Efficiency in Public Education – The Role of Socioeconomic Variables

Kalyan Chakraborty

Emporia State University AIS,

Box 4057, 1200 Commercial Street, Emporia, KS 66801, USA.

Tell: 620-341-5913 Fax: 620-341-6346 E-mail: kchakrab@emporia.edu

Abstract

This study measures the efficiency of public education using a stochastic frontier model that estimates an educational production function and an inefficiency effects function that controls the socio-economic and environmental factors simultaneously. The model developed by Battese and Coelli (1995) is used in this study and applied to a panel dataset. The study finds that although the mean efficiency scores obtained from the model are lower than the efficiency scores from a conventional stochastic frontier model, the estimates are robust and consistent. The empirical application uses three-year panel data from Kansas school districts and finds that Kansas schools are generally efficient and most of the educational inputs under the control of the school administration are either have a low or no influence on students' achievement scores. However, students' socioeconomic factors are found to have significant influence on their achievement scores.

1. Introduction

Since the publication of Coleman's report (1966) a considerable amount of resources have been devoted to public education to raise students' achievement scores but with little improvement. In 2003-04 U.S. spent 4.7 percent of the GDP on public elementary and secondary education (OECD countries average spending is 3.9 percent of GDP), but U.S ranked 16 out of 21 countries with respect to graduation rates (Kirsch et al. 2007). In the International Mathematics and Science Study (TIMSS, 2003) U.S. fourth graders rank 12 and 5, and eighth graders rank 15 and 9, in math and science tests respectively, among 46 nations. Despite the nation's high spending in public education ($8,701 per student in 2005), many students will not generate comparable performance. According to NAEP (2007) more than a quarter of eighth graders scored 'below basic' in reading exam, which by the government's definition means that they are not able to "demonstrate a literal understanding of what they read" and "make some interpretations" (Lips, 2006). For math 30 percent of the eighth graders are below basic. Clearly our schools are not delivering the right quality and quantity of education to a large section of our student population.

Although the quality of education is hard to define precisely due to difficulties in measurement, generally it refers to the "knowledge base and analytical skills that are the focal point of schools" (Hanushek and Luque, 2002). Studies estimating educational production function measure quality of education in terms of standardized test scores, a proxy for educational outcome. Falling students' achievement scores in the face of rising educational input qualities and expenditure per student raises public debate on the level of efficiency of our school districts. This study evaluates the relative performance of the Kansas school districts based on how efficiently the resources are being used to provide quality education to Kansans. The provision of elementary and secondary education is largely financed from property taxes which are determined at the local level. Hence, the inefficiency in the provision of education would amount to misuse of tax dollars.

From school year 2000-01 Kansas introduced a new curricular standard and assessment criteria, Kansas Assessment of Modified Measures (KAMM) for its all school districts at various grade levels. This study attempts to explore the behavior of the school districts' inefficiencies after the introduction of the KAMM. This information would be of considerable importance in situations where significant policy changes, reforms, and introduction of new regulations had taken place at some point in time. A relevant question would be: how did the new policy impact students learning generally measured by their success in achievement tests over the years? One of the objectives of this study is to explore the factors that explained the variation in inefficiencies among school districts across time. Specifically, the study investigates why is it so hard to achieve same educational standard for some schools than others? In order to address these questions/issues the study uses a stochastic frontier models for panel data analysis and applied Kansas school districts' data to measure productive inefficiencies. Two stochastic frontier model have been used in this study; the error component model proposed by Battese and Coelli (1992) (henceforth will be called as Model-1), and the 'inefficiency effects model' also proposed by Battese and Coelli (1995)

(henceforth will be called as Model-2). Model-2 is the primary focus but for contrast, Model-1 is also estimated.

The next section provides the background for this study including a brief survey of the studies done in the literature. The third section defines the educational production function and the models followed by a section devoted to the dataset. The fifth section discussed the empirical results. The summary and conclusions are in the concluding section.

2. The Background

One of the common problems in educational production or cost function estimation is: How does one measure educational outputs? Hanushek (1986) defines the purpose of education is to provide the students with the ability "to perform and cope with society after they leave school" and that "adequate measure of innate abilities have never been available." Kane and Staiger (2002) argue that the test scores are unreliable indicators of performance and vary significantly for small samples. In the absence of a perfect measure of output, standardized test scores appear to be the best available indicator of education quality and are related to educational outcome (Hanushek, 2002). Researches on public education find mixed results on the relationship between inputs and student test scores, for example, Hanushek (1986) finds insignificant relationship while Card and Krueger (1996) find school resources have strong and positive influence on students' educational attainment. However, researchers in public education literature consistently find students' demographic characteristics and family background explain their performance in achievement tests better than any other resources devoted to their education (Krueger, 2003; Hanushek, 2003; Rivkin, 1995).

One other issue in education production function literature is how to control for the impact of socioeconomic factors when measuring school districts' inefficiencies. There are several methods both parametric and non-parametric developed over the years to address this issue directly or indirectly. Each of these techniques has its own advantages and disadvantages as discussed in detail by Chakraborty et al. (2001) and Murillo-Zamorano (2004). The most use non-parametric techniques are two-stage DEA and value added DEA; and the most use parametric technique is stochastic frontier production function. (Hoxby, 2000; Grosskopf et al., 1997; 2001; Chakraborty et al., 2001; McCarty and Yaisawrang, 1993; Linna, 1998; Duncombe et al., 1997; Kirjavainen and Loikkanen, 1998; Noulas and Ketkar, 1998; Ruggiero and Vitaliano, 1999; Ruggiero, 2001; Saito and McIntosh, 2003; Bishop and Brand, 2003) The past studies recognize that inaccurate identification of production set, improper measure of inputs and outputs, and failure to account for simultaneity in inputs, outputs, and socioeconomic factors can seriously undermine the accuracy and reliability of efficiency estimates (Klien, 2007). This is because it is most likely that the efficiency and the economic performance of school districts are difficult to accurately assess unless the production function is correctly defined in terms of all measurable inputs and outputs.

