MANAGEMENT SCIENCE
Vol. 37, No. 12, December 1991
Printed in U.S.A.

RESOURCE-USE EFFICIENCY IN PUBLIC SCHOOLS: A STUDY OF CONNECTICUT DATA*

SUBHASH C. RAY
Department of Economics, University of Connecticut, Storrs, Connecticut 06268

This study combines Data Envelopment Analysis (DEA) with regression modelling to estimate relative efficiency in the public school districts of Connecticut. Factors affecting achievements are classified as school inputs and other socio-economic factors. DEA is performed with the school inputs only. Efficiency measures obtained from DEA are subsequently related to the socio-economic factors in a regression model with a one-sided disturbance term. The findings suggest that while productivity of school inputs varies considerably across districts this can be ascribed to a large extent to differences in the socio-economic background of the communities served. Variation in managerial efficiency is much less than what is implied by the DEA results.
(DATA ENVELOPMENT ANALYSIS; NONDISCRETIONARY INPUTS; REGRESSION ANALYSIS)

1. Introduction

The method of Data Envelopment Analysis (DEA) was introduced by Charnes, Cooper, and Rhodes (CCR) (1978, 1981) in order to measure relative efficiency of decision making units (hereafter DMUs) functioning in the nonmarket environment. In a previous study using public school data from Texas published in this journal Bessent et al. (1982) used DEA and identified for individual units the levels of potential increase in individual outputs simultaneously with cutbacks in inputs used for different units. Such information can be very valuable for managerial planning.

In the present study we extend the existing literature in two ways. First, we take explicit recognition of the fact that the outputs used in our study are test scores and are bounded from above. It is imperative that the levels of output at the most efficient levels of use of the existing resources implied by DEA should not exceed the prescribed upper limits. We make the necessary modification in the basic linear programming (LP) formulation developed by CCR (Charnes, Cooper and Rhodes 1978, 1981) and also provide an intuitive interpretation of the dual of the revised LP problem.

Secondly, the community socio-economic status variables expected to affect achievement are separated from the school inputs and excluded from the DEA models. Instead, these are used in a subsequent regression analysis with the measure of efficiency obtained from DEA with school inputs alone as the dependent variable. This procedure seems preferable for two reasons. First, it allows one to treat the nonschool factors as fixed and indivisible nondiscretionary inputs and avoids the free disposability assumption with respect to these inputs implicit in the earlier applications. Second, it permits considerable flexibility in the choice of the nonschool inputs to be considered. Because the DEA is based on the outputs and school inputs alone, unlike in the previous applications we do not need to get fresh solutions for all the LP problems every time any nonschool input is included or excluded. This can result in significant computational economy.

Schooling can be considered to have been effective when the pupils are enabled to acquire adequate levels of different cognitive skills. A student's skill levels are typically measured by the individual's achievement scores in various standardized tests. Usually the average test scores of the students in different kinds of skills (like mathematical or

* Accepted by Arnold Barnett, former Departmental Editor; received April 1986. This paper has been with the author 22 months for 5 revisions.

0025-1909/91/3712/1620$01.25

verbal ability) are treated as indices of outputs of a school. On the other hand, efficiency relates to productivity in resource utilization. A school is efficient if it is found to produce the maximum level of student achievements from a given bundle of resources used. A high-achievement school may be effective but inefficient if it is using more of the school inputs (like teachers or physical facilities) than is necessary to produce that level of achievement for its pupils. On the other hand, a relatively poor school may be making the best use of its inputs (and thus be considered efficient) but may still prove to be ineffective because the achievement levels fall short of what is deemed to be acceptable.

Our results show that if one looks at the school inputs only, the output levels indicate considerable inefficiency in resource utilization in the large and urban school districts. However, once adjustments are made for the socio-economic disadvantages in these communities, these districts do not appear to be significantly less efficient than others. The implication is that only limited gains in achievement can be expected from improved management of the existing levels of school inputs in these districts and for substantial improvements major socio-economic programs may be called for.

This paper is organized as follows. In §2, we present the basic DEA model and the revision. In §3 we briefly discuss the findings of our DEA application with Connecticut data. This is followed by the regression results and their implications in §4. The conclusions are summarized in §5.

