.2102
	$\gamma = \frac{\sigma_U^2}{\sigma_U^2 + \sigma_V^2}$	0.9034	0.0167	54.0832
Log Likelihood Function		743.6904		
LR test of the one-sided error (H_0: $\gamma = 0$)		168.1766		
number of restrictions		8		

*Variable is scaled:
SALARY is SALARY/1,000
ADM is ADM/100
ADM^2 is ADM^2/1,000,000

Table 5.12
Maximum-Likelihood Estimates for Parameters of Stochastic Production Frontier and Inefficiency Models for Oklahoma School Districts (Panel Data Model)
Dependent Variable: ln (IT8)

Variable	Parameter	Coefficient	Standard Error	t-ratio
Stochastic Production Frontier				
Constant	β_0	-10.31359	1.17694	-8.76306
MIN	β_1	-0.13522	0.02090	-6.47153
LUNCH	β_2	-0.27192	0.02439	-11.14877
ln(I)	β_3	1.27378	0.73915	1.72332
ln(O)	β_4	2.33484	0.73806	3.16346
$[\ln (I)]^2$	β_5	0.27427	0.09465	2.89777
$[\ln (0)]^2$	β_6	0.22877	0.06653	3.43876
ln(I)ln(O)	β_7	-0.72149	0.13854	-5.20790
DUM1	β_8	0.09548	0.00741	12.88991
DUM2	β_9	0.04530	0.00732	6.18906
Inefficiency Equation				
Constant	δ_0	0.11763	0.34224	0.34370
*SALARY	δ_1	0.09261	0.12188	0.75986
YRSEXP	δ_2	-0.46873	0.13903	-3.37143
DEG	δ_3	-0.03362	0.01144	-2.93899
*ADM	δ_4	-0.01350	0.00561	-2.40657
*ADM^2	δ_5	0.00353	0.00141	2.49903
STR	δ_6	-0.00556	0.00631	-0.88244
Variance Parameters	$\sigma^2 = \sigma_U^2 + \sigma_V^2$	0.05623	0.01551	3.62678
	$\gamma = \frac{\sigma_U^2}{\sigma_U^2 + \sigma_V^2}$	0.92430	0.01769	52.26312
Log Likelihood Function		996.80100		
LR test of the one-sided error (H_0: $\gamma = 0$)		157.85600		
number of restrictions		8		

*Variable is scaled:
SALARY is SALARY/1,000
ADM is ADM/100
ADM^2 is ADM^2/1,000,000

Table 5.13
Maximum-Likelihood Estimates for Parameters of Stochastic Production Frontier and Inefficiency Models for Oklahoma School Districts (Panel Data Model)
Dependent Variable: ln (IT11)

Variable	Parameter	Coefficient	Standard Error	t-ratio
Stochastic Production Frontier				
Constant	β_0	-4.54858	0.98441	-4.62062
MIN	β_1	-0.19946	0.02614	-7.63135
LUNCH	β_2	-0.27029	0.03056	-8.84366
ln(I)	β_3	-0.10112	0.70362	-0.14372
ln(O)	β_4	2.23393	0.73143	3.05420
$[\ln (I)]^2$	β_5	0.14312	0.09843	1.45401
$[\ln (0)]^2$	β_6	-0.00318	0.06895	-0.04613
ln(I)ln(O)	β_7	-0.26462	0.14836	-1.78364
DUM1	β_8	0.06751	0.00975	6.92451
DUM2	β_9	0.08693	0.00817	10.64080
Inefficiency Equation				
Constant	δ_0	0.21630	0.28949	0.47419
*SALARY	δ_1	0.03665	0.09988	0.36693
YRSEXP	δ_2	-0.42693	0.08133	-5.24936
DEG	δ_3	-0.00026	0.00322	-0.08010
*ADM	δ_4	-0.01539	0.00321	-4.79965
*ADM^2	δ_5	0.00004	0.00001	4.89964
STR	δ_6	-0.00801	0.00575	-1.39272
Variance Parameters	$\sigma^2 = \sigma_U^2 + \sigma_V^2$	0.04012	0.00431	9.30646
	$\gamma = \frac{\sigma_U^2}{\sigma_U^2 + \sigma_V^2}$	0.91177		
Log Likelihood Function		849.07500		
LR test of the one-sided error (H_0: $\gamma = 0$)		255.24600		
number of restrictions		8		

*Variable is scaled:
SALARY is SALARY/1,000
ADM is ADM/100
ADM^2 is ADM^2/10,000

The likelihood ratio (LR) tests of the one sided error was conducted against OLS. According to Coelli (1995a), the Wald and two-sided LR tests are of incorrect size. A test based on the third moment of OLS residuals (in the COLS technique) was found to be of correct size. However, the one-sided LR test appeared to have higher power. The test statistic for the LR test involving $\gamma = 0$ has a mixed Chi-square distribution. In the case of the one-sided LR test, Coelli (p. 252) suggests that "the critical value for a test of size α is equal to the critical value of the chi-square (χ^2) distribution for a standard test of size 2α (e.g. the total value for a 5 percent test is reduced from 3.84 to 2.71)."

According to the LR tests, presented in Table 5.14, the OLS specification was rejected for all grades. Rejection of OLS and the appearance of the high estimated values of the γ's and their t-ratios suggest that inefficiencies do exist. Therefore, the traditional average response function does not adequately represent the production structure of Oklahoma school districts under study.

Table 5.14
Likelihood Ratio Test of the Hypotheses for OLS Specification Involving Parameters of the Inefficiency Model

Grade	Hyp	Loglikelihood H_0	χ^2 Statistic df = 8	Critical Value $2\alpha = .1$	Decision Rule
3	H_o: $\gamma = \delta_0 = \delta_1 = \delta_2 = \ldots = \delta_6 = 0$	455.399	128.888	13.3616	Reject H_o
7	H_o: $\gamma = \delta_0 = \delta_1 = \delta_2 = \ldots = \delta_6 = 0$	659.602	168.176	13.3616	Reject H_o
8	H_o: $\gamma = \delta_0 = \delta_1 = \delta_2 = \ldots = \delta_6 = 0$	917.873	157.856	13.3616	Reject H_o
11	H_o: $\gamma = \delta_0 = \delta_1 = \delta_2 = \ldots = \delta_6 = 0$	721.452	254.903	13.3616	Reject H_o

5.3 The Inefficiency Equation

The inefficiency model for Oklahoma school districts is assumed to be district specific (equation (4.5) in Chapter IV with all of its assumptions). The quadratic term ADM^2 has the right sign and is statistically significant. Thus, inclusion of ADM^2 in the inefficiency mode is appropriate.

To test the existence of a stochastic, as well as deterministic component in the error term, a likelihood ratio test is conducted. The results are presented in Table 5.15.

Table 5.15
Likelihood Ratio Test of the Hypotheses for Existence of a Stochastic Component in the Error Term

Grade	Hyp	Loglikelihood H_0	χ^2 Statistic df = 2	Critical Value $2\alpha = .1$	Decision Rule
3	$H_0{:}\gamma_0 = 0$	473.77	92.14	4.605	Reject H_0
7	$H_0{:}\ \gamma_0 = 0$	677.538	132.304	4.605	Reject H_0
8	$H_0{:}\ \gamma_0 = 0$	940.845	111.912	4.605	Reject H_0
11	$H_0{:}\ \gamma_0 = 0$	760.24	177.328	4.605	Reject H_0

The results in Table 5.15 suggest that the error term in the model has a stochastic component for all grades.

The null hypothesis that the inefficiency effects are not a linear function of the right hand side variables in equation (4.5), H_0: $\delta_1 = \ldots\ldots\delta_6 = 0$ is also rejected for all grades. Results are presented in Table 5.16.

Table 5.16
Likelihood Ratio Test of the Hypotheses of Linear Restrictions for Parameters of the Inefficiency Model

Grade	Hyp	Loglikelihood H_0	χ^2 Statistic df = 6	Critical Value $2\alpha = .1$	Decision Rule
3	H_0: $\delta_1 = \delta_2 = \ldots = \delta_6 = 0$	455.609	70.535	10.644	Reject H_0
7	H_0: $\delta_1 = \delta_2 = \ldots = \delta_6 = 0$	661.78	113.067	10.644	Reject H_0
8	H_0: $\delta_1 = \delta_2 = \ldots = \delta_6 = 0$	917.873	97.783	10.644	Reject H_0
11	H_0: $\delta_1 = \delta_2 = \ldots = \delta_6 = 0$	721.452	139.869	10.644	Reject H_0

The rejection of the null hypothesis in Table 5.16 indicates that the joint effects of these explanatory variables on the level of technical inefficiencies is significant.

According to inefficiency parameters estimate, presented in Tables 5.10-5.13, increasing teacher salary reduces inefficiency by a statistically significant amount for the 3rd and 7th grades, which is consistent with Adkins and Moomaw (1997). For grades 8 and 11 the effect of salary on efficiency is negative but statistically the effects are not significant. Years of experience affects technical efficiency positively in all grades. Adkins and Moomaw's results agree with these findings except for the 3rd grade.

Teachers holding advanced degrees have a positive effect on efficiency for grades 8 and 11 and a negative effect on grades 3 and 7. However, the effects are statistically significant for grades 3 and 8 and not grades 7 and 11. This may suggest that for lower grades, teachers holding a bachelors degree are more desirable than the ones with more advanced degrees.

The effect of the student/teacher ratio on efficiency is not statistically significant for any grade. This implies that the positive or negative effect of increasing student/ teacher ratio on the efficiency is negligible.

In summary, the estimation of the model under this study generally supports Adkins and Moomaw's (1997) results and particularly is consistent with their results in that larger school districts have a greater degree of technical efficiency and the differences seem to stem from changes in data availability in the years considered in these studies.

The optimum district size based on each grade results are computed and presented in Table 5.17.

Table 5.17
The Optimum ADM Based on
Each Grade's Results from Panel Data

Grade	Optimum ADM
3	21385
7	20135
8	19142
11	20676

Table 5.17 suggests that the optimum district size is between 19,142 and 21,383.

However, these findings are conditional on the homoscedasticity of the model. If the data are heteroscedastic, as the results from White's tests suggest, then inferences may not be statistically valid. To the author's knowledge, there is no estimator of panel data for the stochastic frontier model where efficiencies are determined by exogenous variables that account for heteroscedasticity.

To determine the degree of robustness of this model, a DEA is performed that permits some flexibility in this specification. In addition, DEA allows more suitable modeling of district level production function since it allows for multiple output.

5.4 DEA and Second Stage Tobit

In this section, the results of DEA and second stage Tobit regression are discussed.

5.4.1 DEA Model

The first stage output-oriented DEA model:

$$Scores_{it} = L\ (MIN, LUNCH, I, O) \qquad (5.2)$$

Includes the same outputs and inputs as the SFR model. The SFR model estimation suggests that MIN and LUNCH have a negative effect on the test scores. Since output-oriented DEA is a maximization problem then the complement of MIN and LUNCH instead of the variables themselves are considered:

$$Scores_{it} = L\ (MIN^*, LUNCH^*, I, O)$$

where:

$Scores_{it}$	IT3, IT7, CRT8, CRT11
MIN*	percentage of nonminority students (1-MIN)
LUNCH*	percentage of students not eligible for subsidized or reduced LUNCH (1-LUNCH)
I, O	expenditures

Equation (5.2) is estimated using DEAP (2.1) software developed by T.J. Coelli (1996). Table 5.18 presents basic information on the distribution of efficiency scores generated by the DEA model under constant return to scale (CRS) and variable return to scale (VRS) assumptions for 1996-1999.

Table 5.18
Summary Statistics for DEA Efficiency Scores

	1996-1997		1997-1998		1998-1999		Panel	
	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
Mean	.8918	.9256	.8743	.9102	.8557	.8837	.8706	.9065
SD	.8486	.672	.9287	.7165	.8956	.7279	.9108	.7262
Minimum	.581	.673	.529	.647	.561	.605	.529	.605
Maximum	1	1	1	1	1	1	1	1

There are considerable similarities between efficiency generated under CRS and VRS. The average efficiency assuming CRS for the panel is 87 percent (suggesting an average inefficiency of 13 percent) and, assuming VRS, the average is 90 percent (suggesting an average inefficiency of 10 percent). There is small variation in VRS results, as expected (Kirjavainen and Loikkanen, 1998).

Interestingly, under both CRS and VRS assumptions, the average efficiency scores for the sample has declined and the variation has increased every year, suggesting that school districts actually became less efficient with more variation in the level of efficiency among districts throughout 1996-1999 academic years.

5.4.2 Tobit Regression Model

Tobit regressions are computed using the LIMDEP 7.0 software, which allows for the existence of heteroscedasticity in the model. All the explanatory variables in equation (4.10) as well as the independent variable, CRS efficiency, are considered as the possible source of this misspecification. However, CRS efficiency scores and student/teacher ratio and the size of the school districts as measured by ADM are likely sources of heteroscedasticity. Thus, a heteroscedastic Tobit regression with these variables as sources of heteroscedasticity is computed.

To test the heteroscedasticity hypothesis, a likelihood ratio test is employed:

H_0: Homoscedasticity

H_1: At least one of the variables is a source of heteroscedasticity.

The test statistic:

$$\lambda = -2[\log(\text{likelihood } H_0) - \log(\text{likelihood } H_1)] \sim \chi_J^2$$

where J equals the number of variables considered as a potential source of heteroscedasticity (J = 3) $\lambda = -2[902.5622 - 929.997] = 54.86 \sim \chi_3^2$ suggests that H_0 should be rejected (critical $\chi_3^2 = 7.814$ 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 assumptions of homoscedastic and heteroscedastic error terms in the model, are presented in Table 5.19.

Table 5.19
Tobit Coefficient Estimates of the Efficiency Model
Dependent Variable: Efficiency Estimates from the First-Stage DEA CRS Model

	Homoscedastic		Heteroscedastic	
Variable	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
Constant	.666951	7.584	.726393	8.830
SALARY	-.000005	-1.829	.099482	-1.433
DEG	.089591	4.091	.099482	4.580
YRSEXP	.003067	4.126	.002656	3.900
STR	.018588	12.985	.013556	9.057
ADM	-.000035	-.174	.000008	.048
ADM^2	-.000000	-.160	-.000000	-.488

In both models, DEG, YRSEXP, and STR have statistically significant positive effect on efficiency and the effect of SALARY, ADM, and ADM^2 are all negligible. To determine the magnitude of these effects, the marginal effects of the explanatory variables for heteroscedastic Tobit is computed and presented in Table 5.20.