Methods for estimating efficiency such as, two-stage DEA, value added DEA, and stochastic frontier models directly or indirectly attempt to control the impact of socioeconomic factors on students achievement scores but they do not account for the simultaneous behavior of inputs, outputs, and socioeconomic factors determining (in)efficiency. In other words, these

models cannot explain the direction and magnitude of the change in school district inefficiency arising out of a change in any of the socioeconomic variables. Contrary to the above, the current study uses a model proposed by Battese and Coelli (1995) called 'inefficiency effects model' that estimated a stochastic production function and an 'inefficiency effects function' simultaneously. In this model conventional educational inputs under the control of the district administration are used in the production function while the socioeconomic variables are used in the inefficiency function. Considering the importance of the socioeconomic factors affecting students' achievement scores efficiency estimates in this study are robust and reliable.

Except the study by Currier (2007) 'inefficiency effects model' has rarely been used in the education production function literature to assess productive efficiency. In her study Currier (2007) estimates technical efficiency for Oklahoma school districts using Battese and Coelli (1995) model but the specification of the 'inefficiency effects function' is incorrect. For example, instead of using socioeconomic variables in the 'inefficiency effects function' the author uses them as inputs in the education production function. The current study contributes to the literature by improving our existing knowledge on how the socioeconomic and other variables affect simultaneously in the measurement of efficiency when 'inefficiency effects model' is applied. The study is important because socioeconomic and environmental factors play a significant role in students' achievement score. The major advantage of this model is its unique ability to estimate the impact of socioeconomic factors and other educational inputs simultaneously using panel data. The major focus in this study is the results obtained from Model-2. However for comparative analysis efficiency estimates from a stochastic frontier model (Model-1) are also reported. These two models are discussed in detail in the next section.

3. Educational Production Function and the Models

The concept of educational production function assumes that the school districts maximize the educational outcome of its students given their budget constraint. However, there is no unique measure of educational outcome which is perfect. Conventionally, it is assumed that for the production of education school districts use inputs that are associated with instructional and non-instructional activities within and outside the control of the school management. School inputs that are associated with achievement scores are generally measured as the student-teacher ratio, educational attainment of the teachers, teaching experience, and various instructional and non-instructional expenditures. Non-school inputs include the socioeconomic status of the students and other environmental factors influencing students' productivity such as, family income, number of parents in the home, parental education, percent of student population belongs to minority, and the percent of students qualified for free and subsidized lunch. The environmental factors are often measured by geographic location (e.g., rural vs. urban), the net assessed value per student, and low English proficiency.

Stochastic production frontier was developed and extended by Aigner, Lovell, and Schmidt (1977), Meeusen and van den Broeck (1977), and Jondrow et al. (1982). The basic idea

behind the stochastic frontier model is that the error term is composed of two parts: (1) the systematic component (i.e., a traditional random error) that captures the effect of measurement error, other statistical noise, and random shocks; and (2) the one-sided component that captures the effects of inefficiency. Several extensions of the stochastic frontier models have been proposed over the years (Battese and Coelli, 1992; 1995; Kumbhakar, 1997; Kumbhakar and Lovell, 2000; Greene, 2001, 2002a, 2002b). In the stochastic frontier model, a nonnegative error term representing technical inefficiency is subtracted from the traditional random error in the classical linear model. The general formulation of the model is:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_k x_{ik} + \varepsilon_i \qquad (1)$$

Where, y_i is output and the x_j are inputs. It is postulated that $\varepsilon_i = v_i - u_i$ where $v_i \sim N(0,\sigma_v^2)$ and $u_i \sim \left| N(0,\sigma_u^2) \right|$, $u_i \geq 0$, and the u_i and v_i are assumed to be independent.

The error term (ε_i) is the difference between the standard white-noise disturbance (v_i), and the one-sided component (u_i). The term v_i allows for randomness across firms and captures the effect of measurement error, other statistical noise, and random shocks outside the firm's control. The component u_i captures the effect of inefficiency (Forsund et al. 1980).

Most of the earlier stochastic production frontier studies estimated only the mean technical inefficiency of firms because the residual for individual observations could not be decomposed into the two components. Jondrow et al. (1982) solved the problem by defining the functional form of the distribution of the one-sided inefficiency component and deriving the conditional distribution of $[u_i|v_i - u_i]$ for two popular distribution cases (i.e., the half normal and exponential) to estimate firm-specific technical inefficiency (Note1).

3.1 Model-1

Following Battese and Coelli (1992) consider a generalize frontier production function for education as:

$$y_{it} = \exp(x_{it}\beta + v_{it} - u_{it}) \qquad (2)$$

where y_{it} denotes the output of the ith school district in the tth time period, x_{it} represents a (1xk) vector of inputs and other explanatory variables for the ith school district at tth time, β is a (kx1) vector of unknown parameters to be estimated, v_{it}s are assumed to be iid$\sim N(0,\sigma_v^2)$ random variables and u_{it}s are non-negative unobserved random variables associated with the technical efficiency of production, where u_{it} is defined as: $u_{it} = \{\exp[-\eta(t-T)]\}u_i$; $i = 1,2,\ldots 283$; $t = 1, 2, \ldots T$. Where $u_{it} \sim N^+(\mu,\sigma^2)$ and η is a parameter to be estimated. As $t \rightarrow T$, $u_{it} \rightarrow u_T$.