2. Data Envelopment Analysis

If education is viewed as an industry transforming inputs into outputs, each school or district (treated as a DMU) can be viewed as a multi-product firm. Let $q_t = (q_{1t}, q_{2t}, \ldots, q_{mt})$ and $x_t = (x_{1t}, x_{2t}, \ldots, x_{nt})$ represent the output and input bundles of DMU t ($t = 1, 2, \ldots, N$). As suggested by CCR (1978, 1981), a scalar measure of the productive efficiency of DMU t can be obtained as the optimal solution of the conceptual model:

$$\max h_t = \left[\sum_{r=1}^{m} v_{rt} q_{rt} \Big/ \sum_{i=1}^{n} u_{it} x_{it} \right]$$

$$\text{subject to} \quad \left[\sum_{r=1}^{m} v_{rt} y_{rj} \Big/ \sum_{i=1}^{n} u_{it} x_{ij} \right] \le 1 \qquad (j = 1, 2, \ldots, N) \qquad \text{and}$$

$$v_{rt},\ u_{it} \ge 0 \quad (r = 1, 2, \ldots, m;\ i = 1, 2, \ldots, n). \tag{1}$$

The weights u_{it} and v_{rt} are 'virtual prices' to be used to aggregate the outputs and inputs of DMU t. In this sense, h_t is a measure of factor productivity. By construction this measure of efficiency cannot exceed 1.

In practice, however, one does not need to solve the problem above. CCR (1978, 1981) have shown that the same efficiency measure can be obtained from the solution of the following simpler linear programming (LP) problem:

$$\max f_t = z_t + \delta \sum_{r=1}^{m} s_{rt}^{+} + \delta \sum_{i=1}^{n} s_{it}^{-}$$

$$\text{subject to} \quad \sum_{j=1}^{N} \lambda_j q_{rj} - s_{rt}^{+} = z_t q_{rt} \qquad (r = 1, 2, \ldots, m),$$

$$\sum_{j=1}^{N} \lambda_j x_{ij} + s_{it}^{-} = x_{it} \qquad (i = 1, 2, \ldots, n),$$

$$\lambda_j,\ s_{rt}^{+},\ s_{it}^{-} \ge 0 \qquad (j = 1, 2, \ldots, N;\ r = 1, 2, \ldots, m;\ i = 1, 2, \ldots, n). \tag{2}$$

It can be shown that at the optimal solution $h_t^* = 1/f_t^*$. This LP problem has a simple intuitive interpretation. We can visualize a hypothetical DMU constructed as a linear combination of some (or all) the units in the sample. Any individual unit j is assigned weight λ_j (≥ 0) in this construction. Let $y_r^* = \sum_{j=1}^N \lambda_j y_{rj}$ and $x_i^* = \sum_{j=1}^N \lambda_j x_{ij}$ be the rth output produced and ith input used by this synthetic DMU. By construction $y_r^* \geq z_t y_{rt}$ and $x_i^* \leq x_{it}$. Thus the synthetic DMU could produce at least z_t times each output produced now by unit t without using any more of any input than what this unit is using. For example, let $z_t = 1.25$. Then it is possible to increase every output by at least 25% without any increase in inputs. Indeed when $s_{rt}^+ > 0$ output r can be expanded beyond 25%. A charitable measure of efficiency of DMU t would then be $1/1.25 = 0.8$. It should be noted that because δ is arbitrarily small (typically 0.000001) f_t will essentially be determined by z_t and h_t is approximately equal to $1/z_t$. The role of the other terms is to ensure that a DMU does not get an efficiency rating of 1 even when $z_t = 1$ so long as it is possible to increase any individual output or to reduce any individual input.

As stated earlier, when the outputs are bounded from above one has to ensure that the implied levels of outputs at the fully efficient level of operation of a DMU do not exceed the prescribed upper limits. In terms of the hypothetical DMU described above, it would require that $q_r^* \leq \bar{q}_r$ ($r = 1, 2, \ldots, m$) where $\bar{q}_r$ is the upper limit of output r. In our specific context the upper limits denote all correct or perfect answers. Failure to reach these maximum levels of achievement would imply that there is some scope for improvement. We accordingly modified the LP model (2) above. Our measures of efficiency are obtained from the following optimization problem:

$$\max z_t + \delta \sum_{r=1}^{m} s_{rt}^+ + \delta \sum_{i=1}^{n} s_{it}^- + \delta \sum_{r=1}^{m} \sigma_{rt}^+$$

$$\text{subject to} \quad \lambda_1 q_{r1} + \lambda_2 q_{r2} + \cdots + \lambda_t q_{rt} + \lambda_N q_{rN} - s_{rt}^+ = z_t q_{rt} \qquad (r = 1, 2, \ldots, m);$$

$$\lambda_1 q_{r1} + \lambda_2 q_{r2} + \cdots + \lambda t q_{rt} + \lambda_N q_{rN} + \sigma_{rt}^+ = \bar{q}_r \qquad (r = 1, 2, \ldots, m),$$

$$\lambda_1 x_{i1} + \lambda_2 x_{i2} + \cdots + \lambda_t x_{it} + \cdots + \lambda_N x_{iN} + s_{it}^- = x_{it},$$

$$\lambda_j, s_{rt}^+, s_{it}^-, \sigma_{rt}^+ \geqq 0 \qquad (j = 1, 2, \ldots, N; r = 1, 2, \ldots, m; i = 1, 2, \ldots, n). \qquad (3)$$

Here $\bar{q}_r$ is the upper limit of the rth output and σ_{rt}^+ is the amount by which the rth output of the hypothetical process fails to reach the capacity level. In this revised model DMU t would be assigned an efficiency rating 1 only when $q_{rt} = q_{rt}^* = \bar{q}_r$ for all r and $x_{it} = x_{it}^*$ for all i.

The dual of the LP problem in (3) is worth looking into because of its interesting interpretation. The dual problem would be:

$$\min \sum_{i=1}^{n} \mu_i x_{it} + \sum_{r=1}^{m} \pi_r \bar{q}_r$$

$$\text{subject to} \quad \sum_{i=1}^{n} \mu_i x_{ij} + \sum_{r=1}^{m} \pi_r q_{rj} - \sum_{r=1}^{m} y_r q_{rj} \geqq 0 \qquad (j = 1, 2, \ldots, t, \cdot, N);$$

$$\sum_{r=1}^{m} y_r q_{rt} = 1;$$

$$y_r, \mu_i, \pi_r \geqq \delta \qquad (r = 1, 2, \ldots, m; i = 1, 2, \ldots, n). \qquad (3a)$$

Here μ_i and y_r are the accounting prices of input i and output r, respectively. Additionally, π_r is the overhead charge per unit of output r. The imputed value of the observed output bundle of firm t is set equal to 1. The inequality constraints imply that the imputed cost of the variable inputs plus overhead charges cannot be lower than the value of the output bundle of any DMU j ($j = 1, 2, \ldots, t, \cdot, N$). The objective function is the total of variable cost and overhead cost for DMU t. It may be rewritten as

$$\sum_{i=1}^{n} \mu_i x_{it} + \sum_{r=1}^{m} \pi_r(\bar{q}_r - q_{rt}) + \sum_{r=1}^{m} \pi_r q_{rt}.$$

This expression at its minimum would be $\sum_{r=1}^{m} \pi_r(\bar{q}_r - q_{rt}) + 1$ because

$$\sum_{i=1}^{n} \mu_i x_{it} + \sum_{r=1}^{m} k_r q_{rt} \geqq \sum_{r=1}^{m} y_r q_{rt} = 1.$$

Therefore the objective function becomes 1 only when $q_{rt} = \bar{q}_r$ for all r apart from satisfying the minimality of the variable input bundle. Our computations of relative efficiency of each school district were based on model (3).

2.1. *Efficiency and Nondiscretionary Inputs*

Link between the nondiscretionary inputs in the production process and the measure of efficiency obtained by DEA was discussed for the single-output case in Ray (1988a) and can be briefly described as follows. For DMU t, let q_{0t} represent its scalar output; x_t and w_t are the vectors of controllable and nondiscretionary inputs. Consider the production function

$$q_{0t} = F(x_t, w_t) \quad \text{and let} \quad F(x_t, w_t) = g(x_t) \cdot h(w_t). \tag{4}$$

Such multiplicative separability holds for some commonly used production functions like the Cobb Douglas. Assume that $h(w_t)$ lies between (0, 1) for all values of w_t. If we specify that the function $g(\cdot)$ is linear (or log-linear in order to permit variable returns to scale) it can be shown (Ray 1988a, pp. 170–174) that the DEA measure of efficiency (h_t) corresponds to the component $h(w_t)$ of the production function above. One can generalize this to the multiple output case by replacing the scalar output q_{0t} by a linear (log-linear) output function $T(q_t)$ where q_t is the vector of outputs.