Table 5.20
Tobit Slope (Marginal Effect) Estimates of the Efficiency Model

Variable	Slope	t-statistic
Constant	6.52323	7.538
SALARY	-.000003	-1.432
DEG	.089337	4.305
YRSEXP	.002385	3.803
STR	.020637	5.607
ADM	.000292	1.386
ADM^2	-.000000	-.489

The effects of teachers salary and the size of school districts are insignificant. For one unit increase in each, teachers holding advanced degrees, teachers years of experience, and student/teacher ratio, the efficiency will improve by .08, .002, and .02, respectively.

5.5 DEA vs. SFR

In order to compare the Stochastic Frontier Regression (SFR) and the Data Envelopment Analysis (DEA) models, the data set on Oklahoma school districts is used in several different ways to facilitate the comparison between the results from these models.

To begin with, the four output categories are used directly in the DEA model. SFR does not allow for multiple outputs. Thus, for comparison purposes, the dependent variable in SFR is computed as the logarithm of the average of all the test scores in the panel. The results of the SFR model is presented in Table 5.21.

Table 5.21
Maximum-Likelihood Estimates for Parameters of Stochastic Production Frontier and Inefficiency Models for Oklahoma School Districts (Panel Data Model)
Dependent Variable: ln (Average of all Test Scores)

Variable	Parameter	Coefficient	Standard Error	t-ratio
Stochastic Production Frontier				
Constant	β_0	-12.69181	1.15505	-10.98811
MIN	β_1	-0.14036	0.01719	-8.16405
LUNCH	β_2	-0.30118	0.01928	-15.62106
ln(I)	β_3	1.95900	0.66211	2.95871
ln(O)	β_4	2.24843	0.65266	3.44504
$[\ln (I)]^2$	β_5	0.15694	0.08202	1.91335
$[\ln (0)]^2$	β_6	0.15582	0.05031	3.09743
ln(I)ln(O)	β_7	-0.57053	0.11323	-5.03870
DUM1	β_8	0.04144	0.00721	5.74854
DUM2	β_9	0.03146	0.00588	5.34921
Inefficiency Equation				
Constant	δ_0	0.55581	0.17862	3.11170
*SALARY	δ_1	-0.20812	0.06965	-2.98807
YRSEXP	δ_2	-0.45472	0.08305	-5.47490
DEG	δ_3	0.00148	0.00214	0.69187
*ADM	δ_4	-0.00908	0.00240	-3.78644
*ADM^2	δ_5	0.00223	0.00058	3.86501
STR	δ_6	0.01119	0.00448	2.49717
Variance Parameters	$\sigma^2 = \sigma_U^2 + \sigma_V^2$	0.16949	0.00247	6.85311
	$\gamma = \frac{\sigma_U^2}{\sigma_U^2 + \sigma_V^2}$	0.84875	0.02975	28.52959
Log Likelihood Function		127.39500		
LR test of the one-sided error (H_0: $\gamma = 0$)		187.34200		
number of restrictions		8		

*Variable is scaled:
SALARY is SALARY/1,000
ADM is ADM/100
ADM^2 is ADM^2/1,000,000

The rankings of the school districts generated by the DEA CRS and SFR efficiency estimation are computed. The Spearman rank correlation between the two models is computed to be .6.

To see whether or not a different functional form would affect the computation of the efficiency scores for the SFR model, a Cobb-Douglas SFR is estimated. The efficiency scores for SFR under translog and Cobb-Douglas specification are almost identical. Therefore, subsequent analysis is based on the results of SFR with translog specification.

Furthermore, the SFR and DEA CRS models are estimated cross-sectionally for each grade and each year of the 3-year period under study. In order to examine the effect of technological changes over the years under study, the panel data for each grade is estimated under both SFR and DEA models.

Thus, 12 different cross-section SFR and DEA CRS (N = 366) and 4 panel (N = 1098) are estimated to obtain the efficiency scores under each model. To obtain the correlation between the results of the two models, school districts are ranked based on their efficiency scores and Spearman Correlation between these rankings are computed. The results are presented in Table 5.22.

Table 5.22
Spearman Rank Correlation Coefficients of DEA CRS and SFR Efficiency Scores

Grade	1996-1997	1997-1998	1998-1999	Panel
3	.82	.85	.77	.80
7	.79	.86	.82	.82
8	.72	.70	.79	.76
11	.87	.81	.80	.80

These results suggest that the efficiency rankings generated by both models are generally similar for the cross-sections as well as the panels which suggests that sample size and technological changes do not seem to have a strong influence on the efficiency ranking.

To compare and examine the effects of the explanatory variables in equation (4.10) on the efficiency, the parameter estimates of SFR and heteroscedastic Tobit are compared.

Recall that for SFR, equation (4.5) has inefficiency as the dependent variable, while in DEA second stage Tobit regression the dependent variable is efficiency rather than inefficiency. Thus, in order to simplify the comparison between the two models, the signs of the SFR parameter estimates are inverted. The results for both models are presented in Table 5.23.

Table 5.23
Parameter Estimates; Dependent Variable: Efficiency Estimates

	SFR		DEA-Tobit Heteroscedastic	
	Coefficient	**t-ratio**	**Coefficient**	**t-ratio**
Constant	-.5558	-3.1117	.72639	8.830
SALARY	.2081	2.988	-.000004	-1.433
YRSEXP	.4547	5.4749	.002656	3.900
DEG	-.0014	-.69187	.099482	4.580
ADM	.0090	3.7864	.000008	.048
ADM^2	-.0022	-3.865	-.000080	-.488
STR	-.01119	-2.497	.013556	9.057

According to the results from Table 5.23, the DEA and SFR models suggest teachers' years of experience and the size of the school districts measured by ADM have a positive effect on efficiency. However, the effect of ADM in the DEA model is not statistically significant. As for the effect of teachers' salary on efficiency, SFR suggests a positive and significant effect vs. DEA suggests the opposite.

Teachers holding advanced degrees, according to the DEA model, affect efficiency positively and statistically significant where SFR suggests just the opposite. With respect to the student/teacher ratio, the effect on efficiency is positive and significant in the DEA model and negative and significant in the SFR model.

One possible reason for a lower correlation between the multiple output DEA and SFR models, as well as differences in Table 5.23, could be the fact that SFR simply does not allow for multiple outputs. Thus, taking the average of all outputs to overcome this restriction may not be appropriate.

Therefore, in the case of multiple outputs, DEA models with second stage Tobit regression may be more reliable to explain the efficiency differences among school districts. The efficiency rankings based on SFR and first-stage DEA CRS for multiple output panel and single output panel data for the third grade are presented in Appendix I.

CHAPTER VI

DATA ENVELOPMENT ANALYSIS

6.1 Model I Specification

In this section a variant of the DEA model where the output measures are expanded is used to achieve a more robust result. The model is:

$$Score_{it} = f(YRSEXP_{it}, DEG_{it}, I_{it}, O_{it}) \quad (6.1)$$

where output score includes all the measures in Chapter V, IT3, IT7, CRT8, and CRT11, plus:

CRT5 Average CRT scores for grade 5

ACT Average ACT scores for all seniors in the district.

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).

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 noninstructional expenditures, I and O, respectively are the traditional inputs in the model.

6.1.1 DEA Estimation

The model in equation (6.1) is estimated using the DEA method for the panel data (1996-1999). Thus the number of observations for the panel is N = 1062. Given the nature of the production frontier function in this study, i.e., multiple outputs, multiple inputs, etc. DEA is better suited for estimation of efficiencies than SFR (Coelli, 1998; Kirjavainen, et al., 1998). The DEA will yield estimates of district efficiencies than can subsequently be modeled as functions of other district characteristics.

6.1.2 Tobit Model

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 is:

$$EFF_{it} = \alpha_0 + MIN_{it}\alpha_1 + LUNCH_{it}\alpha_2 + HHINCOME_{it}\alpha_3 + PVALUATION_{it}\alpha_4 + POVERTY_{it}\alpha_5 + DEGADULTS_{it}\alpha_6 + SED_{it}\alpha_7 + SALARY_{it}\alpha_8 + ADM_{it}\alpha_9 + ADM^2_{it}\alpha_{10} + STR_{it}\alpha_{11} + e_{it} \qquad (6.2)$$

where:

EFF	is the efficiency score generated by DEA
MIN	is as described in Chapter III
LUNCH	is as described in Chapter III
HHINCOME	is Average Household Income (1990) ($)
PVALUATION	is assessed value of property within the boundaries of the district per student ($)
POVERTY	is Poverty Rate (1990)
DEGADULTS	is percentage of adults age 20+ with education beyond high school diploma (1990)
SED	is percentage of students in special education
SALARY	is as described in Chapter III
ADM	is as described in Chapter III
ADM^2	is as described in Chapter III
STR	is as described in Chapter III
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. Elimination of *dependent* districts along with the availability of data results in 354 observations for each of the years in the study. Summary statistics for various years under study are presented in Tables 6.1-6.4

Table 6.1
Summary Statistics for Oklahoma School Districts 1996-1997

Variable	Mean	Std. Dev.	Minimum	Maximum
MIN	0.26537	0.16403	0.00000	0.98749
SED	0.11899	0.02888	0.03974	0.21684
LUNCH	0.49013	0.15790	0.04477	1.00
FTE	89.61236	199.25301	13.01000	2096.88100
STR	15.98945	2.19718	8.76589	21.97037
SALARY	30451.65520	982.65798	27819.10720	35332.95750
DEG	0.33534	0.13530	0.04890	0.78470
YRSEXP	20.40876	3.63724	8.50000	30.66670
I	2874.74186	453.55734	2090.59046	5354.80683
O	1820.17738	456.82875	1099.57657	5242.41826
IT3	61.96328	9.44381	33.00000	92.00000
CRT5	53.87571	5.42624	30.66667	66.66667
IT7	55.59605	7.98356	32.00000	79.00000
CRT8	52.62100	5.33220	34.66667	65.00000
CRT11	73.63319	8.54481	42.20000	94.25000
DROPRATE	0.04049	0.03004	0.00000	0.15780
ACT	19.70480	1.40683	15.40000	23.50000
ADM	1582.49590	3830.25520	157.27000	41196.32000

Table 6.2
Summary Statistics for Oklahoma School Districts 1997-1998

Variable	Mean	Std. Dev.	Minimum	Maximum
MIN	0.26958	0.16377	0.00000	0.99530
SED	0.12191	0.03083	0.04969	0.25809
LUNCH	0.51488	0.16116	0.04668	0.84213
FTE	89.61380	197.19582	13.43000	2106.71755
STR	16.01358	2.20429	8.61200	21.44213
SALARY	29704.06090	1070.71722	27119.53570	34067.73720
DEG	0.33048	0.13191	0.03968	0.78571
YRSEXP	12.67952	2.05703	5.83011	19.21429
I	3013.66509	502.57997	2186.20878	6167.84176
O	1922.65997	454.78303	1185.43086	5029.66495
IT3	62.30508	9.98143	35.00000	91.00000
CRT5	66.18456	7.50435	40.50000	83.33333
IT7	55.24576	8.39864	30.00000	75.00000
CRT8	61.33192	7.40468	39.16667	82.16667
CRT11	71.73898	9.67590	29.00000	93.60000
DROPRATE	0.04083	0.02762	0.00000	0.15090
ACT	20.01855	1.39185	16.00000	23.70000
ADM	1591.40328	3828.05241	161.01000	41309.39000

Table 6.3
Summary Statistics for Oklahoma School Districts 1998-1999

Variable	Mean	Std.Dev.	Minimum	Maximum
MIN	0.27816	0.16532	0.00000	0.99820
SED	0.12717	0.03099	0.06536	0.26357
LUNCH	0.52477	0.16158	0.04544	0.90961
FTE	91.08316	202.66391	13.00000	2165.05590
STR	16.02215	2.11174	8.16278	20.79424
SALARY	30147.39220	1058.27137	27634.59590	33844.86490
DEG	0.31351	0.12963	0.03687	0.80621
YRSEXP	12.74271	2.04143	5.89998	17.88462
I	3142.63675	564.11693	2249.14232	6310.27679
O	2000.65208	467.58847	1321.99328	4680.62600
IT3	62.54237	9.89031	35.00000	93.00000
CRT5	76.02024	9.88747	36.66667	100.00000
IT7	56.26271	8.18043	31.00000	85.00000
CRT8	70.03955	8.94987	36.66667	93.00000
CRT11	64.77475	9.08113	40.28571	93.00000
DROPRATE	0.04078	0.02952	0.00000	0.15410
ACT	19.91385	1.37204	15.63000	23.51000
ADM	1598.49167	3831.98743	148.53000	41471.46000

Table 6.4
Summary Statistics for Oklahoma School Districts 1996-1999

Variable	Mean	Std. Dev.	Minimum	Maximum
MIN	0.27104	0.16430	0.00000	0.99820
HHINCOME	21313.03960	5753.18331	10833.00000	45790.00000
PVALUATION	19982.70710	14535.03490	3639.00000	172102.57900
POVERTY	0.18851	0.07467	0.03320	0.41180
DEGADULTS	0.11420	0.05580	0.02290	0.40100
SED	0.12269	0.03041	0.03974	0.26357
LUNCH	.50993	0.16073	.04477	1.00
FTE	90.10311	199.52986	13.00000	2165.05590
STR	16.00839	2.16948	8.16278	21.97037
SALARY	30101.03610	1081.48607	27119.53570	35332.95750
DEG	0.32644	0.13251	0.03687	0.80621
YRSEXP	15.27700	4.51403	5.83011	30.66670
I	3010.34790	519.93564	2090.59046	6310.27679
O	1914.49647	465.24712	1099.57657	5242.41826
IT3	62.27024	9.76836	33.00000	93.00000
CRT5	65.36017	11.96686	30.66667	100.00000
IT7	55.70151	8.19245	30.00000	85.00000
CRT8	61.33082	10.24532	34.66667	93.00000
CRT11	70.04898	9.86907	29.00000	94.25000
DROPRATE	0.04070	0.02905	0.00000	0.15780
ACT	19.87907	1.39512	15.40000	23.70000
ADM	1590.79695	3826.49269	148.53000	41471.46000

Comparison of Tables 6.1, 6.2, and 6.3 suggests that for the years under study MIN, SED, and LUNCH have been increasing. The increase in LUNCH could possibly be due to lower household income. However, the data on household income for each year is not available, thus the relationship cannot be confirmed. Another possibility for the increase in LUNCH is that more eligible students actually applied for assistance.

SALARY, DEG and YRSEXP have declined over time. Since SALARY is determined based on teachers' degree and experience, it is reasonable to expect these variables to move in the same direction. However, the sharp decline in YRSEXP from 20 years in 1996-1997 to almost 13 years in 1997-1998 could possibly be due to retirement of a group of highly experienced teachers.