Under this specification inefficiencies in periods prior to T depends on the parameter η. If η is positive then $\exp\{-\eta(t-T) = \exp\{\eta(T-t)\}$ is always greater than 1, and increases with the distance of the period t and the last period T. If η is positive, then it implies technical inefficiencies decrease over time and if η is negative, then technical inefficiencies increase over time.(Note 2) The general interest in this model is testing the null hypothesis that the technical inefficiency effects are not present in the model, which is expressed as H_0: γ = 0,

where $\gamma = \dfrac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2}$.

3.2 Model-2

To explore the sources of inefficiency this study used the Battese and Coelli (1995) model where technical inefficiency effects u_{it} in equation (2) is further defined as:

$$u_{it} = z_{it}\delta + c_{it} \qquad\qquad (3)$$

where z_{it} is a (1xM) vector of explanatory variables associated with technical inefficiency effects; δ is a (Mx1) vector of unknown parameters and c_{it} is non negative unobserved random variable obtained by truncation of the $c_{it} \sim N^+(0,\sigma_c^2)$ such that $c_{it} \geq - z_{it}\delta$. This is an alternative specification of u_{it} being a non-negative truncation of the N($z_{it}\delta$, σ^2). Thus the means may be different for different schools/districts and time periods, but the variances are assumed to be the same. For details please see Battese and Broca (1997) and Battese and Coelli (1995).

4. The Dataset

The empirical application of the models uses data from 304 school districts in Kansas. The information on inputs and outputs are obtained from the Kansas State Department of Education (KSDE). Information on standardized test scores in math and reading are collected from the Center for Educational Testing and Evaluation, University of Kansas, Lawrence. Math tests are administered in all districts at the 4th, 7th, and 10th grades, and reading tests at grades 5th, 8th, and 11th. Test score data are available at the school level but the information on most of the other variables used in this study are available only at the district level. For consistency, test score data are aggregated at the district level. The state introduced Kansas Assessment of Modified Measures (KAMM) for all school districts beginning in the 2000-01 school year.

School and non-school inputs used in this study are measured as operating expenditure per-pupil (full-time equivalent, FTE), student-teacher ratio, student-administrative staff ratio, average contract salary for teachers and administrative staff, and district enrollment (FTE). Full time equivalent student (FTE) is based on the percent of time a student is enrolled in grades K-12. The operating expenditure includes expenditure for instruction, administration, and plant maintenance and operation and does not include capital and debt services. The

variables measuring teachers' quality are percent of teachers (FTE) with MA and/or PhD degrees and the percent of teachers (FTE) with ten or above years of teaching experience. The variables for controlling the socioeconomic conditions of the students are: percent of students belonging to minority, percent of students enrolled in a special-ed program, and percent of students qualified for free and subsidized lunches.

The above information was collected for all 304 school districts beginning school year 2002-2003 and until 2004-05. However, due to mergers and consolidations some of the districts do not exist across the entire study period. Further, non-availability of the test score data for all six output measures for some districts limits their inclusion in the dataset. As a result, the number of school districts eligible for inclusion in the dataset is 283. Hence, this study uses a balanced panel data for 283 districts observed over a period of three years (2002-05). The definition and the descriptive statistics for each of these variables are reported on Table 1.

Table 1. Descriptive Statistics and the Definition of Variables used in the Study

(Three-year Average) (Observations, 849 School Districts)

Description of the Variables	Mean	Standard Deviation	Minimum	Maximum
Outputs (Y)				
Seventh grade math score (MATH7)	56.53	7.17	32.80	78.35
Tenth grade math score (MATH10)	50.31	6.31	33.55	73.20
Fifth grade reading score (READ5)	84.08	3.40	68.90	92.70
Eighth grade reading score (READ8)	83.59	2.83	72.40	90.91
Eleventh grade reading score (READ11)	81.12	3.11	64.29	89.80
Variable Inputs (X)				
Student-teacher ratio (STR)	13.28	2.44	6.50	22.40
Student-administrative staff ratio (SAR)	83.16	24.95	21.18	252.00
Teachers with ten or above years of experience (%) (EXPR)	64.21	1.61	28.23	92.80
Teachers with MA and/or PhD degree (%) (MAPH)	34.08	13.08	2.90	76.81
Average contract salary of the teachers ($) (TSAL)	40,675	3,396	30,707	73,749
Average contract salary of the administrative staff ($) (PSAL)	65,065	7,656	36,240	99,974
Socioeconomic/Environmental Factors (Z)				
Students belongs to minority (%) (MIN)	10.97	12.99	0.00	80.77
Students with disabilities (%) (DIS)	14.74	4.19	0.00	56.91
Students with free or reduced lunch (%) (POV)	37.06	12.78	2.14	78.25
Other Variables				
Operating expenditure per student ($) (EXPD)	9,799	1,426	6,551	17,580

Enrollment 7-12 (FTE) (ENROL712)	755	1,827	18	20,467
Enrollment Total (FTE) (ENROLT)	1,635	4,151	123	48,760

5. Empirical Results

The empirical results obtained from Model-1 are based on the estimation of the following equation.

5.1 Model-1

$$Ln(Y_{it}) = \beta_0 + \beta_1 STR + \beta_2 SAR + \beta_3 EXPR + \beta_4 MAPHD + \beta_5 TSAL + \beta_6 PSAL + \beta_7 ENROL$$

$$+ \beta_8 EXPD + \beta_9 MIN + \beta_{10} DIS + \beta_{11} POV + \beta_{12} YR + \varepsilon_{it} \qquad (4)$$

Results from stochastic production frontier model (Model-1, equation 4) and inefficiency effects model (Model-2, equation 5, 6) are reported in Table-2.