The function $h(w_t)$ defines the maximum efficiency level attainable corresponding to a given configuration (w_t) of nondiscretionary inputs. But if there is inefficiency due to poor management the actual efficiency level attained will be even lower. We can, therefore, relate the measured efficiency (h_t) to the nondiscretionary inputs (w_t) in a regression model as:

$$h_t = h(w_t) + \epsilon_t, \qquad \text{where} \qquad \epsilon_t \leq 0. \tag{5}$$

Note that ϵ_t represents pure and avoidable inefficiency caused by mismanagement. Without such managerial inefficiency h_t attains its highest value defined by $h(w_t)$. Otherwise, $h_t < h(w_t)$. It should be stressed that the maximum attainable efficiency depends on w_t and need not be 1 for every DMU. If h_t equals $h(w_t)$ we should conclude that there is no managerial inefficiency and failure to reach 100% efficiency (if $h_t < 1$) is due to external factors beyond the control of the DMU.

The practical steps in measuring efficiency will involve:

(i) specifying the output vector (q) and the input vector (x) and obtaining measures of efficiency by DEA computing one LP model (from (3)) for each decision-making unit in the sample;

(ii) specifying the external factors affecting output but not under control of the decision-making unit; and

(iii) estimating a suitable statistical relation between measured efficiency (from (i)) and the external factors (specified in (ii)).

3. Educational Production Analysis and DEA

Despite severe problems of measurement (discussed, for example, in Hanushek 1976, 1979 or Cohn 1979) there exists a significant volume of literature investigating the relation between outputs and inputs of educational institutions through production functions. Outputs are typically measured by some standardized achievement test scores although it is generally acknowledged that such scores do not appropriately measure certain outcomes like creativity or attitude. Various kinds of school resources and socio-economic characteristics of the pupil's family are used to measure inputs.

These studies have two major disadvantages. First, the production function specified is estimated as a standard regression model incorporating the usual two-sided random disturbance term. As a result, the predicted values from the fitted regression models provide the expected (or average) level of outcome from a given bundle of school and nonschool inputs—not the maximum achievable outcome. Therefore, such models cannot be used to measure efficiency.[1]

Secondly, most of these studies specify a single output production function although some of them (e.g., Levin 1970 and Michelson 1970) estimate simultaneous equation models to estimate multiple output production technologies.[2]

Both of these concerns can be addressed if one uses DEA instead of standard regression analysis. Designed explicitly to measure efficiency, DEA attempts to determine maximal rather than average output levels producible from given input bundles. Moreover, DEA can treat the observed output-input vectors as multiple-output production processes and eliminates the need for aggregating outputs into a single scalar index.

Among previous applications of DEA with educational data are the studies by CCR (1978, 1981), Bessent and Bessent (1980), and Bessent et al. (1982).[3] In all of those studies nondiscretionary inputs (measured by the socio-economic status (SES) variables) and school inputs were used together in the DEA models for individual units. In the present study only the school inputs are used at the DEA stage and the SES variables feature only in the subsequent regression analysis. When school inputs and socio-economic factors are treated alike in the DEA models, one can identify possible economies in the use of inputs like teacher (full time equivalent) as well as in the socio-economic factors like average education of mother or percentage of nonminority population. Even though slacks in the nondiscretionary inputs may not be used in the objective function of LP (3) in order to compute efficiency, it is somewhat difficult to conceptualize such slacks. Our approach treats the nondiscretionary inputs as fixed (and indivisible) and uses them only in the regression model.

Another problem about using DEA with all inputs (controllable and nondiscretionary) is that whenever any input is excluded or a new input is included all the individual LP problems have to be solved again in order to get a new set of efficiency ratings. In our case this will not be necessary unless the set of school inputs to be used in the analysis is changed. Exclusion of any socio-economic status variable from the study or inclusion of a new one can easily be handled by estimating an additional regression model. The left-hand side variable is the efficiency measure based on outputs and school inputs alone

[1] Levin (1976) uses LP to estimate a frontier function and it can be used to measure efficiency.