Except for CRT11 which has declined every year, all other outputs as measured by test scores, have been increasing. The increase in test scores is especially noticeable in CRT5 and CRT8. However, CRTs are renormed every year. Thus, it is difficult to explain the differences in the performance of Oklahoma schools over time. Since district efficiency is necessarily a value between 0 and 1, a Tobit model is used to estimate the parameters. The variable, efficiency, is truncated from below at 0 and from above by 1. This also ensures that predictions from the model will lie in this interval.

6.2 Results

6.2.1 DEA

The results of the DEA estimation are obtained using DEAP(2.1) software developed by T. J. Coelli and are presented in Table 6.5. The table 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 VRS assumption, the possibility of scale of operation is considered and the efficiency measures are affected by it.

Table 6.5
Summary Statistics for DEA Efficiency Scores, Model I (Panel)

	CRS	VRS
Mean	.82155	.91076
SD	.10845	.06146
Minimum	.436	.706
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 6.6.

Table 6.6
Frequencies of School Districts in Classes Based on Efficiency Scores of the DEA Model I (Panel)

	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

Table 6.6 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.

6.2.2 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 (6.2) by the Tobit regression method, using LIMDEP (7.0) software. In equation (6.2) 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. All the explanatory variables as well as the dependent variable in equation (6.2) are considered. Except for DEGADULTS, LUNCH, and SALARY; all coefficients are statistically significant at the 5 percent level and are likely sources of heteroscedasticity. To test this hypothesis, the likelihood ratio test for heteroscedasticity is performed:

H_0: homoscedasticity

H_1: at least one of the variables is a source of heteroscedasticity

The computed likelihood ratio is:

$$\lambda = -2[\log(\text{likelihood}\, H_0) - \log(\text{likelihood}\, H_1)] \sim \chi^2_J$$

$$\lambda = -2[1017.326 - 1115.880] = 197.108 \sim \chi^2_{11}$$

This ratio suggests that H_0 should be rejected (critical χ^2_{11} = 19.675 at $\alpha = .05$), therefore there is substantial evidence that at least one of the variables "explain" the existence of heteroscedasticity in the Tobit regression.

The results of Tobit regression under the assumption of homoscedasticity and heteroscedasticity are presented in Table 6.7.

Table 6.7
Tobit Regression Coefficient Estimates of the Efficiency Model I
Dependent Variable: Efficiency Estimates from the
First-Stage DEA Model under CRS Assumption (Panel)

	Homoscedastic		Heteroscedastic	
Variable	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
Constant	1.468993	16.729	1.347191	15.554
MIN	-.124989	-6.306	-.148579	-8.543
LUNCH	-.106253	-3.728	-.140669	-5.474
HHINCOME	.000002	2.618	.000002	2.717
PVALUATION	-.000000	-2.941	-.000001	-3.926
POVERTY	-.175100	-2.656	-.103581	-1.665
DEGADULTS	.179362	2.89	.222383	4.362
SED	-.134157	-1.549	-.363291	-4.111
SALARY	-.000028	-11.003	-.000022	-9.252
ADM	-.000008	-3.977	-.000002	-2.154
ADM^2	.000000	3.094	.000000	1.76
STR	.017379	10.584	.015819	10.386

The comparison of the results of the two models in Table 6.7 suggests that under the assumption of homoscedasticity the coefficient of all variables, except for SED, are statistically significant. However, when heteroscedasticity is considered, SED becomes statistically significant also, but POVERTY and ADM^2 become statistically insignificant.

Based on the heteroscedastic Tobit regression results in Table 6.7, 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. Therefore, districts with high property valuation could potentially have low income families. 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 6.8.

Table 6.8
Tobit Slope (Marginal Effect) Estimates of the Efficiency Model I

Variable	Slope	t-statistic
Constant	1.343758	14.668
MIN	-.155574	-9.246
LUNCH	-.145606	-6.076
HHINCOME	.000002	3.286
PVALUATION	-.000001	-4.395
POVERTY	-.077231	-1.256
DEGADULTS	.235664	4.629
SED	-.409743	-4.713
SALARY	-.000023	-9.360
ADM	-.000002	1.700
ADM^2	.000000	1.755
STR	.016689	10.893

The results of table 6.8 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 not significantly different from zero. Also, school district size (ADM) does not seem to have a significant effect on efficiency, which is consistent with Kirjavainen, et al. (1998).

6.3 Model II Specification

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

$$Score_{it} = f(I_{it}, O_{it}) \quad (6.3)$$

Where outputs; score, and inputs; I, O are as defined in equation (6.1). 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.

6.3.1 Results

The results of the DEA estimation are presented in Table 6.9. The table contains the basic information on the distribution of the efficiency scores generated by DEA under CRS and VRS assumptions.

Table 6.9
Summary Statistics for DEA Efficiency Scores, Model II (Panel)

	CRS	VRS
Mean	.75133	.88691
SD	.11801	.06065
Minimum	.332	.677
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.

The efficiency equation estimated in the second stage using the Tobit regression method is equation (6.2) including YRSEXP and DEG as explanatory variables:

$$EFF_{it} = \alpha_0 + MIN_{it}\alpha_1 + LUNCH_{it}\alpha_2 + HHINCOME_{it}\alpha_3 + PVALUATION_{it}\alpha_4 + POVERTY_{it}\alpha_5 + DEGADULTS_{it}\alpha_6 + SED_{it}\alpha_7 + SALARY_{it}\alpha_8 + ADM_{it}\alpha_9 + ADM^2_{it}\alpha_{10} + STR_{it}\alpha_{11} + YRSEXP_{it}\alpha_{12} + DEG_{it}\alpha_{13} + e_{it} \quad (6.4)$$

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 (6.4) are considered as the possible source of this misspecification. Except for SALARY, YRSEXP, and DEG all of the variables are likely sources of heteroscedasticity. To test this hypothesis a likelihood ratio test is performed:

H_0: homoscedasticity

H_1: at least one of the variables is a source of heteroscedasticity

The test statistic:

$$\lambda = -2[\log(\text{likelihood}\, H_0) - \log(\text{likelihood}\, H_1)] \sim \chi^2_J$$

where J equals the number of variables considered as a potential source of heteroscedasticity (J = 10)

$$\lambda = -2[1283.976 - 1336.689] = 105.426 \sim \chi^2_{10}$$

suggests that H_0 should be rejected (critical $\sim \chi^2_{10} = 18.307$ at $\alpha = .05$), therefore, there is substantial evidence that at least one of the variables "explain" the existence of heteroscedasticity in the Tobit regression.

The Tobit coefficient estimates computed under the assumption of homoscedastic and heteroscedastic error terms in the model are computed and presented in Table 6.10.

Table. 6.10
Tobit Regression Coefficient Estimates of the Efficiency Model II
Dependent Variable: Efficiency Estimates from the
First Stage DEA Model under CRS Assumption (Panel)

	Homoscedastic		Heteroscedastic	
Variable	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
Constant	.709585	9.143	.810359	10.048
MIN	-.082703	-4.859	-.098662	-5.748
LUNCH	-.206432	-8.373	-.225267	-10.036
HHINCOME	.000002	2.207	.000001	1.753
PVALUATION	-.000002	-8.209	-.000002	-8.040
POVERTY	-.117564	-2.079	-.133885	-2.415
DEGADULTS	.304638	5.377	.331352	7.313
SED	-.24076	-3.231	-.292922	-3.734
SALARY	-.000005	-2.265	-.000006	-2.354
ADM	-.000008	-1.977	-.000002	-.592
ADM^2	-.000000	1.153	.000000	.314
STR	.023476	16.512	.019400	14.219
YRSEXP	-.001703	-3.258	-.001514	-3.096
DEG	.000223	1.165	.000034	.259

The results in Table 6.10 suggest that under both assumptions, MIN, LUNCH, POVERTY, DEGADULTS, SED, and STR are the only variables where their effects on efficiency are significantly different from zero. To examine the magnitude of the effects, the marginal effects of these variables on efficiency, based on the heteroscedastic Tobit model, are computed and presented in Table 6.11.

Table 6.11
Tobit Slope (Marginal Effect) Estimates of the Efficiency Model II

Variable	Slope	t-statistic
Constant	.8103364	10.048
MIN	-.098816	-5.766
LUNCH	-.225242	-10.039
HHINCOME	.000001	1.763
PVALUATION	-.000002	-8.052
POVERTY	-.133537	-2.411
DEGADULTS	.331766	7.328
SED	-.293425	-3.741
SALARY	-.000006	-2.354
ADM	-.000002	-.590
ADM^2	.000000	.314
STR	.019410	14.243
YRSEXP	-.001514	-3.096
DEG	.000033	.259

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 6.11 suggest that a one percent increase in each MIN, LUNCH, POVERTY, and SED decreases efficiency by almost .1, .23, .13, and .29, respectively; and a one percent increase in DEGADULTS increases efficiency by almost .33. Also, for each unit increase in student/teacher ratio (STR), efficiency increases by .02. This is consistent with Kirjavainen, et al. (1998).

6.4 Conclusion

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, 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 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 affects efficiency positively; however, the relationship is not strong. The optimal school district size as measured by ADM is computed to be around 21,460 in both models.

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 these two models. The efficiency rankings are presented in Appendix II.

CHAPTER VII

SUMMARY AND CONCLUSIONS

The primary objective of this dissertation was to estimate production efficiency of Oklahoma school districts in light of possible empirical specification problems caused by possible structural changes in data collection and by the possible heteroscedasticity in the error term. The existing literature indicates a relatively small number of applications of the stochastic production frontier approach to school districts, none of which considers the existence of heteroscedasticity.

A review of production frontier studies suggests the use of different methods of estimation based on the available data and model specification. The Stochastic Frontier Regression (SFR) and Data Envelopment Analysis (DEA) are the estimators considered in this study.

First, in Chapter V a two equation stochastic production frontier was estimated using both SFR and DEA estimators, and the results are compared to see how robust these methods are. The time period under study consists of a three-year period that encompasses the 1996-97 through the 1998-99 academic school years. In this chapter the data set includes observations on 366 so-called *independent* (K-12) school districts and the results suggest that:

1. There are varying degrees of technical inefficiency among Oklahoma school districts. Therefore, the average response function (e.g., OLS) cannot adequately represent the production function of Oklahoma school districts. Thus a two-equation stochastic production frontier model is a preferred alternative model. This conclusion is supported on the basis of hypothesis tests that support the notion that inefficiency effects have both systematic and random components.
2. The existence of heteroscedasticity in the data was supported based on hypothesis tests. Although Jacques and Brorsen claim, without providing evidence, that heteroscedasticity in these data exist, they take the position that it is solely a function of the number of tests taken. Results in Chapter V suggest that there are other factors that contribute to heteroscedasticity in the data and that the existence of such misspecification should be checked before proceeding. An attempt to extend the computer program, Frontier 4.1 (Coelli, 1995,) to account for heteroscedasticity was not successful. However a model assuming homoscedastic error was estimated.

The estimation results of the homoscedastic model suggest that the signs of the coefficients of explanatory variables are in general as expected but these estimates may not be robust in the presence of heteroscedasticity. The results from the potentially misspecified homoscedastic production frontier were compared to a heteroscedastic Tobit model estimated from DEA efficiencies and the results are fairly similar. Perhaps the biases created by heteroscedasticity are not very large.

3. In addition to the problem of heteroscedasticity, since the model consists of multiple outputs, the existing literature suggests the use of distance-functions rather than stochastic frontier functions. Thus, the non-parametric approach to estimate efficiency of Oklahoma school districts was employed, i.e., the DEA approach.

   DEA is not very useful in answering questions of whether money matters; the production function is not parameterized and it yields no estimates of the various spending elasticities. DEA suffers from a lack of well-known statistical properties. So, a second-stage Tobit regression 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), and student/teacher ratio (STR) on the efficiency scores generated by the DEA model. Tobit regression is appropriate since the efficiency scores (dependent variable) are between 0 and 1. However, heteroscedasticty was accounted for in the Tobit regression and thus, the coefficient estimates of the so called efficiency variables are found to be in general consistent with expected hypothesis; however, with the exception of two variables, teachers holding advanced degrees and the student/teacher ratio, the effects of other variables on efficiency are not significantly different from zero.

4. It could be argued that these estimates are more reliable than those of past studies, which were based on the average response function, homoscedastic stochastic production frontier, and DEA with second-stage homoscedastic

Tobit regression. Thus, school districts are ranked based on their efficiency estimates computed using SFR and DEA CRS models and are presented in the Appendix.

Second, based on the findings in Chapter V, DEA may be a more appropriate method of estimation given the nature of the data and objectives of the school districts. Thus, Chapter VI is devoted to DEA estimation of more sophisticated specifications.

In Chapter VI the data consisted of 354 *independent* (K-12) school districts for the three-year period, which are used in various specifications and efficiencies are estimated using DEA. The specification of the model in this chapter include variables not included in the first and/or second stage of the model in Chapter V but seem to affect efficiency measures. Thus, the model considered in this chapter has more output measures in the first stage than the Chapter V model. Also, in the second stage, the efficiency equation, the model includes more explanatory variables which may help explain the efficiency variations among school districts. The analysis of this model also suggests that inefficiency exists among Oklahoma school districts. In the second stage, the Tobit regression model, the efficiency variables included environmental variables that school districts have no control over as well as nontraditional inputs that school districts have control over but were not included in the first stage. Here it seems that environmental variables over which school districts have no control, e.g., percentage of minority students (MIN), percentage of students eligible for reduced or free lunch (LUNCH), the poverty rate in the districts (POVERTY), percentage of students in special education (SED), and percentage of students who have an adult age 20+ with higher than a high school diploma in their household (DEGADULT) are the variables that could possibly explain the efficiency differences among the school districts. The nontraditional inputs (e.g, teacher salary) do not seem to hold much explanatory power over efficiency except for the student/teacher ratio and even that is not very strong.

Therefore, based on the results of the DEA model in Chapter VI, it may be appropriate to conclude that the key factors affecting efficiency measures among Oklahoma school districts are primarily the students' characteristics and family environment, i.e., students' socioeconomic characteristics. Thus, an increase in spending on education may do very little for improving efficiency.

In conclusion this study, based on the results of both Chapters V and VI, suggests that:

- Variables that are not under the control of school districts seem to affect efficiency.
- The method of estimation affects the results.
- District size effects are consistent in all methods and are around 20,000. This is also consistent with Adkins and Moomaw's (1997) results.
- Use of cross-section data may be preferred to panel-data.
- Use of new Census data makes the external variables more up to date.
- Since the data seem to be heteroscedastic because of district size as well as other variables (e.g., student/teacher ratio, efficiency scores, etc.). The DEA may be a more appropriate method of estimation. Also, since the available data is at the district level and cannot be disaggregated at the school level, the DEA method is probably a better approach.