Table 2. Estimated Parameters of the Stochastic Production Function and Time Varying Inefficiency Function - Dependent Variable = Ln(Composite Scores) (Obs, 849)

Variable	Model-1 (Coelli-1992)		Model-2 (Coelli-1995)	
	Coefficient	t-statistics	Coefficient	t-statistics
Intercept	5.812	185.71*	5.851	186.529*
STR	-0.0013	-1.078	-0.0013	-1.124
SAR	0.0008	1.184	0.0006	0.986
EXPR (%)	0.0013	0.792	0.0005	0.337
MAPHD (%)	-0.0008	-0.510	-0.0006	-0.430
TSAL ($ '000)	-0.0000	-0.003	0.0002	0.382
PSAL ($ '000)	0.0004	1.453	0.0003	1.351
ENROL712 (FTE, '000)	9.6517	7.219*	8.6278	1.830**
EXPD ($ '000)	-3.8584	-7.217*	-3.4490	-1.828**
YEAR	0.0289	13.928*	0.0116	2.471*
Intercept	-----	-----	-0.0273	-0.938
MIN (%)	-0.0005	-2.759*	0.0008	2.645*
DIS (%)	0.0005	0.113	0.0001	0.000
POV (%)	-0.0006	-3.683*	0.0014	3.576*
YEAR	----	----	-0.0124	-1.084
Sigma squared	0.0066	3.365*	0.0033	66.074*
Gamma	0.7351	7.665*	0.6895	7.820*
Mu	-0.1396	-1.89**	----	----
Eta	-1.4437	-3.459*	----	----
Log Likelihood	1444.70	----	1448.89	----
LR for one-sided error	14.300	----	70.104	----
Degree of Freedom	3	----	6	----

* indicates significant at 5 percent; ** indicates significant at 10 percent

For both models the functional relationship such as, Cobb-Douglas and trans-log were tested but were rejected in favor of semilog model (dependent variable in log form). In a semilog model slope coefficients defines semielasticity. Since the main focus for this study is efficiency, a detailed discussion on the estimated parameters of the production function are not presented. For Model-1 (column 2 and 3) while most of the input variables have expected signs, only six out of thirteen coefficients are statistically significant. The coefficient on the time variable is positive and significant which suggests that output (test scores) increases over time however this does not necessarily imply that the efficiency increases over time. The variance parameters reflect the suitability of stochastic frontier model.

To check if district specific inefficiency exists in the data that necessitates the use of frontier model, the composite hypothesis is tested as, $H_0 : \gamma = \mu = \eta = 0$. The usual likelihood ratio (LR) test reports the test statistics with mixed chi-squared distribution (Battese and Coelli, 1992), and the critical value for a given level of significance is lower than the one reported in the chi-squared table (Kumbhakar, 2005). At five percent level of significance the critical value is 7.81. The value of the test statistics reported in Table 2 is 14.30 which are larger than the critical value hence, the null hypothesis is rejected. It is concluded that the standard average response function is not adequate for analyzing the production behavior of Kansas schools.

The null hypothesis $H_0 : \gamma = 0$, is also rejected because the test statistic is significantly different from zero at five percent level of significance, which implies that the stochastic frontier specification fits the data better than deterministic frontier. The estimated value of the parameter η is negative and significant which implies that the inefficiencies of school districts increase over time. The mean efficiencies decrease marginally, for example, for school year 2003, 2004, and 2005 the means are 99.8, 99.3, and 97.4 percent, respectively. The overall mean efficiency for Model-1 is 98.8 percent (Table 3) implying Kansas school districts are 1.2 percent inefficient in the production of educational output.

5.2 Model-2

$$Ln(Y_{it}) = \beta_0 + \beta_1 STR + \beta_2 SAR + \beta_3 EXPR + \beta_4 MAPHD + \beta_5 TSAL$$

$$+ \beta_6 PSAL + \beta_7 ENROL + \beta_8 EXPD + \beta_9 YR \qquad (5)$$

$$Z_{it} = \delta_0 + \delta_1 MIN + \delta_2 DIS + \delta_3 POV + \delta_4 YR + \varepsilon_{it} \qquad (6)$$

Column 4 and 5 of Table 2 reports the results from Model-2, estimating the equations (5, 6). Similar to Model-1 the likelihood ratio test statistic reconfirms the rejection of $H_0 : \gamma = \delta_0 = \delta_1 = ... = \delta_5 = 0$ (where δ's are the parameters associated with socioeconomic

variables, z's) at five percent level of significance. This implies that the stochastic frontier specification is the appropriate framework for analyzing school district performances. The γ parameter is significantly different from zero implying the exogenous variables in the inefficiency function are able to explain a substantial part of the unconditional variance of the one-sided error term. Out of 15 coefficients in Model-2, 6 are significant either at five or ten percent levels. The coefficient on the expenditure per student variable (EXPD) is -3.45, which implies that an increase of $1,000 expenditure per student would lower the achievement scores by 3.45 percent. The overall results are consistent with the literature, for example, the past studies have found that the school inputs generally have a mixed or no influence on students' achievement scores. For example, Hanushek (2002, 2003) argues that teacher's educational levels and experience (the major criteria for determining teachers' salaries in heavily unionized school teaching in the U.S.) have increased over the years while the student-teacher ratio has decreased resulting in expenditure per student increased although the test scores remained the same.