[2] A multiple output but single equation model using canonical ridge regression was estimated by Chizmar and Zak (1983).

[3] See also Bessent et al. (1984).

and will not change unless these variables are changed. This can reduce the burden of computation considerably.

3.1. *The Connecticut Study*

This study uses DEA to measure efficiency for 122 districts operating high schools in Connecticut during 1980–81.[4] Three kinds of personnel input are included. These are full time equivalents of: (i) classroom teachers per pupil (X_1), (ii) support staff (like counselors, psychologists, reading consultants etc.) per pupil (X_2), and (iii) administrative staff per pupil (X_3). The output measures used are district averages of scores of 9th grade students in the State Wide Proficiency Tests administered in October 1980. The output variables are: (i) mathematics score (Q_1), (ii) language arts score (Q_2), (iii) writing score (Q_3), and (iv) reading score (Q_4). Maximum possible values were 100 for Q_1, Q_2, and Q_4 and 8 for Q_3.

122 LP models—one for each school district—were solved using the school data. The measured efficiency ratings[5] of individual districts vary from a low of 0.6309 for Hartford to a high of 0.9999 for Brooklyn and New Milford. These extremely high ratings for the two towns can be explained in the following way. For Brooklyn the input of x_2 (support staff per pupil) is 0.00146 whereas the mean for all districts was 0.00375. This input level is lower than the average by a margin of 2.5 times the standard deviation. As a result, even though its output measures are around the average levels this input acts as a constraint and results in an extremely high rating. Similarly, the input of administrative staff (X_3) is very low for New Milford which is reflected in its efficiency measure. It would be instructive to find out how these two districts manage to function with such low levels of these inputs (compared to the other districts).

Reliability of the efficiency ratings obtained was checked by performing sensitivity analysis in several ways. The DEA models were solved several times with the following changes: (i) exclusion of input X_3, (ii) exclusion of the regional school districts, and (iii) inclusion of the dollar value of nonpersonnel expenditure as an additional input. In all cases correlation between the original and the new measures of efficiency exceeded 0.93. We can conclude, therefore, that efficiency of school inputs measured by our original DEA is quite reliable.

4. Regression Analysis

Different variables measuring the socio-economic characteristics of the communities served by individual districts were identified as nondiscretionary inputs to be used in the second-stage regression model with the DEA efficiency measure (h_t) as the dependent variable. These reflect 'advantages' and 'disadvantages' in the homelife of the typical student.[6] Among the 'advantage' variables considered were: (i) parental education measured by the percentage of the population 25 years or older with 4 or more years of college education (COLL) and measures of affluence like (ii) 1979 per capita income ($ thousands) (PCY) and (iii) median value of owner occupied housing units ($ thousands in 1980) (HVALU). The 'disadvantage' variables included: (iv) percentage of students in the district from ethnic minority groups (MIN), (v) percentage of students from families receiving Aid to Families with Dependent Children (AFDC), (vi) percentage

[4] The data for DEA and also for the subsequent regression analysis were taken from State of Connecticut Board of Education (1982, 1983). Descriptive statistics about the output and input data are reported in Table 1 of Ray (1988b).

[5] DEA efficiency ratings for individual districts are reported in Ray (1988b).

[6] See Ray (1988b) for a discussion about the choice of variables. Descriptive statistics about the socio-economic variables are reported in Table 2 of May (1988b).

of families with income below the poverty level in 1979 (POV), and (vii) percentage of children in the community from single-parent families (SPRNT).

Anticipated sign of the coefficient is positive for an 'advantage' variable and negative for a 'disadvantage' variable. A linear regression model incorporating all of these variables above proved to be very unsatisfactory. Due to severe multicollinearity caused by high pairwise correlation between many of the explanatory variables many of the coefficients were statistically insignificant. Several alternative models were estimated excluding different variables in different specifications. The variable MIN was included in each model because of its high statistical significance.