Subsequent research could proceed in several different directions. Notably, it may be useful to isolate nontraditional inputs which school districts have control over, from socioeconomic variables over which school districts have no control in the efficiency model (second-stage) to see whether any policy implications can be drawn from these estimations. Also, assuming that proxies for input and output prices can be extracted from the data, estimating a cost function may reveal useful information about the efficiency of the school districts in Oklahoma.

REFERENCES

Abdul Rahman, Rozina (1996). "The Determinants of Oklahoma School Districts' Efficiency: A Frontier Analysis," Dissertation, Oklahoma State University, Stillwater, Oklahoma.

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

Aigner, D.J. and S.F. Chu (1968). "On Estimating the Industry Production Function," *American Economic Review*, Vol. 58, No. 4.

Aigner, D.J., C.A.K. Lovell, and P. Schmidt (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics*, Vol. 6, No. 1.

Anderson, R.G. and J.G. Thursby (1986). "Confidence Intervals for Elasticity Estimators in Translog Models," *The Review of Economics and Statistics*, Vol. 68, No. 4.

Barrow, M.M. (1991). "Measuring Local Education Authority Performance: A Frontier Approach," *Economics of Education Review*, Vol. 10, No. 1.

Battese, G.E. (1992). "Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics," *Agricultural Economics*, Vol. 7.

Battese, G.E. and G.S. Corra (1977). "Estimation of a Production Frontier Model: With Application to the Pastoral Zone of Eastern Australia," *Australian Journal of Agricultural Economics*, Vol. 21.

Battese, G.E. and T.J. Coelli (1992). "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India," *The Journal of Productivity Analysis*, Vol. 3.

Battese, G.E. and T.J. Coelli (1993). "A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects," Working Papers in Econometrics and Applied Statistics, No. 69, Department of Econometrics, University of New England, Armidale, NSW, Australia

Battese, G.E. and T.J. Coelli (1995). "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data," *Empirical Economics*, Vol. 20.

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

Beattie, B.R. and C.R. Taylor (1993). *The Economics of Production*, Krieger Publishing Company, Florida.

Bessent, A.M. and G.W. Bessent (1980). "Comparing the Comparative Efficiency of Schools Through Data Envelopment Analysis," *Education Administration Quarterly*, Vol. 16.

Bessent A., W. Bessant, J. Kennington, and B. Reagan (1982). "An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District," *Management Science*, Vol. 28.

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

Caudill, S., J.M. Ford, and D.M. Gropper (1995). "Frontier Estimation and Firm-specific Inefficiency Measures in the Presence of Heteroscedasticity," *Journal of Business and Economic Statistics*, Vol. 13, No. 1

Chirikos, T.N. and A.M. Sears (2000). "Measuring Hospital Efficiency: A Comparison of Two Approaches," *Health Services Research*, Vol. 36, No. 6.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau (1973). "Transcendental Logarithmic Production Frontiers," *Review of Economics and Statistics*, Vol. 55, No. 1.

Chung, J.W. (1994). *Utility and Production Function*, Blackwell Publishers, Cambridge.

Cobb, S. and P. Douglas (1928). "A Theory of Production," *American Economic Review*, Vol. 18.

Coelli, T.J. (1994). "A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation." Draft.

Coelli, T.J. (1995a). "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis," *The Journal of Productivity Analysis*, Vol. 6.

Coelli, T.J. (1995b). "Recent Developments in Frontier Modeling and Efficiency Measurement," *Australian Journal of Agricultural Economics*, Vo. 39, No. 3.

Coelli, T.J. (1996). "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.

Coelli, T.J. (1997). "A Multi-Stage Methodology for the Solution of Oriented DEA Models," paper presented to the Taipei International Conference on Efficiency and Productivity Growth, Taipei, June 20-21.

Coelli, T.J. and G. Battese (1996). "Identification of Factors Which Influence the Technical Inefficiency of Indian Farmers," *Australian Journal of Agricultural Economics*, Vol. 40, No. 2, (August).

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

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

Cornes, R. (1992). *Duality and Modern Economics*, Cambridge University Press.

Dean, J. (1951). *Managerial Economics*, Prentice-Hall Publishers: Englewood Cliffs, N.J.

Deller, S.C. and E. Rudnicki (1993). "Production Efficiency in Elementary Education: The Case of Maine Public Schools," *Economics of Education Review*, Vol. 12, No. 1.

Diewert, W.E. (1971). "An Application of the Shephard Dualtiy Theorem: A Generalized Leontief Production Function," *Journal of Political Economy*, Vol. 79.

Diewert, W.E. (1974). "Application of Duality Theory," in *Frontiers of Quantitative Economics*, Vol. II, ed.

Diewert, W.E. and T.J. Wales (1987). "Flexible Functional Forms and Global Curvature Conditions," *Econometrica,* Vol. 55, No. 1.

Färe, R. and C.A.K. Lovell (1978), "Measuring the Technical Efficiency of Production," *Journal of Economic Theory,* Vol. 19.

Farrell, M.J. (1957). "The Measurement of Productive Efficiency," *Journal of Royal Statistical Society*, Series A, Vol. 120, No. 3.

Fuss, M., D. McFadden, and Y. Mundlak (1978). *Production Economics: A Dual Approach to Theory and Application,* North Holland, Amsterdam, pp. 219-268.

Greene, W.H. (1980a). "Maximum Likelihood Estimation of Econometric Frontier Functions," *Journal of Econometrics*, Vol. 13.

Greene, W. H. (1980b), "On the Estimation of a Flexible Frontier Production Model," *Journal of Econometrics*, Vol. 13.

Greene, W.H. (1990), "A Gamma-Distributed Stochastic Frontier Model," *Journal of Econometrics*, Vol. 46.

Greene, W.H. (2000). *Econometric Analysis*, Fourth Edition, Prentice-Hall.

Griffiths, W.E., R.C. Hill, and G.G. Judge (1993). *Learning and Practicing Econometrics*, John Wiley & Sons.

Hanushek, E.A. (1979). "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *The Journal of Human Resources*, Vol. 14, No. 3.

Hanushek, E.A. (1986). "The Economics of Schooling: Production and Efficiency in the Public Schools," *Journal of Economic Literature*, Vol. 24.

Hanushek, E.A. (1987). "Educational Production Functions," *Economics of Education: Research and Studies*, George Psacharopoulos, ed., Oxford; New York: Pergamon Place.

Hanushek, E.A. (1989a). "Expenditures, Efficiency, and Equity in Education: The Federal Government's Role," *American Economic Association Papers and Proceedings*, Vol. 79, No.2.

Hanushek, E.A. (1989b). "The Impact of Differential Expenditures on School Performance," *Education Research*, Vol. 18.

Hanushek, E.A. (1994). "Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald," *Education Research*, Vol. 23.

Hanushek, E.A. (1996). "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.

Hanushek, E.A., S. Rivkin, and L. Taylor (1995). "Aggregation and the Estimated Effects of School Resources," Rochester Center for Economic Research, Working Paper No. 397.

Hardi, K. (1999). "Estimation of Doubly Heteroscedastic Stochastic Frontier Cost Function," *Journal of Business & Economic Statistics*, Vol. 17. No. 3.

Heckman, J., C. Heinrich, and J. Smith (1997). "Assessing the Performance of Performance Standards in Public Bureaucracies," *American Economic Association Papers and Proceedings*, Vol. 87, No. 2.

Henderson, J.M. and R.E. Quandt (1980). *Microeconomic Theory: A Mathematical Approach*, Third Edition, McGraw-Hill.

Himmelblau, D.M. (1972). *Applied Nonlinear Programming*, McGraw-Hill.

Hoxby, C.M. (1994a). "Do Private Schools Provide Competition for Public Schools?" NBER Working Paper No. 4978, Cambridge, MA: National Bureau of Economic Research, Inc.

Hoxby, C.M. (1994b). "Does Competition Among Public Schools Benefit Students and Taxpayers?" NBER Working Paper No. 4979, Cambridge, MA: National Bureau of Economic Research, Inc.

Huang, C.J. and J. Liu (1992). "Estimation of Non-Neutral Stochastic Frontier Production Function," *The Journal of Productivity Analysis*, Vol. 3.

Jacques, C. and V.W. Brorsen (1997). "Relationship between School District Expenditures in Various Categories and Student Performance," Department of Agricultural Economics, Oklahoma State University, Working Paper.

Johnson, J. (1960). "The Structure of Educational Costs: Multiproduct Cost Functions for Primary and Secondary Schools in Latin America," *Economics of Education Review*, Vol. 5, No. 1.

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*, Vol. 19.

Kabir, Md. M. (1998). "Estimation of Stochastic Cost Frontier Models: The Case of Nebraska School Districts," Dissertation, University of Nebraska, Lincoln, Nebraska.

Kenedy, P. (1992). *A Guide to Econometrics*, Third Edition, The MIT Press, Cambridge, MA.

Kirjavainen, T. and H.A. Loikkanen (1998). "Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit Analysis," *Economics of Education Review*, Vol. 17, No. 4.

Kumbhakar, S.C. (1990). "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency," *Journal of Econometrics*, Vol. 26.

Kumbhakar, S.C. (1997). "Efficiency Estimation with Heteroscedasticity in a Panel Data Model," *Applied Economics*, Vol. 29.

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*, Vol. 9.

Leibenstein, H. and S. Maital (1992). "Empirical Estimation and Partitioning of X-Inefficiency: A Data-Envelopment Approach," *American Economic Association Papers and Proceedings*, Vol. 82, No. 2.

McCarthy, T.A. and S. Yaisawarng (1993). "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.

Metzger, M.R. and S. Barta (1999). "The Determinants of Oklahoma K-12 Educational Performance," report to the Oklahoma State Senate Staff.

Moomaw, R.L. and M. Yusof (1995). "The Determinants of Oklahoma K-12 Educational Performance," report to the Oklahoma State Senate Staff.

Nerlove, M. (1963). "Returns to Scale in Electricity Supply," in C.F. Christ et al. (eds.), *Measurement in Econometrics: Studies in Mathematical Economics and Econometrics*, in memory of Yehuda Grumfield, Stanford University Press.

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

Pindyck, R.S. and D.L. Rubinfeld (1998). *Econometric Models and Economic Forecasts*, Irwin/McGraw-Hill, Fourth Edition.

Pitt, M.M. and L.F. Lee (1981). "Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics,* Vol. 9.

Reifschneider, D. and R. Stevenson (1991). "Systematic Departure from the Frontier: A Framework for the Analysis of Firm Inefficiency," *International Economic Review*, Vol. 32, No. 3.

Rubinfeld, D.L. (1985). "The Economics of the Local Public Sector," in *Handbook of Public Economics*, A. Auerbach and M. Feldstein (eds.), Amsterdam: North Holland.

Schmidt, P. (1975). "On the Statistical Estimation of Parametric Frontier Production Functions," *Review of Economics and Statistics*, Vol. 58.

Schmidt, P. (1985-86). "Frontier Production Function." *Econometric Reviews*, Vol. 4, No. 2.

Schmidt. P. and R.C. Sickles (1984). "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics*, Vol. 2.

Stevenson, R.E. (1980). "Likelihood Function for Generalized Stochastic Frontier Estimation," *Journal of Econometrics*, Vol. 13.

Taubman, P. (1989). "Role of Parental Income in Educational Attainment," *American Economic Association Papers and Proceedings*, Vol. 79, No. 2.

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

Terrell, D. (1996). "Incorporating Monotonicity and Concavity Conditions in Flexible Functional Forms," *Journal of Applied Econometrics*, Vol. II.

Yuengert, A.W. (1993). "The Measurement of Efficiency in Life Insurance: Estimates of a Mixed Normal-Gamma Error Model," *Journal of Banking and Finance*, Vol. 17.

Zarembka, P. (1970). "On the Empirical Relevance of the CES Production Function," *Review of Economics and Statistics*, Vol. 52.

Zellner, A., J. Kmenta, and J. Dreze (1966). "Specification and Estimation of Cobb-Douglas Production Models," *Econometrica*, Vol. 34.

APPENDIX I

APPENDIX 1A

Ranking of School Districts by their Efficiency Estimates Generated by the Single Output Model, 3rd Grade ITBS Scores, Using Panel Data

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
ACHILLE	495	172	161	139	118	293	279
ADA	2780	143	120	106	65	260	238
ADAIR	908	321	263	183	154	93	66
AFTON	446	257	266	118	104	36	40
ALEX	369	250	278	84	169	308	321
ALINE-CLEO	217	192	251	8	62	54	132
ALLEN	413	76	179	29	76	350	343
ALTUS	4703	228	186	129	103	159	145
ALVA	1067	110	220	105	187	174	256
AMBER-POCASSET	452	94	124	41	138	131	175
ANADARKO	2032	324	294	319	294	324	273
ANTLERS	1154	175	108	90	44	96	95
ARAPAHO	294	220	274	219	309	24	108
ARDMORE	3414	114	231	153	220	135	216
ARKOMA	476	319	266	343	317	287	267
ARNETT	186	37	308	16	281	22	313
ASHER	227	365	363	87	157	9	1
ATOKA	937	171	242	161	206	261	245
BALKO	158	298	359	364	366	365	366
BARNSDALL	479	341	331	323	301	268	211
BARTLESVILLE	6456	109	89	168	146	246	213
BATTIEST	351	304	317	355	358	349	353
BEAVER	413	177	272	314	355	317	338
BEGGS	966	281	260	291	276	231	201
BENNINGTON	256	77	183	77	132	90	191
BERRYHILL	1002	134	63	30	13	259	184
BETHANY	993	138	85	259	157	180	127
BETHEL	1027	153	111	230	165	95	63
BIG PASTURE	266	326	324	286	306	314	326
BILLINGS	174	262	302	36	206	217	328
BINGER-ONEY	366	208	258	76	189	64	160
BIXBY	3113	57	21	191	56	229	75
BLACKWELL	1713	245	168	316	254	34	26
BLAIR	372	308	259	361	339	316	181
BLANCHARD	1197	232	107	210	79	278	201
BLUEJACKET	257	271	243	104	79	257	153
BOISE CITY	415	200	253	320	336	311	352
BOKOSHE	285	2	1	54	58	30	15
BOONE-APACHE	682	164	191	175	194	161	199
BOSWELL	443	249	324	206	236	134	143