The coefficient on the *YEAR* variable in the stochastic frontier part of Model-2 (equation 5) is positive and significant which implies that the output increases over time. However, when the 'inefficiency effect function' in Model-2 is considered (equation 6) the coefficient on *YEAR* is negative but insignificant. This implies that the level of inefficiency for Kansas schools tend to decrease (or efficiency tend to increase) over time. The mean efficiency for Kansas schools increases marginally over the study period. For example, for school year 2003, 2004, and 2005 the mean efficiencies from Model-2 are 95.3, 96.0, and 96.1 percent, respectively, and the overall mean efficiency is 95.9 percent (Table 3). Positive and highly significant coefficients on MIN and POV variables in inefficiency effects function (equation 6, Model-2) suggest inefficiency increases due to an increase in these variables. Between the two model results, Model-1 generates higher efficiency scores although it does not incorporate the simultaneous behavior of the socioeconomic factors and the school inputs measuring inefficiency. This study considers efficiency scores obtained from Model-2 are more robust and consistent than Model-1 due to the structure of the model.

In order to explore further the relationship between the efficiency scores and some of the school inputs and socioeconomics variables across district sizes, the raw data on these variables are examined in Table 3. It is hypothesized that the expenditure per student are generally higher for small rural school districts and lower for city/urban districts. For this purpose the districts are grouped into five sizes based on enrollments. It is observed that the efficiency scores from Model-1 are consistently higher for all sizes. It is also revealed that the small (<400) districts have the highest expenditure per student, have favorable school and socioeconomic inputs, but are relatively less efficient. Rose and Barrow (2006) find that smaller class size and higher expenditure does not necessarily imply better quality for those school districts where most of its students come from low income families. Because these schools may have a larger share of special education or English-Language-Learner students than schools serving more privileged children that have fewer special education classrooms.

Table 3. Efficiency Estimates vis-à-vis Selected School and Socioeconomic Factors

Grouped by District Size (2003-2005)

Variables	Small	Medium	Large	X-Large	XX-Large	Sample
	<400	401-1,000	1,001-1,730	1,731-10,000	>10,000	Average
Model-1	0.988	0.990	0.988	0.990	0.990	0.989
Model-2	0.957	0.961	0.961	0.957	0.944	0.959
EXPEND ($)	10,773	9,723	9,089	8,728	9,476	9,799
STR (%)	10.92	13.12	14.70	16.65	16.61	13.28
EXPER (%)	64.48	65.98	65.33	59.51	57.55	64.21
MIN (%)	8.04	8.50	8.49	20.62	36.26	10.97
POV (%)	10.91	35.93	34.03	34.88	36.88	37.07
USD	86	110	35	45	7	283

For Model-2 the XX-Large group with the lowest efficiency level (94.4 percent) has the highest percentage of its students from minority and has an above average student-teacher ratio and fewer experienced teachers. Table 3 demonstrates that variations in educational inputs generally have a less significant effect than variations in socioeconomic factors explaining the differences in efficiency levels across district sizes. These findings are consistent with the current literature. Klein (2007) argues that higher ability and better prepared students appear to score higher on tests and students' demographic characteristics matters most in their success. Dahl and Lochner (2005) find parental socioeconomic status has a causal effect on children's educational outcome. The NAEP (2007) data for Kansas 4th and 8th graders show that for both math and reading exams students who are eligible for national lunch program are fifty percent less 'proficient' than the other students.

The relationship between the raw data on socio-economic variables and the efficiency scores are explored using Table 4. The percentage of students in minority and poverty are divided into three sub-groups – Group-1 (below mean), Group-2 (between the mean and one standard deviation), and Group-3 (mean + one standard deviation). It is interesting to note that for Model-1 efficiency levels do not reflect any significant changes due to variations in POV and MIN variables across district sizes. However, the efficiency scores from Model-2 are consistently higher for the lowest levels of 'minority' and 'poverty' across all district sizes and the vice versa. The efficiency scores dropped sharply when 'minority' and 'poverty' exceed 'one-standard deviation above the mean' limit. It is revealed that the efficiency scores from Model-1 tend to be less responsive to variations in 'minority' and 'poverty,' across all sizes. Although, the use of stochastic frontier model (Model-1) seems to be justified, it is less efficient in controlling the socioeconomic factors while measuring efficiency. Figure 1 and 2 depicts Model-2 efficiency scores against MIN and POV levels across district sizes. Efficiency levels in Figure 1 are high for all districts when MIN is at its lowest level and efficiency levels fall sharply as MIN increase above 25 percent. The impact is highest for the XX-Large districts. Figure 2 shows similar trend for POV levels across district sizes. Efficiency levels decrease as POV increases and the impact is highest for XX-Large districts.

Table 4. Responsiveness of Efficiency Scores from Changes in Minority and Poverty

Sub-groups based on District Size (2003-2005)

	Minority (%)			Poverty (%)		
	0 - 11	12 – 25	> 25	0 - 37	38 - 51	> 51
Model*	Group-1	Group-2	Group-3	Group-1	Group-2	Group-3
Small (< 400)						
Model-1	0.988	0.988	0.984	0.987	0.988	0.989
Model-2	0.961	0.946	0.926	0.964	0.957	0.938
Medium (401 – 1,000)						
Model-1	0.990	0.990	0.989	0.990	0.990	0.991
Model-2	0.963	0.948	0.947	0.967	0.953	0.938
Large (1,001 – 1,730)						
Model-1	0.988	0.989	0.985	0.987	0.991	0.985
Model-2	0.964	0.952	0.937	0.964	0.959	0.939
X-Large (1,731 – 9,999)						
Model-1	0.989	0.988	0.993	0.987	0.993	0.992
Model-2	0.967	0.959	0.942	0.964	0.954	0.941
XX-Large (> 10,000)						
Model-1	0.989	0.991	0.989	0.990	----	0.989
Model-2	0.980	0.974	0.903	0.976	----	0.903