Judging by the goodness of fit (R^2) and the 't' ratios of individual coefficients the most preferred model was:

$$\begin{aligned}\hat{h}_t = {} & \underset{(t = 57.7)}{0.87348} + \underset{(t = 3.58)}{0.00174}\ \mathrm{COLL}_t - \underset{(t = -3.8)}{0.00195}\ \mathrm{MIN}_t - \underset{(t = -3.8)}{0.00204}\ \mathrm{SPRNT}_t, \\ & R^2 = 0.6368 \text{ number of observations} = 102. \end{aligned} \tag{6}$$

It was stated earlier that the regression model specified should include a random disturbance term which is nonpositive. That would ensure that the predicted value of $h(\hat{h}_t)$ never falls below the observed value (h_t). As such the regression reported above uses standard least squares procedure and the predicted value from the model above does not always satisfy the required inequality with respect to the actual value. However, a simple adjustment of the intercept term enables us to achieve that objective. Greene (1980) has shown that one can obtain consistent estimators of the parameters by adding the largest positive residual to the intercept and recomputing the residuals by subtracting this value from each residual obtained by least squares. These adjusted residuals will thus become all nonpositive. In our case the largest residual was 0.136686. We accordingly added this term to the intercept to obtain the adjusted measure $\hat{\hat{h}}_t = \hat{h}_t + 0.136686$. For illustration consider the district of Meriden. For this district the DEA measure h is 0.798996. The unadjusted predicted value ($2\hat{h}$) is 0.806514 and the adjusted value ($\hat{\hat{h}}_t$) is 0.94320. Table 1 shows the frequency distributions of the adjusted and unadjusted measures of inefficiency.

The regression results can be used as follows. The adjusted predicted value ($\hat{\hat{h}}$) estimates the maximum efficiency in use of school inputs attainable given a specific level of

TABLE 1

Frequency Distribution of Inefficiency

Range	Adjusted	Unadjusted
0–5%	6	6
5%–10%	19	15
10%–12.5%	13	13
12.5%–15%	34	22
15%–17.5%	23	18
17.5%–20%	6	14
20%–25%	1*	9
25%–30%	0	2
30%–35%	0	2
>35%	0	1**
Mean	12.64%	14.63%

* 21.16%

** 36.9%

the socio-economic characteristcs. For the town of Meriden just considered this efficiency would be 0.9432 *if there were no managerial inefficiency*. However, the realized efficiency level was 0.798996. Hence $(\hat{\hat{h}} - h)$ or 0.144204 is the extent of managerial inefficiency not caused by external factors. The point to be emphasized is that inefficiency should be measured as the shortfall of h from $\hat{\hat{h}}$ and not from 1. Significance of this adjustment can be clearly seen from the case of a town like Hartford. For this district the DEA measure was 0.630998 which would suggest that Hartford had inefficiency measuring up to 0.369002 in utilization of its school resources. However, the adjusted predicted value was 0.75417 resulting in a measured value of 0.12317 for managerial inefficiency. Thus a large part of what appeared to be inefficient management of resources by the administration is accounted for by disadvantageous socio-economic status of the families served. Similar conclusions also emerge for other urban school districts like New Haven or Waterbury.

One problem in this regression model is that the predicted value $(\hat{h})$ may not lie below 1. The problem is further aggravated for the upward adjusted value $\hat{\hat{h}}$. As is usually done in respect of the linear probability model, we can replace the adjusted value $\hat{\hat{h}}$ by 1 when that happens and measure managerial inefficiency as $(1 - h)$. Note that $\hat{\hat{h}}$ is more likely to exceed 1 for the more advantaged districts. Hence this procedure will probably underestimate the level of inefficiency in such cases. It is quite interesting to note that in spite of such likely bias in favor of the affluent and better educated communities the level of managerial inefficiency at a town like Greenwich is found to be 0.118002, only slightly better than at Hartford. The implication is that given the home background of its pupils the school district of Hartford is not performing far worse than Greenwich in managing its resources. Hartford is not a particularly bad school district; it is a significantly disadvantaged community. These adjusted inefficiency measures can be used to set realistic targets for improvement in student achievement level by better management only. Further improvement would require some form of social reform not within the jurisdiction of the school administration.