District	ADM*	1996-1997 Rank SFR	1996-1997 Rank DEA	1997-1998 Rank SFR	1997-1998 Rank DEA	1998-1999 Rank SFR	1998-1999 Rank DEA
BRAGGS	257	260	247	328	289	182	122
BRAMAN	160	366	366	312	305	366	365
BRAY-DOYLE	441	135	163	276	311	177	263
BRIDGE CREEK	1031	136	28	279	161	247	121
BRISTOW	1616	117	162	272	278	240	281
BROKEN ARROW	14499	191	71	194	77	202	76
BROKEN BOW	1792	277	290	245	247	184	216
BUFFALO	354	312	341	162	312	83	291
BUFFALO VALLEY	242	160	170	321	317	273	197
BURNS FLAT-DILL CITY	685	290	326	310	334	309	319
BUTNER	288	14	34	12	53	2	1
BYNG	1692	241	235	215	233	352	345
CACHE	1263	105	122	155	155	223	182
CADDO	410	96	178	197	205	204	160
CALERA	564	201	148	115	110	130	103
CAMERON	503	274	243	324	283	87	112
CANADIAN	403	325	322	224	212	189	230
CANEY VALLEY	810	327	281	222	152	346	286
CANTON	480	33	165	180	271	110	245
CANUTE	267	85	39	198	74	53	56
CARNEGIE	770	331	303	339	325	328	301
CARNEY	265	328	280	170	198	357	344
CASHION	423	116	61	274	195	322	219
CATOOSA	2362	199	124	237	204	207	180
CEMENT	275	275	200	186	224	283	305
CENTRAL	448	51	80	22	41	89	87
CENTRAL HIGH	353	294	276	352	347	302	290
CHANDLER	1140	247	218	148	135	102	90
CHATTANOOGA	294	309	337	211	299	171	271
CHECOTAH	1628	336	328	296	280	120	144
CHELSEA	1044	210	194	277	244	292	263
CHEROKEE	405	276	124	89	20	103	9
CHEYENNE	282	69	354	218	361	281	363
CHICKASHA	2980	296	235	295	229	313	260
CHISHOLM	937	70	40	225	159	140	63
CHOCTAW/NICOMA PARK	4627	99	51	236	163	151	76
CHOUTEAU-MAZIE	936	311	276	305	270	334	300
CIMARRON	368	181	222	142	178	234	225
CLAREMORE	3632	137	88	165	108	194	142
CLAYTON	398	348	353	204	266	354	357
CLEVELAND	1674	165	165	249	222	270	223
CLINTON	2036	349	317	345	322	356	346
COALGATE	686	347	348	252	248	243	288
COLBERT	798	287	282	75	92	307	236
COLCORD	693	103	174	96	165	62	84

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
COLEMAN	179	63	225	235	283	74	154
COLLINSVILLE	1593	234	139	253	173	232	115
COMANCHE	1036	269	227	333	323	345	322
COMMERCE	811	255	200	340	297	82	39
COPAN	452	338	268	359	333	256	145
CORDELL	693	23	14	228	184	71	44
COVINGTON-DOUGLAS	335	215	296	273	337	296	329
COWETA	2454	120	71	109	67	58	25
COYLE	377	71	188	208	292	4	36
CRESCENT	640	74	59	166	108	219	138
CROOKED OAK	820	337	305	121	105	43	32
CROWDER	513	46	82	160	227	15	65
CUSHING	2029	285	349	233	340	255	341
CYRIL	457	354	315	79	47	185	156
DALE	624	209	142	78	53	315	221
DAVENPORT	455	310	283	59	79	70	87
DAVIS	926	129	154	20	45	20	59
DEER CREEK	1343	127	49	117	29	112	20
DEER CREEK-LAMONT	258	149	283	203	265	238	316
DEPEW	416	363	364	265	258	337	327
DEWAR	442	212	121	354	302	362	256
DEWEY	1168	300	215	156	60	266	145
DIBBLE	576	256	172	344	327	361	354
DICKSON	1111	108	110	4	6	35	35
DOVER	202	154	262	55	239	198	303
DRUMMOND	304	88	67	199	126	197	164
DRUMRIGHT	676	340	299	232	214	213	213
DUKE	201	193	208	282	302	209	240
DUNCAN	3882	107	99	130	148	178	182
DURANT	3015	18	24	17	21	45	58
EDMOND	16018	60	32	45	22	47	28
EL RENO	2680	82	56	158	134	50	52
ELGIN	1200	45	71	132	102	136	69
ELK CITY	2205	259	77	100	14	99	106
ELMORE CITY-PERNELL	552	283	159	178	106	312	232
EMPIRE	545	186	85	348	328	242	190
ENID	6888	133	119	113	86	128	91
ERICK	269	243	330	337	360	251	325
EUFAULA	1130	43	113	125	165	73	100
FAIRLAND	492	278	300	289	222	341	294
FAIRVIEW	834	29	9	62	17	116	102
FARGO	205	229	313	124	249	104	179
FLETCHER	481	118	30	52	6	158	41
FORT COBB-BROXTON	430	204	270	85	195	52	150
FORT SUPPLY	158	3	1	270	320	60	254

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
FOX	382	13	117	47	140	19	73
FOYIL	563	299	291	244	209	355	347
FREDERICK	1130	265	264	122	192	330	322
FRONTIER	385	226	346	308	363	56	336
FT GIBSON	1887	27	34	43	33	55	48
FT TOWSON	452	16	66	13	88	14	56
GANS	286	364	365	360	359	8	44
GARBER	369	282	278	23	26	66	97
GEARY	415	297	338	363	364	57	208
GLENCOE	358	295	310	322	279	144	105
GLENPOOL	2117	253	226	294	213	310	269
GORE	599	34	96	94	175	295	292
GRANDFIELD	324	345	342	207	183	340	333
GRANITE	305	22	67	128	229	262	311
GROVE	2076	86	75	171	144	92	91
GUTHRIE	3294	102	151	181	179	147	168
GUYMON	2053	93	87	110	99	288	281
HAILEYVILLE	524	303	294	326	317	285	260
HAMMON	265	179	320	365	365	29	267
HARRAH	2235	125	95	123	96	86	76
HARTSHORNE	830	28	84	306	297	129	173
HASKELL	949	344	311	358	353	343	335
HAWORTH	614	50	111	311	324	331	317
HEALDTON	682	231	159	214	161	248	168
HEAVENER	978	173	194	193	242	81	175
HENNESSEY	818	254	228	53	160	233	208
HENRYETTA	1236	156	128	61	40	111	71
HILLDALE	1561	150	77	69	28	84	41
HINTON	602	19	19	35	29	11	14
HOBART	924	238	247	108	169	254	249
HOLDENVILLE	1235	6	43	66	151	78	156
HOLLIS	744	40	26	5	8	5	1
HOMINY	802	292	273	341	325	344	318
HOOKER	562	151	233	217	260	160	218
HUGO	1604	122	144	169	184	69	96
HULBERT	542	145	103	21	12	211	193
HYDRO	360	169	100	266	224	332	269
IDABEL	1703	342	327	353	342	216	178
INDIAHOMA	230	264	211	257	285	358	362
INDIANOLA	451	78	183	58	90	253	227
INOLA	1184	178	97	221	145	220	136
JAY	1716	31	65	83	119	153	149
JENKS	8812	106	10	157	11	187	31
JONES	1109	49	37	51	49	97	60
KANSAS	636	225	231	262	220	181	195

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
KELLYVILLE	1208	187	100	338	308	335	287
KEOTA	506	279	230	14	39	67	98
KETCHUM	593	330	312	226	193	276	262
KIEFER	454	356	256	347	195	214	41
KINGFISHER	1247	139	82	92	25	139	60
KINGSTON	913	343	344	350	354	351	339
KIOWA	362	176	320	267	307	61	98
KONAWA	783	167	221	184	237	304	305
KREMLIN-HILLSDALE	270	339	345	227	289	228	265
LATTA	637	21	27	48	55	10	1
LAVERNE	445	61	261	220	295	124	265
LAWTON	18298	183	204	164	198	176	223
LE FLORE	288	221	175	229	217	286	158
LEEDEY	209	119	335	195	338	244	358
LEXINGTON	922	314	247	243	172	305	225
LIBERTY	536	288	256	260	217	319	297
LINDSAY	1084	30	1	82	24	201	210
LITTLE AXE	1392	7	18	72	93	33	50
LOCUST GROVE	1411	152	60	107	52	267	197
LOMEGA	178	98	269	173	244	26	100
LONE GROVE	1381	162	91	172	82	115	81
LONE WOLF	219	20	61	6	48	28	46
LUTHER	767	39	44	240	203	323	283
MADILL	1227	346	305	285	208	318	277
MANGUM	721	222	117	284	264	348	349
MANNFORD	1471	203	135	200	113	199	124
MARIETTA	909	316	188	241	173	277	124
MARLOW	1429	163	147	97	64	94	69
MAUD	430	355	350	80	42	329	299
MAYSVILLE	475	161	151	263	235	284	301
MCALESTER	2868	72	200	95	214	68	199
MCCURTAIN	242	41	25	127	135	175	173
MCLOUD	1750	168	91	293	209	169	120
MEDFORD	329	323	307	143	180	210	329
MEEKER	885	273	219	231	186	237	132
MERRITT	457	9	16	34	94	12	47
MIAMI	2516	90	50	147	113	173	132
MILBURN	272	216	175	33	78	236	229
MILLWOOD	1068	131	1	140	1	245	1
MINCO	529	104	40	189	121	224	160
MOORE	18082	236	131	192	94	183	76
MOORELAND	470	196	211	313	328	156	271
MORRIS	1036	148	80	315	253	339	284
MORRISON	454	26	103	28	65	179	177
MOSS	256	313	328	281	273	250	252

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**
MOUNDS	706	170	131	234	148	265	205
MOUNTAIN VIEW-GOTEBO	370	132	289	141	312	188	336
MULDROW	1492	123	129	146	141	133	112
MULHALL-ORLANDO	244	89	234	31	246	168	293
MUSKOGEE	6782	190	186	126	180	117	155
MUSTANG	6309	55	15	111	38	85	22
MWC/DEL CITY	15399	111	98	154	121	126	130
NAVAJO	576	56	34	98	61	227	170
NEW LIMA	275	112	197	258	262	206	207
NEWCASTLE	1093	244	17	145	4	150	17
NEWKIRK	732	362	355	317	269	342	311
NINNEKAH	549	291	240	246	229	363	361
NOBLE	2608	258	199	298	249	280	184
NORMAN	12492	83	67	91	86	113	83
NOWATA	1042	357	340	300	229	289	186
OAKS-MISSION	388	12	1	50	17	38	12
OILTON	326	266	283	42	43	41	36
OKARCHE	308	268	238	103	98	13	10
OKAY	487	1	1	1	1	1	1
OKEENE	371	158	181	177	216	272	319
OKEMAH	977	263	158	190	180	258	235
OKLA CITY	38543	80	52	86	36	101	30
OKMULGEE	2302	270	286	302	281	299	285
OKTAHA	614	44	204	159	252	121	240
OLIVE	427	280	308	325	332	205	170
OOLOGAH-TALALA	1469	144	113	167	124	191	126
OWASSO	5878	48	29	63	29	63	23
PADEN	275	305	247	342	316	162	132
PANAMA	675	38	177	10	19	25	38
PAOLI	258	87	103	223	190	49	52
PAULS VALLEY	1341	246	239	303	310	279	288
PAWHUSKA	1112	207	153	255	237	333	332
PAWNEE	870	141	116	288	268	167	141
PERKINS-TRYON	1160	159	89	292	249	100	74
PICHER-CARDIN	456	333	334	346	345	320	305
PIEDMONT	1279	15	1	24	5	77	18
PIONEER-PLEASANT VALE	576	35	57	247	187	31	28
PLAINVIEW	1286	58	154	39	45	98	127
POCOLA	856	233	173	176	152	226	201
PONCA CITY	5568	84	77	144	111	138	106
POND CREEK-HUNTER	363	64	106	149	168	230	331
PORTER CONSOLIDATED	475	121	190	138	202	164	245
PORUM	494	320	300	356	350	235	227
POTEAU	1963	224	115	209	116	303	172
PRAGUE	990	174	144	38	97	51	49

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**
PRESTON	457	54	38	309	272	252	127
PRYOR	2359	198	74	119	35	148	84
PURCELL	1321	284	139	329	240	186	87
PUTNAM CITY	18938	92	48	114	71	125	62
QUAPAW	562	147	182	216	200	282	279
QUINTON	497	359	360	357	357	275	254
RATTAN	499	75	208	268	285	123	240
RED OAK	257	334	343	366	362	364	364
RINGLING	528	350	352	301	292	325	342
RINGWOOD	338	126	131	134	132	44	67
RIPLEY	504	188	215	19	117	146	233
ROCK CREEK	545	335	338	271	261	274	295
ROFF	331	211	254	238	243	338	340
ROLAND	1222	79	54	205	129	119	86
RUSH SPRINGS	598	25	23	68	75	27	21
RYAN	275	361	361	9	67	72	166
SALINA	823	322	331	330	343	269	273
SALLISAW	1966	301	297	275	262	294	303
SAND SPRINGS	5324	66	67	151	141	127	131
SAPULPA	4145	124	131	133	106	155	122
SAVANNA	520	115	222	88	100	152	211
SAYRE	752	113	75	131	83	40	16
SEILING	445	166	171	297	296	291	324
SEMINOLE	1467	219	191	327	288	142	137
SENTINEL	390	91	208	56	171	88	204
SEQUOYAH	1171	142	53	71	34	200	110
SHARON-MUTUAL	233	146	335	136	304	137	296
SHATTUCK	261	101	228	278	347	298	360
SHAWNEE	3833	289	291	174	211	218	250
SILO	543	5	22	27	88	7	34
SKIATOOK	2008	217	157	185	125	193	187
SMITHVILLE	307	36	1	49	126	79	139
SNYDER	558	329	331	152	254	76	230
SOPER	267	100	240	15	131	48	150
SPERRY	1116	73	33	40	27	59	33
SPIRO	1341	240	167	201	83	249	111
STERLING	365	214	180	137	85	105	94
STIGLER	1192	230	193	112	123	215	233
STILLWATER	5537	95	57	73	49	118	76
STILWELL	1537	202	196	25	70	6	8
STRATFORD	583	239	217	283	275	190	222
STRINGTOWN	243	24	93	74	143	203	108
STROTHER	402	358	356	101	163	32	117
STROUD	818	318	148	239	147	306	159
STUART	257	351	356	248	299	336	349