*Efficiency scores using Coelli-1992 (Mod-1); Efficiency scores using Coelli-1995 (Mod-2)

Based on the Model-2 results the overall mean efficiency is 95.9 percent indicating that the educational production could be increased by 4.1 percent if all school districts operate at 100 percent efficiency level. An intuitive interpretation for 4.1 percent overall inefficiency in terms of tax dollars provide would be: average per student expenditure could be reduced by 4.1 percent without reducing the current level of student achievement while all school districts could still remain on the production frontier and be fully efficient. For example, overall average expenditure per student per year for the sample is $9,799 (Table 3) hence the total expenditure could be reduced by $401.76 (= $9799*4.1%) per year per student. For a three-year period and for all 283 school districts with average district size of 1,630 students the amount would be $558 million. It can be inferred that between 2003 and 2005 a total of $558 million misallocation of funds resulted from Kansas school districts not operating on the frontier.
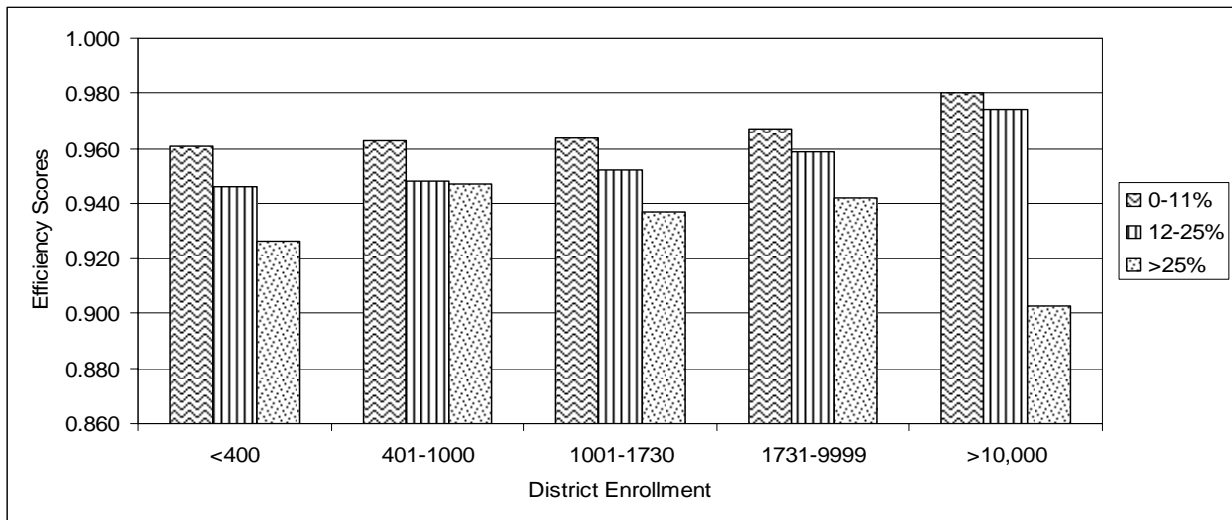
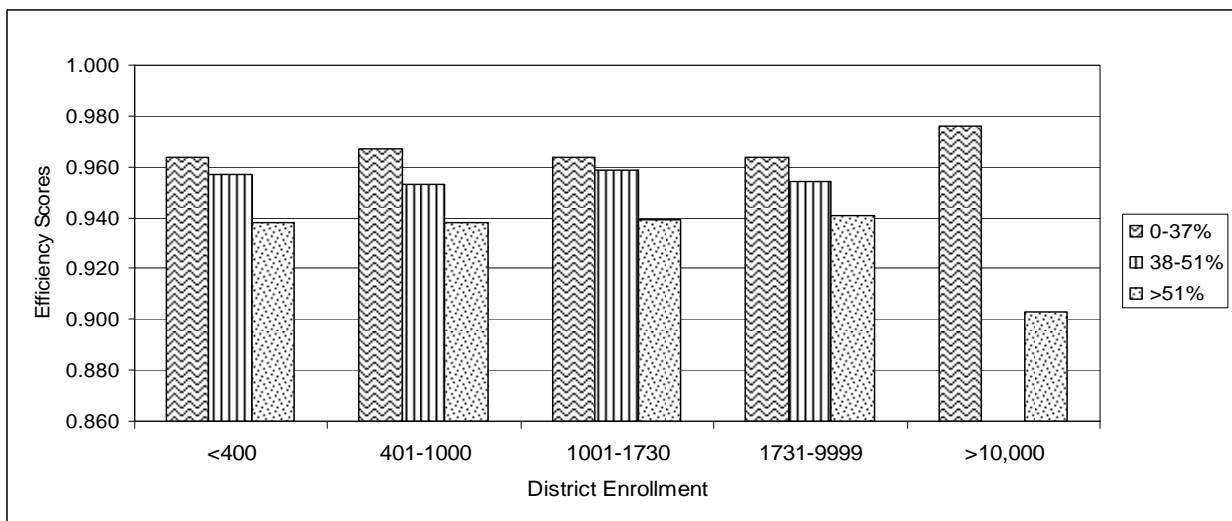Figure 1. Model-2 Efficiency Scores and Students' Minority Levels

Figure 2. Model-2 Efficiency Scores and Students' Poverty Levels

6. Summary and Conclusions

The studies measuring inefficiency in public education consistently found students' socioeconomic factors play a far greater role than school inputs explaining variations in their achievement scores. It is recognized in the literature that any measure of inefficiency that fails to account for the simultaneity in inputs, outputs, and socioeconomic factors would amount to inaccurate and unreliable estimates (Grosskopf et al., 2001; Saito and McIntosh, 2003; Klien, 2007). Unlike the past studies the current study uses an econometric model that specifically addresses the simultaneity issue. This study uses Battese and Coelli (1995) model estimates the parameters of a stochastic production function and an inefficiency effects function simultaneously.