The estimated regression model does provide some help in this respect. For example the coefficient of the variable *COLL* indicates that if an additional 2% of the adult population could be induced to complete a 4-year college degree the predicted efficiency of schools would increase by 0.00348. Outputs can be increased by either using more resources or by improving the productivity of the existing resources. Special plans directed towards preventing drop outs in colleges can be considered as a means of improving efficiency at schools. The cost of such programs can be compared with alternative plans of simply allocating more funds for resources at the high schools.

5. Conclusion

Our study of the data from Connecticut school districts shows that efficiency in utilization of school inputs varies systematically with the socio-economic characteristics of towns. Once account is taken of these external factors the average level of managerial inefficiency is found to be 12.64%. This provides a benchmark for improvement in the levels of achievement from existing resources through better management. However, given the particularly low levels of achievement in the urban school districts of the large cities even such gains would prove to be inadequate. In such cases one needs to consider either allocating more resources to these districts or undertaking broader socio-economic programs designed to relax the constraints presently adversely affecting productivity of school inputs.

A word of caution is warranted. Our results are based on input output data from a single year and may have been influenced by random elements and measurement errors in the data. Lack of sufficiently detailed information precluded any adjustment for difference in quality of inputs. Nor was it possible to control for the prior level of achievement of the pupils. Our findings should therefore be interpreted as broad indicators.[7]

[7] The author is indebted to Arnold Barnett and two anonymous referees for valuable comments. Responsibility for errors remaining rests on the author alone.

References

BESSENT, A. M. AND E. W. BESSENT, "Determining the Comparative Efficiency of Schools through Data Envelopment Analysis," *Educational Admin. Quart.*, (1980), 57–75.

———, W. BESSENT, J. KENNINGTON AND B. REAGAN, "An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District," *Management Sci.*, 28 (1982), 1355–1367.

———, ———, J. ELAM AND D. LONG, "Educational Productivity Council Employs Management Science Methods to Improve Educational Quality," *Interfaces*, 14, 6 (1984), 1–8.

CHARNES, A., W. COOPER AND E. RHODES, "Measuring the Efficiency of Decision Making Units," *European J. Oper. Res.*, 2 (1978), 429–444. See also "Corrections," *European J. Oper. Res.*, 3, 339.

———, ——— AND ———, "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through," *Management Sci.*, 27 (1981), 668–697.

CHIZMAR, J. F. AND T. A. ZAK, "Modeling Multiple Outputs in Educational Production Functions," *Amer. Economic Rev.*, 73, 2, (1983), 59–94.

COHN, E., *The Economics of Education*, (3rd ed.), Ballinger Publishing Co, Cambridge, MA, 1979.

GREENE, W., "Maximum Likelihood Estimation of Econometric Frontier Functions," *J. Econometrics*, 13 (1980), 26–57.

HANUSHEK, E., "Comments," in J. T. Froomkin, D. T. Jamison, and R. Radner (Eds.), *Education as an Industry*, Ballinger Publishing Co. for National Bureau of Economic Research, Cambridge, MA, 1976.

———, "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *J. Human Resources*, 14, 3 (1979), 351–388.

LEVIN, H. M., "A New Model of School Effectiveness," in U.S. Department of Health, Education, and Welfare, *Do Teachers Make a Difference?*, U.S. Government Printing Office, Washington, D.C., 1970.

———, "Concepts of Economic Efficiency and Educational Production," in J. T. Froomkin, D. T. Jamison, and R. Radner (Eds.), *Education as an Industry*, Ballinger Publishing Co. for National Bureau of Economic Research, Cambridge, MA, 1976.

MICHELSON, S., "The Association of Teacher Resources with Children's Characteristics," in U.S. Department of Health, Education, and Welfare, *Do Teachers Make a Difference?* U.S. Government Printing Office, Washington, D.C., 1970.

RAY, S. C., "Data Envelopment Analysis, Non-discretionary Inputs and Efficiency: An Alternative Interpretation," *Socio-Economic Planning Sci.*, 22 (1988a), 167–178.

———, "A Study of Resource Utilization in Connecticut Schools: Socio-economic Factors and Inefficiency," (mimeo), Department of Economics, University of Connecticut, Storrs, CT, 1988b.

State of Connecticut Board of Education, *Condition of Education 1980-81. Vol. 2. Town and School District Profiles*, 1982.

———, *Condition of Education 1981-82. Vol. 2. Town and School District Profiles*, 1983.