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
SULPHUR	1408	206	183	102	175	149	160
TAHLEQUAH	3431	62	63	150	135	42	54
TALIHINA	663	227	222	196	100	196	118
TALOGA	193	315	362	179	349	172	348
TECUMSEH	2135	307	271	304	254	301	245
TEMPLE	283	352	347	331	328	80	195
THACKERVILLE	275	32	40	318	273	359	356
THOMAS-FAY-CUSTER UNIFIED DIST	507	252	322	280	321	165	309
TIMBERLAKE	383	65	252	188	267	91	253
TIPTON	413	180	197	335	352	353	359
TISHOMINGO	966	213	54	120	49	107	26
TONKAWA	747	155	135	18	14	17	23
TULSA	41326	205	245	213	224	212	205
TUPELO	264	272	297	93	111	271	297
TURNER	340	302	351	261	344	122	250
TURPIN	508	286	316	256	289	154	236
TUSHKA	344	59	139	212	190	326	333
TUTTLE	1193	17	11	46	16	23	13
UNION	11927	68	31	99	63	108	51
UNION CITY	313	128	13	362	345	290	213
VALLIANT	1005	52	123	11	56	3	11
VANOSS	513	4	19	3	10	39	68
VARNUM	287	97	138	32	23	221	244
VELMA-ALMA	626	189	124	250	240	264	275
VERDEN	324	197	207	60	69	297	315
VIAN	871	360	358	332	312	143	118
VICI	331	218	317	336	356	321	351
VINITA	1572	130	100	287	219	208	166
WAGONER	2308	223	169	65	58	195	140
WAKITA	191	182	286	251	334	145	277
WALTERS	721	242	156	299	227	170	104
WAPANUCKA	207	8	47	2	1	106	193
WARNER	800	157	213	182	233	114	187
WASHINGTON	644	248	93	269	138	241	112
WATONGA	1005	194	150	70	90	163	189
WATTS	374	306	245	163	113	37	1
WAUKOMIS	439	11	45	44	29	75	55
WAURIKA	512	185	143	37	36	18	19
WAYNE	456	81	108	7	9	65	93
WAYNOKA	293	53	264	187	312	46	165
WEATHERFORD	1999	184	163	135	119	109	71
WELCH	350	47	135	57	156	21	219
WELEETKA	455	140	204	290	285	360	355
WELLSTON	694	235	237	202	175	239	192

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**
WESTERN HEIGHTS	3020	267	275	254	259	300	305
WESTVILLE	998	317	304	349	341	327	314
WETUMKA	476	42	129	64	150	16	81
WEWOKA	858	67	46	307	254	347	310
WILBURTON	1090	195	144	264	200	157	115
WILSON1	501	332	200	81	71	141	150
WILSON2	355	353	314	334	351	132	239
WISTER	438	261	254	116	129	263	275
WOODLAND	576	251	293	242	276	192	256
WRIGHT CITY	476	10	12	67	128	225	259
WYANDOTTE	688	293	288	26	73	222	240
WYNNEWOOD	862	237	214	351	328	166	148

* ADM is the average ADM for the cross-sections

APPENDIX IB

Ranking of School Districts by their Efficiency Estimates Generated by the Multiple Output Model, Using Panel Data

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
ACHILLE	495	110	1	188	1	190	147
ADA	2780	15	62	10	1	25	85
ADAIR	908	61	1	43	42	15	1
0FTON	446	227	102	249	185	43	50
ALEX	369	306	307	277	240	325	268
ALINE-CLEO	217	168	281	17	175	122	244
ALLEN	413	156	72	58	120	262	212
ALTUS	4703	54	167	104	215	113	249
ALVA	1067	72	246	70	237	155	300
AMBER-POCASSET	452	106	121	55	173	152	169
ANADARKO	2032	276	303	323	334	241	222
ANTLERS	1154	82	1	137	103	48	24
ARAPAHO	294	19	65	9	100	5	74
ARDMORE	3414	120	259	145	211	165	254
ARKOMA	476	343	315	335	235	349	327
ARNETT	186	58	347	126	358	129	358
ASHER	227	249	115	40	201	3	1
ATOKA	937	254	230	147	185	297	292
BALKO	158	296	365	316	365	265	353
BARNSDALL	479	310	270	266	149	339	180
BARTLESVILLE	6456	17	88	22	72	28	100
BATTIEST	351	331	331	317	335	318	345
BEAVER	413	295	340	294	349	283	339
BEGGS	966	328	296	278	287	183	180
BENNINGTON	256	37	181	100	200	138	279
BERRYHILL	1002	265	149	124	91	157	32
BETHANY	993	164	117	291	143	247	193
BETHEL	1027	203	165	253	226	128	129
BIG PASTURE	266	329	250	305	292	315	315
BILLINGS	174	314	307	68	264	306	351
BINGER-ONEY	366	80	97	91	257	23	76
BIXBY	3113	126	75	235	135	144	76
BLACKWELL	1713	201	157	255	208	185	80
BLAIR	372	261	56	300	140	309	202
BLANCHARD	1197	244	162	243	170	312	212
BLUEJACKET	257	301	248	269	84	207	28
BOISE CITY	415	79	223	261	320	171	311
BOKOSHE	285	13	1	82	1	9	1
BOONE-APACHE	682	98	95	175	213	263	289
BOSWELL	443	209	342	285	327	90	166

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**
BRAGGS	257	256	187	282	54	59	1
BRAMAN	160	342	47	363	360	361	327
BRAY-DOYLE	441	131	185	227	257	96	199
BRIDGE CREEK	1031	213	82	268	125	256	123
BRISTOW	1616	152	178	136	215	156	234
BROKEN ARROW	14499	22	130	16	76	31	156
BROKEN BOW	1792	280	321	258	285	197	277
BUFFALO	354	226	341	232	355	174	350
BUFFALO VALLEY	242	196	105	342	317	355	228
BURNS FLAT-DILL CITY	685	178	220	199	255	303	315
BUTNER	288	112	115	226	177	37	1
BYNG	1692	64	91	29	129	140	217
CACHE	1263	199	236	184	195	257	260
CADDO	410	24	1	93	71	161	66
CALERA	564	95	49	62	33	49	67
CAMERON	503	337	282	364	363	246	53
CANADIAN	403	338	319	229	223	227	282
CANEY VALLEY	810	271	159	267	66	302	85
CANTON	480	163	311	322	345	344	347
CANUTE	267	175	121	142	1	286	225
CARNEGIE	770	128	71	164	189	151	96
CARNEY	265	356	322	264	232	352	234
CASHION	423	151	109	296	220	292	43
CATOOSA	2362	132	59	242	179	218	190
CEMENT	275	349	262	351	340	230	280
CENTRAL	448	143	134	79	1	86	33
CENTRAL HIGH	353	305	261	262	178	281	158
CHANDLER	1140	148	124	90	103	105	79
CHATTANOOGA	294	274	313	248	330	305	305
CHECOTAH	1628	260	267	225	230	87	163
CHELSEA	1044	289	238	325	307	224	188
CHEROKEE	405	217	58	141	66	181	1
CHEYENNE	282	105	366	118	366	196	366
CHICKASHA	2980	162	181	163	180	253	270
CHISHOLM	937	191	175	247	167	213	119
CHOCTAW/NICOMA PARK	4627	71	93	167	156	85	119
CHOUTEAU-MAZIE	936	315	253	201	183	242	139
CIMARRON	368	221	287	198	251	248	254
CLAREMORE	3632	62	95	54	85	62	123
CLAYTON	398	323	348	120	248	288	324
CLEVELAND	1674	215	205	244	230	269	169
CLINTON	2036	234	200	195	101	232	184
COALGATE	686	205	265	84	124	42	41

District	ADM*	1996-1997 Rank SFR	1996-1997 Rank DEA	1997-1998 Rank SFR	1997-1998 Rank DEA	1998-1999 Rank SFR	1998-1999 Rank DEA
COLBERT	798	308	334	170	160	255	151
COLCORD	693	195	216	299	286	220	48
COLEMAN	179	30	270	161	284	223	285
COLLINSVILLE	1593	230	193	143	118	101	53
COMANCHE	1036	299	245	298	223	267	241
COMMERCE	811	140	140	158	73	26	1
COPAN	452	287	159	346	251	331	241
CORDELL	693	45	1	186	125	158	123
COVINGTON-DOUGLAS	335	285	336	250	282	316	338
COWETA	2454	192	188	130	105	150	108
COYLE	377	206	238	287	336	12	91
CRESCENT	640	251	190	237	220	271	225
CROOKED OAK	820	360	329	340	306	329	153
CROWDER	513	245	87	239	283	136	97
CUSHING	2029	283	360	228	353	209	360
CYRIL	457	243	72	181	150	252	247
DALE	624	272	213	172	152	324	261
DAVENPORT	455	208	78	166	127	153	174
DAVIS	926	133	139	25	1	66	70
DEER CREEK	1343	44	47	30	1	77	59
DEER CREEK-LAMONT	258	200	316	309	315	310	331
DEPEW	416	365	332	365	349	360	174
DEWAR	442	303	228	359	294	357	1
DEWEY	1168	262	155	273	161	279	143
DIBBLE	576	290	200	362	348	362	299
DICKSON	1111	233	226	56	1	215	147
DOVER	202	34	175	75	301	159	317
DRUMMOND	304	186	127	106	89	168	248
DRUMRIGHT	676	350	297	260	257	272	272
DUKE	201	115	82	156	206	254	332
DUNCAN	3882	32	88	44	170	106	193
DURANT	3015	11	69	15	66	21	131
EDMOND	16018	1	62	1	38	1	44
EL RENO	2680	135	141	144	195	135	112
ELGIN	1200	122	164	154	143	177	64
ELK CITY	2205	145	1	69	1	84	190
ELMORE CITY-PERNELL	552	292	162	293	251	341	303
EMPIRE	545	325	197	358	278	199	81
ENID	6888	38	188	19	111	30	135
ERICK	269	69	266	224	264	146	298
EUFAULA	1130	90	147	138	48	56	78
FAIRLAND	492	185	256	193	111	91	1

District	ADM*	1996-1997 Rank SFR	1996-1997 Rank DEA	1997-1998 Rank SFR	1997-1998 Rank DEA	1998-1999 Rank SFR	1998-1999 Rank DEA
FAIRVIEW	834	57	1	47	1	41	94
FARGO	205	149	301	279	311	92	70
FLETCHER	481	264	1	171	1	173	39
FORT COBB-BROXTON	430	63	152	24	156	22	135
FORT SUPPLY	158	3	1	295	328	226	311
FOX	382	183	200	327	291	244	207
FOYIL	563	324	227	314	246	337	166
FREDERICK	1130	273	302	265	267	322	295
FRONTIER	385	77	355	67	362	95	365
FT GIBSON	1887	92	137	146	185	73	139
FT TOWSON	452	137	159	165	140	24	36
GANS	286	366	364	356	313	178	163
GARBER	369	320	299	94	58	107	145
GEARY	415	330	361	286	322	11	195
GLENCOE	358	317	291	271	120	118	1
GLENPOOL	2117	235	214	173	123	238	231
GORE	599	166	97	160	88	231	225
GRANDFIELD	324	291	184	207	114	251	284
GRANITE	305	23	85	307	329	205	174
GROVE	2076	85	105	81	129	78	158
GUTHRIE	3294	160	250	215	248	261	289
GUYMON	2053	83	105	83	82	187	210
HAILEYVILLE	524	345	335	360	342	290	215
HAMMON	265	68	339	324	336	76	357
HARRAH	2235	138	147	202	208	166	185
HARTSHORNE	830	150	250	301	307	311	325
HASKELL	949	353	336	349	326	351	352
HAWORTH	614	304	248	361	324	343	334
HEALDTON	682	197	93	194	180	110	72
HEAVENER	978	248	294	218	294	109	272
HENNESSEY	818	239	207	128	201	97	53
HENRYETTA	1236	97	86	89	80	170	109
HILLDALE	1561	88	1	125	114	125	91
HINTON	602	167	84	109	1	58	27
HOBART	924	93	109	48	150	149	123
HOLDENVILLE	1235	51	52	111	174	100	156
HOLLIS	744	118	113	71	63	89	1
HOMINY	802	326	256	347	301	317	220
HOOKER	562	210	312	230	293	164	197
HUGO	1604	43	1	49	117	50	106
HULBERT	542	136	119	36	1	132	119
HYDRO	360	56	62	302	180	333	258
IDABEL	1703	313	328	275	164	235	215
INDIAHOMA	230	344	310	319	305	358	347

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**
INDIANOLA	451	180	223	213	58	314	199
INOLA	1184	236	126	252	192	260	199
JAY	1716	26	1	78	127	104	154
JENKS	8812	27	1	41	51	35	33
JONES	1109	229	171	211	147	141	117
KANSAS	636	174	211	204	195	124	254
KELLYVILLE	1208	146	1	303	204	301	229
KEOTA	506	352	274	318	210	335	277
KETCHUM	593	70	1	13	1	40	21
KIEFER	454	364	292	366	300	346	207
KINGFISHER	1247	91	90	77	35	88	69
KINGSTON	913	334	343	337	343	334	293
KIOWA	362	127	345	168	261	39	203
KONAWA	783	86	142	122	248	126	182
KREMLIN-HILLSDALE	270	267	326	272	315	239	288
LATTA	637	25	1	11	32	7	1
LAVERNE	445	81	346	85	270	27	135
LAWTON	18298	5	144	2	129	8	185
LE FLORE	288	258	211	344	346	330	285
LEEDEY	209	172	356	133	359	68	361
LEXINGTON	922	340	294	292	238	323	280
LIBERTY	536	129	72	169	156	212	160
LINDSAY	1084	134	1	113	1	275	300
LITTLE AXE	1392	107	1	231	175	264	127
LOCUST GROVE	1411	327	214	308	194	356	334
LOMEGA	178	29	303	37	235	147	233
LONE GROVE	1381	46	1	102	57	102	119
LONE WOLF	219	161	220	99	167	233	195
LUTHER	767	119	69	336	290	353	311
MADILL	1227	293	230	155	118	219	209
MANGUM	721	253	135	256	276	289	340
MANNFORD	1471	111	1	178	94	175	117
MARIETTA	909	187	78	179	137	273	187
MARLOW	1429	255	285	206	191	182	116
MAUD	430	363	336	297	37	366	362
MAYSVILLE	475	302	292	209	133	195	203
MCALESTER	2868	65	237	66	225	82	241
MCCURTAIN	242	257	178	223	275	282	272
MCLOUD	1750	278	275	304	257	201	203
MEDFORD	329	266	324	151	294	163	341
MEEKER	885	218	194	281	276	259	234
MERRITT	457	47	1	236	206	162	109
MIAMI	2516	33	52	59	99	81	138
MILBURN	272	335	314	289	273	300	249
MILLWOOD	1068	297	1	332	1	321	1