One of the major conclusions drawn from this study is that higher expenditure and lower performance for Kansas schools is similar to the conclusions reached by other researchers from their studies using data from other states including the nation as a whole. Nationwide class sizes became smaller, teachers' educational level has increased, and the expenditure per student has increased without any significant increase in student outcome. With a view to improving the quality of education government policy and proponents of education reform have increased school funding steadily over the years without any evidence that increased resource would improve performance (Hanushek, 2003). It is evident from this study that school inputs such as, student-teacher ratio, teacher's experience and educational background are not significantly related to student test scores (Model-2).

The variables measuring socioeconomic factors in the inefficiency effects model are positive and significant implying increased (decreased) levels for these factors would raise inefficiency (efficiency) for Kansas school districts. The magnitude of the socioeconomic variables increased in general for all Kansas school districts during the study period. The results from this study reported in Table 2 and 3 suggest that the efficiency estimates differ significantly depending upon the model use and how well the model addresses the measurement error, omitted variables, and endogeneity. In order to formulate any policy for education reform based on efficiency analysis it is essential to consider several studies and not just what is observed from a single study.

In the past studies have found lack of competition in the provision of public good is one of the main reasons for its poor performance and quality. If the parents are allowed to choose the school for their children as proposed by some education reformers, that good schools would attract more students and bad schools would lose students. However, there are counter arguments on how would private schools compete against public schools and what would be the impact of racial de-segregation, the problem presumed to have been resolved after years of effort. Most recent cross-country studies on education find that the development of students' cognitive skill or the quality of education is more important for growth of human capital and economic wellbeing of an individual than quantity of schooling (Hanushek and Woessmann, 2008; Woessmann, 2002, 2003). Higher quality schools positively contribute to higher cognitive skills just like non-school factors such as, families, peers, neighborhoods, and health status contribute to student learning (Hanushek, 2008).

Since children's educational achievement depends on so many factors – many of which are beyond the control of the school districts, any policy aimed at achieving higher quality will always have some limits. It is suggested that funds should be directed to do more research on how students learns rather than just spending more money on public education without knowing if additional resource would increase students' performance.

References

Aigner, D. J., C. A. K. Lovell & Schmidt. P. (1977). Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics,* 6, pp.21-27.

Bishop, P., & Brand. S. (2003). The efficiency of museums: A stochastic frontier production function approach, *Applied Economics,* 35, pp.1853-58.

Battese, G. E., & Coelli. T. J. (1995). A Model for technical inefficiency effects in a stochastic

frontier production for panel data, *Empirical Economics*, 24, pp.325-332.

Battese, G. E., & Coelli. T. J. (1992) Frontier production functions, technical efficiency, and panel data with application to paddy farmers in India, *Journal of Productivity Analysis,* 3, pp.153-169.

Battese, G.E.,& Broca, S. S. (1997). Functional forms of stochastic frontier production function and models for technical efficiency effects: A comparative study for wheat farmers in Pakistan, *Journal of Productivity Analysis*, 8, pp.395-414.

Card, D., and A.B. Krueger. (1996) School resources and student outcomes: An overview of the literature and new evidences from North and South Carolina, *Journal of Economic Perspective, 10*(4), 31-50.

Chakraborty, K., B. Biswas, & Lewis, C. W. (2001). Measurement of technical efficiency in public education: A stochastic and nonstochastic production function approach, *Southern Economic Journal, 67*(4), pp.889-905.

Coleman, J.S., E.Q. Cambell, C.J. Hobson, J. McPartland, A.M. Mood, F.D. Weinfield, & York, R. L. 1966. *Equality of Educational Opportunity*, U.S. Government Printing Office, 1966. Washington D.C.

Currier, S. R. (2007). The choice of estimation method and its effect on efficiency measurement on public education: Stochastic frontier analysis vs. data envelopment analysis. *Journal of Economics*, *33*(1), pp.53-73.

Duncombe, W. & Yinger. J. (1997). Why is it so hard to help central city schools? *Journal of Policy Analysis and Management, 16*(1), pp.85-113.

Forsund, E. R., C. A. K. Lovell, & Schmidt, P. 1980. A survey of frontier production functions and their relationship to efficiency measurement. *Journal of Econometrics* 13:5-25. Frontier analysis parametric and nonparametric approaches. Annals 1990-4, A Supplement. 1990. *Journal of Econometrics*. Amsterdam, North Holland.

Greene, W. (2002a) Alternative panel data estimators for the stochastic frontier models,

Department of Economics, Stern School of Business, New York University. Working Paper.

Greene, W. (2002b) Fixed and random effects in stochastic frontier models", Department

of Economics, Stern School of Business, New York University. Working Paper.

Greene, W. (2001) New developments in the estimation of stochastic frontier models with panel data, Paper Presented at the 7th European Workshop on Efficiency and Productivity Analysis, University of Oviedo, Spain. Sept 25.