District	ADM*	1996-1997 Rank SFR	1996-1997 Rank DEA	1997-1998 Rank SFR	1997-1998 Rank DEA	1998-1999 Rank SFR	1998-1999 Rank DEA
MINCO	529	154	67	153	137	123	83
MOORE	18082	10	97	7	82	16	169
MOORELAND	470	220	256	312	274	245	160
MORRIS	1036	171	97	92	1	221	139
MORRISON	454	52	1	27	63	120	114
MOSS	256	361	356	326	247	193	244
MOUNDS	706	241	175	251	184	308	249
MOUNTAIN VIEW-GOTEBO	370	18	157	50	318	52	320
MULDROW	1492	104	137	101	129	29	39
MULHALL-ORLANDO	244	55	200	74	338	72	263
MUSKOGEE	6782	53	204	51	242	63	262
MUSTANG	6309	21	1	28	73	32	60
MWC/DEL CITY	15399	8	197	14	201	14	229
NAVAJO	576	76	1	98	45	148	100
NEW LIMA	275	121	267	192	93	94	53
NEWCASTLE	1093	246	1	157	1	216	46
NEWKIRK	732	336	255	270	146	307	173
NINNEKAH	549	346	327	328	294	364	321
NOBLE	2608	311	285	274	211	249	147
NORMAN	12492	6	133	5	105	4	100
NOWATA	1042	351	283	210	152	268	139
OAKS-MISSION	388	36	1	116	1	169	42
OILTON	326	355	287	73	1	160	26
OKARCHE	308	103	109	149	137	53	23
OKAY	487	232	1	341	1	270	1
OKEENE	371	116	243	105	270	145	332
OKEMAH	977	179	75	162	205	258	263
OKLA CITY	38543	189	78	182	62	274	63
OKMULGEE	2302	288	324	331	333	345	354
OKTAHA	614	259	317	284	331	117	310
OLIVE	427	268	299	311	298	250	217
OOLOGAH-TALALA	1469	74	1	52	49	70	53
OWASSO	5878	31	77	31	109	54	103
PADEN	275	212	91	222	50	38	1
PANAMA	675	159	259	18	41	67	38
PAOLI	258	96	109	63	75	18	28
PAULS VALLEY	1341	319	344	214	311	338	343
PAWHUSKA	1112	158	152	97	156	206	239
PAWNEE	870	67	1	64	1	33	1
PERKINS-TRYON	1160	224	190	212	170	133	112
PICHER-CARDIN	456	358	332	355	314	347	326
PIEDMONT	1279	49	1	72	42	137	50
PIONEER-PLEASANT VALE	576	109	149	114	80	74	61

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
PLAINVIEW	1286	59	207	20	56	19	61
POCOLA	856	294	190	205	215	214	151
PONCA CITY	5568	40	155	21	95	34	87
POND CREEK-HUNTER	363	182	219	134	164	208	342
PORTER CONSOLIDATED	475	237	270	220	244	340	318
PORUM	494	102	60	219	220	328	220
POTEAU	1963	130	152	88	111	108	68
PRAGUE	990	117	117	129	195	191	143
PRESTON	457	66	1	320	280	114	37
PRYOR	2359	73	51	42	1	46	89
PURCELL	1321	142	1	216	76	204	182
PUTNAM CITY	18938	7	124	3	90	6	115
QUAPAW	562	147	216	329	304	287	263
QUINTON	497	359	354	353	357	350	344
RATTAN	499	188	307	241	320	112	308
RED OAK	257	322	262	338	244	354	266
RINGLING	528	357	359	321	287	203	296
RINGWOOD	338	124	127	65	65	45	132
RIPLEY	504	228	225	35	114	184	222
ROCK CREEK	545	279	270	96	33	198	254
ROFF	331	78	130	185	233	167	106
ROLAND	1222	155	120	217	164	240	212
RUSH SPRINGS	598	41	1	61	45	179	89
RYAN	275	339	280	152	192	189	244
SALINA	823	362	363	257	263	116	145
SALLISAW	1966	214	234	123	152	186	249
SAND SPRINGS	5324	39	171	45	143	55	160
SAPULPA	4145	94	130	110	133	69	129
SAVANNA	520	123	197	112	142	65	83
SAYRE	752	225	238	119	87	99	48
SEILING	445	144	297	148	280	237	355
SEMINOLE	1467	101	52	233	161	131	166
SENTINEL	390	28	149	57	218	51	239
SEQUOYAH	1171	194	78	139	66	188	105
SHARON-MUTUAL	233	177	350	288	355	194	330
SHATTUCK	261	50	275	290	361	326	364
SHAWNEE	3833	153	284	87	251	143	268
SILO	543	35	1	117	102	36	22
SKIATOOK	2008	173	104	203	189	119	109
SMITHVILLE	307	14	1	121	213	83	188
SNYDER	558	298	246	189	324	228	336
SOPER	267	321	351	159	227	115	276
SPERRY	1116	89	1	140	110	299	169

		1996-1997		1997-1998		1998-1999	
District	ADM*	Rank SFR	Rank DEA	Rank SFR	Rank DEA	Rank SFR	Rank DEA
SPIRO	1341	300	230	221	185	222	104
STERLING	365	216	194	276	167	234	97
STIGLER	1192	190	234	108	95	139	197
STILLWATER	5537	9	1	8	1	13	65
STILWELL	1537	99	165	32	58	71	20
STRATFORD	583	277	207	334	319	304	305
STRINGTOWN	243	348	233	348	309	313	74
STROTHER	402	332	262	131	105	130	50
STROUD	818	269	1	180	51	278	91
STUART	257	219	287	107	255	202	345
SULPHUR	1408	113	171	60	195	134	203
TAHLEQUAH	3431	12	66	12	42	2	19
TALIHINA	663	139	105	150	55	111	30
TALOGA	193	169	356	176	354	121	359
TECUMSEH	2135	238	253	280	278	276	249
TEMPLE	283	309	317	306	267	280	302
THACKERVILLE	275	341	171	357	270	327	234
THOMAS-FAY-CUSTER UNIFIED DIST	507	75	287	234	332	60	296
TIMBERLAKE	383	108	330	187	309	180	314
TIPTON	413	181	220	330	341	363	363
TISHOMINGO	966	240	142	76	39	75	33
TONKAWA	747	176	144	132	51	176	72
TULSA	41326	270	267	283	218	296	217
TUPELO	264	204	144	86	78	200	231
TURNER	340	307	362	313	364	192	303
TURPIN	508	316	349	196	239	277	319
TUSHKA	344	16	60	26	1	20	1
TUTTLE	1193	48	1	33	35	47	31
UNION	11927	2	1	6	78	10	95
UNION CITY	313	170	1	339	229	319	222
VALLIANT	1005	193	241	135	108	210	24
VANOSS	513	4	1	4	1	44	147
VARNUM	287	157	135	190	39	336	294
VELMA-ALMA	626	223	169	254	264	291	321
VERDEN	324	318	169	259	135	236	88
VIAN	871	333	319	238	228	93	81
VICI	331	100	279	177	287	293	349
VINITA	1572	87	55	174	85	127	132
WAGONER	2308	165	127	34	1	79	99
WAKITA	191	247	353	240	347	98	178
WALTERS	721	211	123	208	95	243	154
WAPANUCKA	207	20	113	23	1	285	308
WARNER	800	60	97	39	58	17	58
WASHINGTON	644	250	1	246	91	295	174

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**	**Rank SFR**	**Rank DEA**
WATONGA	1005	42	57	80	155	57	134
WATTS	374	284	49	38	1	80	1
WAUKOMIS	439	114	186	183	120	103	45
WAURIKA	512	198	194	115	95	61	46
WAYNE	456	207	228	127	45	172	128
WAYNOKA	293	231	352	200	351	142	307
WEATHERFORD	1999	84	207	46	148	64	178
WELCH	350	125	275	245	323	211	323
WELEETKA	455	252	306	263	233	359	270
WELLSTON	694	281	275	343	269	332	289
WESTERN HEIGHTS	3020	275	323	197	261	284	336
WESTVILLE	998	222	243	350	344	217	163
WETUMKA	476	286	216	352	241	365	285
WEWOKA	858	282	205	345	298	348	329
WILBURTON	1090	202	168	333	303	342	282
WILSON1	501	354	178	191	70	266	258
WILSON2	355	347	102	354	352	294	266
WISTER	438	263	1	103	1	298	272
WOODLAND	576	141	241	315	339	320	356
WRIGHT CITY	476	184	67	95	163	154	211
WYANDOTTE	688	242	181	53	1	225	190
WYNNEWOOD	862	312	303	310	242	229	234

* ADM is the average ADM for the cross-sections

APPENDIX II

APPENDIX II

Ranking of School Districts by their Efficiency Estimates Generated by the DEA CRS Model I and Model II using Panel Data

		1996-1997		1997-1998		1998-1999	
District	ADM*	Model I	Model II	Model I	Model II	Model I	Model II
ACHILLE	495	285	232	198	188	194	130
ADA	2780	218	140	149	80	168	125
ADAIR	908	60	32	47	18	46	21
AFTON	446	92	204	50	196	1	140
ALEX	369	1	246	1	192	33	266
ALINE-CLEO	217	209	238	122	163	131	198
ALLEN	413	130	127	282	234	291	312
ALTUS	4703	213	133	209	119	261	200
ALVA	1067	257	199	236	159	233	190
AMBER-POCASSET	452	66	94	108	134	139	156
ANADARKO	2032	337	316	290	312	306	301
ANTLERS	1154	228	152	186	138	303	232
ARAPAHO	294	182	160	194	157	103	61
ARDMORE	3414	295	230	264	243	294	265
ARKOMA	476	233	223	274	260	312	289
ARNETT	186	317	334	312	323	279	322
ASHER	227	334	335	129	293	1	44
ATOKA	937	212	188	207	188	274	243
BALKO	158	348	351	352	351	322	350
BARNSDALL	479	141	226	106	146	62	168
BARTLESVILLE	6456	67	44	66	39	139	73
BATTIEST	351	267	332	271	338	234	345
BEAVER	413	251	280	290	308	179	232
BEGGS	966	225	193	266	210	211	152
BENNINGTON	256	236	264	214	238	347	349
BERRYHILL	1002	25	27	1	18	1	11
BETHANY	993	84	84	29	89	1	69
BETHEL	1027	112	68	85	64	92	55
BIG PASTURE	266	51	207	195	235	152	227
BILLINGS	174	121	303	1	296	108	297
BINGER-ONEY	366	325	301	319	272	238	198
BIXBY	3113	18	16	27	22	1	19
BLACKWELL	1713	138	78	138	107	103	110
BLAIR	372	70	63	181	184	81	116
BLANCHARD	1197	74	50	1	27	56	50
BLUEJACKET	257	93	314	1	236	1	94
BOKOSHE	285	70	130	258	262	264	277
BOONE-APACHE	682	134	140	193	204	223	243
BOSWELL	443	345	349	331	338	288	305
BRAGGS	257	60	171	40	125	1	57
BRAY-DOYLE	441	145	199	246	243	171	186
BRIDGE CREEK	1031	1	1	35	20	38	17
BRISTOW	1616	208	140	254	181	205	170
BROKEN ARROW	14499	49	22	39	10	63	26
BROKEN BOW	1792	309	260	326	288	268	243

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Model I**	**Model II**	**Model I**	**Model II**	**Model I**	**Model II**
BUFFALO	354	317	294	350	342	330	319
BURNS FLAT-DILL CITY	685	227	240	242	204	190	187
BUTNER	288	172	252	161	276	59	191
CACHE	1263	209	196	222	155	187	136
CADDO	410	108	130	91	98	165	173
CALERA	564	183	150	139	76	119	77
CAMERON	503	288	284	279	252	184	184
CANADIAN	403	263	280	198	228	292	334
CANEY VALLEY	810	1	75	1	74	1	46
CANTON	480	333	326	274	331	324	333
CANUTE	267	172	137	134	148	225	231
CARNEGIE	770	239	203	295	278	297	290
CARNEY	265	270	280	68	203	56	132
CASHION	423	36	41	38	48	1	1
CATOOSA	2362	80	47	82	71	157	84
CEMENT	275	300	288	324	333	238	285
CENTRAL	448	231	273	143	121	125	116
CENTRAL HIGH	353	102	144	61	57	1	23
CHANDLER	1140	143	115	95	70	100	53
CHATTANOOGA	294	1	284	89	278	78	153
CHECOTAH	1628	320	299	268	246	272	227
CHELSEA	1044	187	157	117	128	131	123
CHEROKEE	405	82	53	70	36	1	13
CHEYENNE	282	353	353	354	353	354	352
CHICKASHA	2980	166	91	173	101	238	169
CHISHOLM	937	36	35	49	32	36	15
CHOCTAW/NICOMA PARK	4627	74	32	76	24	40	12
CHOUTEAU-MAZIE	936	204	151	178	114	244	178
CIMARRON	368	187	214	227	212	164	114
CLAREMORE	3632	143	73	150	60	147	80
CLAYTON	398	338	339	344	344	345	318
CLEVELAND	1674	82	46	150	122	79	76
CLINTON	2036	247	184	212	209	248	237
COALGATE	686	330	294	274	227	205	154
COLBERT	798	329	303	280	232	201	158
COLCORD	693	312	306	333	301	211	178
COLEMAN	179	286	293	298	257	327	328
COLLINSVILLE	1593	162	91	175	94	90	44
COMANCHE	1036	148	107	139	72	155	82
COMMERCE	811	233	259	227	197	182	142
COPAN	452	1	25	46	58	41	81
CORDELL	693	100	53	251	179	225	165
COVINGTON-DOUGLAS	335	87	211	171	272	201	250
COWETA	2454	200	121	103	37	54	24
COYLE	377	255	296	270	328	101	211
CRESCENT	640	113	171	157	159	95	61
CROOKED OAK	820	270	245	329	318	315	298
CROWDER	513	24	76	222	255	68	141
CUSHING	2029	350	343	328	336	297	320
CYRIL	457	170	167	126	119	205	238
DALE	624	33	51	43	43	95	59

		1996-1997		1997-1998		1998-1999	
District	ADM*	Model I	Model II	Model I	Model II	Model I	Model II
DAVENPORT	455	159	152	136	159	174	145
DAVIS	926	148	129	164	104	171	130
DEER CREEK	1343	27	11	1	1	1	1
DEER CREEK-LAMONT	258	128	251	285	281	275	316
DEPEW	416	95	279	154	282	48	204
DEWAR	442	126	146	322	300	343	332
DEWEY	1168	105	58	108	46	97	54
DIBBLE	576	122	211	101	231	97	243
DICKSON	1111	291	235	85	49	246	185
DOVER	202	198	313	96	332	145	325
DRUMMOND	304	129	116	59	74	142	115
DRUMRIGHT	676	177	154	156	171	157	156
DUKE	201	145	216	124	206	230	278
DUNCAN	3882	201	123	189	122	178	113
DURANT	3015	150	117	177	118	237	174
EAGLETOWN	253	323	345	285	343	227	342
EDMOND	16018	21	11	1	1	1	1
EL RENO	2680	172	102	242	169	196	134
ELGIN	1200	162	98	82	41	79	65
ELK CITY	2205	45	19	70	23	221	162
ELMORE CITY-PERNELL	552	113	94	200	150	188	178
EMPIRE	545	36	51	143	165	156	99
ENID	6888	164	89	119	58	121	57
ERICK	269	316	302	314	311	350	346
EUFAULA	1130	301	267	284	243	282	222
FAIRLAND	492	150	234	1	87	1	47
FAIRVIEW	834	1	1	28	9	137	133
FARGO	205	122	213	55	236	73	227
FLETCHER	481	1	20	1	7	1	1
FORT COBB-BROXTON	430	311	286	307	286	318	307
FOX	382	70	125	218	240	162	194
FOYIL	563	60	112	114	165	106	134
FREDERICK	1130	303	257	268	266	311	286
FRONTIER	385	354	354	348	354	297	354
GANS	286	301	331	249	288	145	268
GARBER	369	21	207	1	104	1	106
GEARY	415	351	346	347	335	344	341
GLENCOE	358	131	196	84	104	1	38
GLENPOOL	2117	206	130	53	33	121	102
GORE	599	185	202	189	146	269	243
GRANDFIELD	324	239	273	264	299	300	343
GRANITE	305	136	181	305	315	249	232
GROVE	2076	138	76	186	102	242	181
GUTHRIE	3294	223	176	218	181	213	170
GUYMON	2053	124	70	142	66	196	136
HAILEYVILLE	524	261	235	195	247	230	258
HAMMON	265	349	348	353	352	353	353
HARRAH	2235	88	38	107	72	152	112
HARTSHORNE	830	322	319	346	330	352	337
HAWORTH	614	85	220	222	247	255	293
HEALDTON	682	100	70	171	131	90	43

		1996-1997		1997-1998		1998-1999	
District	ADM*	Model I	Model II	Model I	Model II	Model I	Model II
HEAVENER	978	288	318	267	317	243	296
HENNESSEY	818	160	148	204	171	129	97
HENRYETTA	1236	209	137	182	110	267	217
HILLDALE	1561	77	42	166	80	166	106
HINTON	602	94	83	100	91	74	50
HOBART	924	230	167	236	178	245	205
HOLDENVILLE	1235	108	57	241	197	205	143
HOLLIS	744	192	148	126	134	133	121
HOMINY	802	215	250	169	232	147	170
HOOKER	562	145	227	201	200	150	187
HUGO	1604	299	232	332	292	320	256
HULBERT	542	347	341	288	215	300	243
HYDRO	360	88	88	161	185	34	111
IDABEL	1703	323	309	273	265	289	261
INDIAHOMA	230	27	160	78	228	128	281
INDIANOLA	451	236	258	182	238	86	96
INOLA	1184	53	34	111	50	63	34
JAY	1716	243	190	313	266	266	207
JENKS	8812	1	1	1	1	1	1
JONES	1109	102	66	97	42	108	73
KANSAS	636	325	322	324	284	330	271
KELLYVILLE	1208	78	45	103	140	177	177
KEOTA	506	298	303	231	257	318	327
KETCHUM	593	29	36	33	28	49	41
KIEFER	454	34	137	1	61	1	92
KINGFISHER	1247	154	80	79	29	150	102
KINGSTON	913	292	286	309	321	209	255
KIOWA	362	352	352	321	349	330	329
KONAWA	783	274	229	300	276	312	281
KREMLIN-HILLSDALE	270	236	261	184	251	194	250
LATTA	637	154	94	161	97	72	30
LAVERNE	445	343	340	262	219	181	136
LAWTON	18298	218	140	227	154	236	191
LEEDEY	209	339	336	345	334	315	334
LEXINGTON	922	226	179	175	96	105	63
LIBERTY	536	59	73	69	67	84	89
LINDSAY	1084	1	1	56	25	249	205
LITTLE AXE	1392	64	28	53	107	116	84
LOCUST GROVE	1411	201	123	234	170	294	263
LOMEGA	178	276	325	217	271	171	213
LONE GROVE	1381	179	102	159	79	166	102
LONE WOLF	219	106	227	167	247	161	224
LUTHER	767	45	48	98	112	147	119
MADILL	1227	305	241	231	188	258	211
MANGUM	721	247	173	317	298	338	325
MANNFORD	1471	20	14	31	12	71	36
MARIETTA	909	141	70	188	131	210	151
MARLOW	1429	203	186	139	102	81	39
MAUD	430	131	157	1	221	1	271
MAYSVILLE	475	152	181	91	150	112	84
MCALESTER	2868	305	248	262	240	284	230

District	ADM*	1996-1997 Model I	1996-1997 Model II	1997-1998 Model I	1997-1998 Model II	1998-1999 Model I	1998-1999 Model II
MCCURTAIN	242	274	298	304	325	337	337
MCLOUD	1750	158	100	129	54	88	40
MEDFORD	329	266	267	202	240	285	260
MEEKER	885	166	136	164	177	163	126
MERRITT	457	47	69	129	94	175	149
MIAMI	2516	157	112	212	122	229	164
MIDWEST CITY-DEL CITY	15399	241	163	231	136	219	154
MILBURN	272	264	329	134	263	39	298
MILLWOOD	1068	331	277	299	312	345	348
MINCO	529	58	81	73	61	36	18
MOORE	18082	88	38	61	20	63	27
MOORELAND	470	197	249	207	212	41	158
MORRIS	1036	117	93	75	45	159	108
MORRISON	454	1	85	150	174	67	174
MOSS	256	172	266	133	254	192	268
MOUNDS	706	78	111	70	67	142	91
MT. VIEW-GOTEBO	370	251	253	337	336	323	315
MULDROW	1492	249	175	238	159	200	144
MULHALL-ORLANDO	244	161	205	260	309	154	286
MUSKOGEE	6782	272	210	334	302	336	303
MUSTANG	6309	26	10	1	8	1	1
NAVAJO	576	1	17	1	15	1	10
NEW LIMA	275	340	333	316	305	279	235
NEWCASTLE	1093	1	1	1	1	49	32
NEWKIRK	732	125	121	128	84	139	126
NINNEKAH	549	55	89	248	226	116	160
NOBLE	2608	96	56	56	64	97	75
NORMAN	12492	96	48	73	35	70	35
NOWATA	1042	216	181	256	219	199	194
OAKS-MISSION	388	245	299	301	345	189	201
OILTON	326	65	288	1	188	59	250
OKARCHE	308	42	61	37	34	1	1
OKAY	487	19	13	1	1	44	20
OKEENE	371	221	214	242	217	278	240
OKEMAH	977	246	179	282	230	269	256
OKLAHOMA CITY	38543	293	221	315	291	304	254
OKMULGEE	2302	313	288	309	286	340	305
OKTAHA	614	280	315	323	327	282	270
OLIVE	427	204	241	111	179	75	90
OOLOGAH-TALALA	1469	51	29	63	30	1	13
OWASSO	5878	34	17	36	10	1	1
PADEN	275	207	207	146	171	134	183
PANAMA	675	272	254	214	155	246	196
PAOLI	258	194	223	121	128	34	33
PAULS VALLEY	1341	307	270	272	270	198	238
PAWHUSKA	1112	278	206	293	225	333	286
PERKINS-TRYON	1160	187	112	137	126	134	99
PICHER-CARDIN	456	346	350	318	325	221	290
PIEDMONT	1279	1	1	1	1	1	1
PIONEER-PLEASANT VALE	576	111	126	63	54	93	48
PLAINVIEW	1286	187	156	178	91	112	70

		1996-1997		1997-1998		1998-1999	
District	ADM*	Model I	Model II	Model I	Model II	Model I	Model II
POCOLA	856	280	218	259	181	186	120
PONCA CITY	5568	113	86	114	53	118	72
POND CREEK-HUNTER	363	1	109	1	138	66	274
PORTER CONSOLIDATED	475	282	277	195	224	89	196
PORUM	494	228	218	218	215	294	259
POTEAU	1963	117	59	143	54	125	59
PRAGUE	990	102	109	180	145	84	67
PRESTON	457	99	119	103	91	44	31
PRYOR	2359	91	43	89	30	190	121
PURCELL	1321	86	37	150	80	106	84
PUTNAM CITY	18938	73	31	50	26	101	42
QUAPAW	562	296	323	238	278	307	301
QUINTON	497	294	276	311	272	304	303
RATTAN	499	284	292	290	307	320	316
RED OAK	257	154	222	227	302	308	347
RINGLING	528	313	309	306	288	312	284
RINGWOOD	338	56	64	79	112	87	84
ROCK CREEK	545	117	160	85	78	175	147
ROFF	331	327	306	334	316	326	275
ROLAND	1222	214	135	159	77	213	161
RUSH SPRINGS	598	32	60	98	116	111	93
SALINA	823	342	337	303	261	230	181
SALLISAW	1966	233	163	235	149	217	162
SAND SPRINGS	5324	199	120	125	44	119	56
SAPULPA	4145	117	66	122	67	115	65
SAVANNA	520	180	243	129	185	76	167
SAYRE	752	171	102	218	130	125	63
SEILING	445	319	270	334	320	348	322
SEMINOLE	1467	134	65	204	143	193	129
SENTINEL	390	257	238	301	263	287	235
SEQUOYAH	1171	39	24	30	14	51	28
SHATTUCK	261	257	269	343	348	341	351
SHAWNEE	3833	222	159	225	174	269	219
SILO	543	40	23	169	116	182	126
SKIATOOK	2008	108	61	117	107	108	78
SMITHVILLE	307	341	342	287	350	253	340
SNYDER	558	315	327	342	341	310	313
SOPER	267	310	328	146	202	290	311
SPERRY	1116	40	38	93	86	53	67
STERLING	365	56	99	1	38	1	37
STIGLER	1192	255	190	280	210	261	210
STILLWATER	5537	23	9	41	16	69	29
STILWELL	1537	336	309	289	212	260	221
STRATFORD	583	251	216	252	255	261	225
STRINGTOWN	243	30	167	209	310	159	226
STROTHER	402	177	186	111	193	185	191
STROUD	818	42	25	88	87	46	48
STUART	257	278	273	277	268	249	314
SULPHUR	1408	243	167	293	221	253	189
TAHLEQUAH	3431	231	176	260	195	275	214
TALIHINA	663	327	308	330	319	349	329

		1996-1997		1997-1998		1998-1999	
District	**ADM***	**Model I**	**Model II**	**Model I**	**Model II**	**Model I**	**Model II**
TALOGA	193	334	338	337	340	286	309
TECUMSEH	2135	241	163	226	165	275	216
TEMPLE	283	218	319	1	284	59	253
THACKERVILLE	275	53	190	56	153	234	292
THOMAS-FAY-CUSTER UNIFIED	507	308	270	320	293	339	293
TIMBERLAKE	383	282	319	253	268	264	278
TIPTON	413	261	243	245	283	223	337
TISHOMINGO	966	184	107	238	152	213	147
TONKAWA	747	48	78	41	52	77	105
TULSA	41326	249	184	250	206	220	214
TUPELO	264	264	264	296	272	329	293
TURNER	340	332	330	341	347	300	298
TURPIN	508	195	261	167	199	130	203
TUSHKA	344	195	189	202	193	249	219
TUTTLE	1193	50	21	50	17	56	25
UNION	11927	1	1	45	13	43	16
UNION CITY	313	1	1	60	110	121	139
VALLIANT	1005	216	163	154	61	55	22
VANOSS	513	180	117	209	126	308	240
VARNUM	287	42	173	1	114	1	79
VELMA-ALMA	626	67	100	63	143	204	218
VERDEN	324	31	81	119	174	83	149
VIAN	871	320	309	340	304	351	331
VICI	331	251	280	257	257	334	336
VINITA	1572	152	87	93	99	169	99
WAGONER	2308	267	194	246	158	227	165
WAKITA	191	74	346	189	312	1	240
WALTERS	721	63	53	66	80	112	82
WAPANUCKA	207	166	223	116	141	257	276
WARNER	800	136	133	189	165	203	174
WASHINGTON	644	1	15	1	46	1	97
WATONGA	1005	223	145	278	201	327	261
WATTS	374	126	230	44	206	142	222
WAUKOMIS	439	116	146	33	50	1	50
WAURIKA	512	166	176	77	137	52	145
WAYNE	456	172	254	48	131	138	94
WAYNOKA	293	276	296	351	346	292	278
WEATHERFORD	1999	164	97	174	89	179	109
WELCH	350	192	288	255	306	241	324
WELEETKA	455	344	344	339	324	335	321
WELLSTON	694	133	155	32	99	121	116
WESTERN HEIGHTS	3020	286	235	296	250	317	283
WESTVILLE	998	297	261	184	217	279	264
WETUMKA	476	106	246	101	253	93	271
WEWOKA	858	303	254	348	329	324	310
WILBURTON	1090	267	198	204	185	255	208
WILSON	501	138	127	157	223	213	267
WILSON1	355	185	324	1	293	1	308
WISTER	438	80	106	146	142	259	243
WOODLAND	576	288	317	327	321	342	343

District	ADM*	1996-1997 Model I	1996-1997 Model II	1997-1998 Model I	1997-1998 Model II	1998-1999 Model I	1998-1999 Model II
WOODWARD	2906	69	30	110	40	134	71
WRIGHT CITY	476	98	105	307	297	273	209
WYANDOTTE	688	187	199	79	85	169	123
WYNNEWOOD	862	257	194	216	163	218	201

*ADM is the average ADM for the cross-sections

VITA

Susanne Rassouli-Currier

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE DETERMINANTS OF SCHOOL EFICIENCY IN OKLAHOMA: RESULTS FROM STOCHASTIC PRODUCTION FRONTIER AND DATA ENVELOPMENT ANALYSIS

Major Field: Economics

Biographical:

Education: Graduated from Hadaf High School, Tehran, Iran in May 1975; received Bachelor of Science Degree in Accounting and Finance from the College of Accounting and Finance, National Iranian Oil Company, Tehran, Iran in March 1980 and a Master of Business Administration from the University of Central Oklahoma, Edmond, Oklahoma in May 1994. Completed the requirements for the Doctor of Philosophy degree with a major in Economics at Oklahoma State University in August 2001.

Experience: University of Central Oklahoma, Department of Decision Sciences, Adjunct Faculty, 1994-1995; Department of Economics, Visiting Professor, 1999-2000; Oklahoma State University, Teaching Associate, 1995-present.

Professional memberships: Oklahoma League of Economics, Southern Economic Association.