Grosskopf, S., K.J. Hayes, L.R. Taylor, & Weber, W. L. (2001). On the determinants of school district efficiency: Competition and monitoring, *Journal of Urban Economics*, 49, pp.453-478.

Grosskopf, S., K.J. Hayes, L.R. Taylor, & Weber, w. L. (1997). Budget-constrained frontier measures of fiscal equality and efficiency in schooling, *Review of Economics and Statistics, 79*(1), pp.116-124.

Hanushek, E.A. (2008). The economic value of education and cognitive skills, Book Chapter in Human Capital of Nations, Edited E.A. Hanushek and L. Woessmann, Cahp-3, pp.39-56

Hanushek, E.A., & Woessmann, L. (2008). The role of cognitive skills in economic development, *Journal of Economic Literature*, *46*(3), pp.607-668.

Hanushek, E.A. (2003). The failure of input based schooling policies, *The Economic Journal,* 113, F64-F98.

Hanushek, E.A., & Luque, J. A. (2002). Efficiency and equity in schools around the World, NBER. Working Paper

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public Schools, *Journal of Economic Literature*, 24, pp.1141-77.

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

Hoxby, C.M. (2000) The effect of class size on student achievement: New evidence from population variation, *Quarterly Journal of Economics*, *115*(3), pp.1239-85.

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

Kane, T.J., & Staiger, D. O. (2002). The Promise and pitfalls of using imprecise school

accountability measures, *Journal of Economic Perspective*, *16*(4), pp.91-114.

Kirsch, I., H. Braun, K. Yamamoto, & Sum, A. (2007). America's perfect storm – Three forces changing our nation's future, Educational Testing Services, Report.

Kirjavainen, T., & Loikkanen, H. A. (1998) Efficiency differences of Finnish senior secondary schools: An application of DEA and Tobit analysis, *Economics of Education Review*, *17*(4), pp.377-394.

Klien, C.C. (2007). Efficiency versus effectiveness: Interpreting education production studies, Department of Economics and Finance, Middle Tennessee State University. Working Paper.

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

Krueger, A.B. (2003). Economic considerations and class size, *The Economic Journal*, 113, F34-F63.

Kumbhakar, S.C. & Sarkar, S. (2005). Deregulation, ownership, efficiency change in Indian banking: An application of stochastic frontier analysis, In *Theory and Application of Productivity and Efficiency – Econometric and DEA Approach* (ed) R. Ghosh and C. Neogi. McMillan India Ltd.

Kumbhakar, S.C. & Lovell, C. A. K. (2000). Stochastic Frontier Analysis. Cambridge University Press.

Kumbhakar, S.C. (1997). Modeling Alocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship, *Journal of Econometrics*, 76, pp.351-356.

Linna, M. (1998) Measuring Hospital Cost Efficiency with Panel Data Models, *Health Economics*, **7**, pp.415-427.

Lips, D. 2006. Education Notebook − The Cost of American Education. The Heritage Foundation at [Online] Available: http://www.heritage.org/ accessed September 12, 2007.

Marlow, M.L. (2000) Spending, school structure, and public education quality: Evidence from California, *Economics of Education Review*, 19, pp.89-106.

McCarty, R., & Yaisawarng, S. (1993). Technical efficiency in New Jersey school districts, In Measurement of Productive Efficiency, eds. H. Fried, K. Lovell, and S. Schmidt, Oxford University Press.

Meeusen, W., & J. van den Broeck. (1977). Efficiency estimation from Cobb-Douglas Production Frontiers with Composed Error, *International Economic Review*, 18, pp.435-444.

Murillo-Zamorano, L.R. (2004). Economic efficiency and frontier techniques, *Journal of Economic Survey*, *18*(1), 33-77.

National Assessment of Educational Progress (NAEP). 2007. National Center for Education Statistics at [Online] Available: http://nces.ed.gov/pubsinfo.asp/ accessed September 12, 2007.

Noulas, A.G., & Ketkar, K. W. (1998). Efficient utilization of resources in public schools: A case study of New Jersey, *Applied Economics*, 30, pp.1299-1306.

Rivkin, S.G. (1995). Black/white difference in schooling and employment, *Journal of Human Resource*, *30*(4), pp.826-52.

Rouse, C.E., & Barrow, L. (2006). U.S. Elementary and Secondary schools: Equalizing opportunity or replicating status quo? 16(2), pp.99-123, [Online] Available: www.futureofchildren.org

Ruggiero, J. (2003). Comment on estimating school efficiency, *Economics of Education Review*, 22, pp.631-334.

Ruggiero, J. (2001). Determining the base cost of education: An analysis of Ohio school districts, *Contemporary Economic Policy*, *19*(3), pp.268-279.

Ruggiero, J., & Vitaliano, D. F. (1999). Assessing the efficiency of public schools using data envelopment analysis and frontier regression, *Contemporary Economic Policy*, July, *17*(3), pp.321-331.

Saito, Y., & McIntosh, C. S. (2003). Monitoring inefficiency in public education, *Journal of Agriculture and Applied Economics*, *35*(3), pp.611-623.

Van den Broeck, J., G. Koop, J. Osiewalski, & Steel, M. F. J. (1994). Stochastic frontier models: A Bayesian perspective. *Journal of Econometrics* 61, pp.273-303.

Woessmann, L. (2002). Schooling and quality of human capital, Heidelberg and New York:

Springer.

Woessmann, L. (2003). Specifying human capital, Journal of Economic Surveys, *17*(3): pp.239-70

Notes

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

2. Kumbhakar and Sarkar (2005) explained η as an overall or average trend in inefficiency over the periods of study across all school districts hence, appropriate for modeling aggregate behavior.