This report analyzes the relationships between such inputs as expenditure per pupil and school district size and the average achievement of pupils in basic subjects. The analysis focuses on production function relationships of education. The results of the analysis illustrate the problems involved in constructing meaningful production functions from data that are essentially cross-sectional. The data are discussed, a model of the educational process is proposed, and the problems concerning its implementation with available data are explored. Empirical findings relate the most important inputs to the average achievement level of 6th grade pupils who were present in the 4th grade, as well as to the average performance gain over the 2-year period. (Author/JF)
THE RELATIONSHIP OF SCHOOL INPUTS TO PUBLIC SCHOOL PERFORMANCE IN NEW YORK STATE

Herbert J. Kiesling *
Consultant to The Rand Corporation, Santa Monica, California

Introduction

In an earlier paper (14) the author related expenditure per pupil and school district size to average achievement performance of pupils in basic subjects in a sample of school districts in New York State. The focus of that paper was upon education considered as a local government service and examined the relationships of expenditure per pupil and school district size to the performance of pupils from roughly similar socio-economic backgrounds. Since the writing of that paper, these New York school data have been analyzed much more carefully with an eye to shedding some light upon production function relationships of education when considered as a firm. This paper outlines some findings from that research.

It has been difficult for economists to establish production functions for firms even in the private sector, ** and it would be presumptuous to claim much progress in establishing production function relationships for public education. While a number of investigators are currently engaged in this kind of work, it is probably safe to say that cross-section analysis is more properly viewed as descriptive of

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation of the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The Rand Corporation as a courtesy to members of its staff.

**See (4), especially papers by Solow and Carter.
past school performance than technical production functions. It may be that the latter can only be correctly established with highly controlled longitudinal experiments. The results given here are meant to give some appreciation of the problems involved in constructing meaningful production functions from data that are essentially cross-sectional in nature and to present a descriptive report of how schools in New York State—based upon a fairly representative sample—seemed to be performing in the late 1950's. After a brief discussion of the data, a model of the educational process will be proposed and problems associated with its implementation with available data discussed. Empirical findings are given relating the most important inputs to the average level of achievement performance of sixth-grade pupils who were present in grade 4 and also the average gain in performance over the two-year period.

Data and Data Limitations

The data upon which this study is based is described to some extent in the author's earlier paper. Fourth grade pupils in 97 school districts were given the Iowa Test of Basic Skills (2nd) in the 1951 school year and retested the next year as fifth graders and retested again the third year when in grade 6.*

A discussion of the strengths and weaknesses of achievement scores in basic subjects as a measure of public school performance was included in the

*This was also done for grades 7-9 and 10-12, but only the results for the elementary grades are discussed here because of space limitations. It is to be noted that all pupils present in a given year were tested, not merely the pupils who had been tested in the first year.
author's earlier paper and will not be repeated here. In that paper the achievement measure used was the average for all pupils present in a given year, a procedure which is potentially dangerous since it is possible that many of the students may have moved into the school district just prior to the year of the test, with test results therefore measuring performance due to the inputs of other school districts besides the one being measured. To avoid this danger the data were resummarized for this paper to represent the average performance for pupils in the sixth grade who were also present in the fourth grade. This assures that the pupils whose performance is being examined were present in the school district being studied for at least three years, and, given the mobility patterns described in the previous note, probably for most of their six-year school careers. Test scores were averaged by school district for sixth-grade level and for gain from grades 4 to 6 for all pupils and for pupils in five groups stratified according to occupation of family breadwinner.

Members of the original New York State Study project** and the author gathered a variety of information concerning characteristics

---

*For most of the school districts in the sample the mobility was not unreasonably great. Most smaller and medium-sized school districts had ratios of numbers of matched pupils to the greatest number of pupils present for tests in any of the three years between 70 and 80 percent. Using the greatest number of pupils present undoubtedly overstated the mobility rate somewhat. In the 19 largest school districts, where mobility is presumably the greatest, the average ratio for the best 15 school districts was 77 percent. The other four large school districts, however, had extremely high mobility rates with the average percentage of matched scores of the total in those districts being only 20 percent. Two of these four districts were not used in the analysis for other reasons, however. (See note below.)

**The "Quality Measurement Project."
of the school districts, the most important of which includes the following variables:

1. Teacher-Pupil Ratio  
2. Principals/Supervisors to Pupil Ratio  
3. Special Staff Personnel to Pupil Ratio  
4. Expenditure per Pupil on Books and Supplies  
5. Median Teacher Salary  
6. Average Salary of Teachers in the Top Salary Decile in Grades Kindergarten Through Six  
7. Average Socio-Economic Index of Occupation of Family Breadwinner of Pupils in Grade 5  
8. Amount of School District Debt per Pupil  
10. School District Size in Average Daily Attendance  
11. School Property Value per Pupil  
12. The Salary of the Superintendent of Schools  
13. Median Salary of Principals  
14. Expenditure per Pupil on Principals, Assistant Principals, and Supervisors  
15. School District Value of Buildings per Classroom  
16. School District Value of Furniture and Equipment per Classroom  
17. Median Years of Teacher Experience in the School District  

A few community characteristics such as tax rate, tax base per pupil, and type (village, urban, etc.) were gathered. Other community characteristics were difficult to obtain because school district boundaries are seldom coterminous with jurisdictions used by the U.S. Census.
The importance of many of these variables is obvious. The ratios of teachers, principals, and special staff personnel are important direct school inputs. Expenditure per pupil on principals, assistant principals, and supervisors is a close proxy variable for the principals variable and possibly more accurate for purposes of measuring resources going into supervision. Books and supplies are meaningful educational inputs. Salaries and experience levels are attributes almost always important to the operation of firms.

As a list from which we hope to choose the variables which will explain school performance, it has real weaknesses, however. If an investigator wishes to capture differences in personnel quality, salary variables will undoubtedly be found wanting, at least with respect to education. (This is also true—and probably more so—with experience levels, especially for teachers.) The fundamental reason for this is the widespread use of the single-salary schedule in education in which minimum salary levels (and often, salaries actually paid) are a function of, and only of, number of years experience. Second, enough market imperfections exist such that salary levels in different locations are not strictly comparable.

* The principals variable does not include "supervisors."

** It appears from work done by Katzman, (13) Turner, (22) and in other analysis of these New York data, that additional experience beyond a certain point (which may come relatively early) in the teacher's professional career has little effect on teaching effectiveness. There were a large number of school districts (about one-third) in this sample for which it was not possible to collect teacher-experience data and the variable is not included in the discussion below. Multiple regression analysis for the approximately 60 school districts for which teacher experience data were available displayed the interesting finding that experience was unrelated to the achievement performance of pupils from the top occupational background but significantly and positively related to the performance of pupils from lower occupational backgrounds.
Some salary variables are probably more affected by this than others. Of those on the list, the best salary variable is undoubtedly mean salary of the top teacher salary decile. This is because single-salary schedules have more effect upon salary floors than ceilings and more variance would be expected because of school district salary policy in the top salary bracket than at any other level.*

The value variables, finally, are only weak first approximations and are highly inadequate. No better data were available, unfortunately. From all the author has been able to observe concerning New York schools at this time it would appear that most of them used highly similar mixes of capital inputs and therefore the value variables might be somewhat better than one would otherwise expect.

An Educational Model: Theoretical Considerations

It is not possible for an economist (especially) to build a model which is satisfactory for understanding the educational process. Much has been said by other writers concerning a lack of a satisfactory theory of learning which would be a necessary ingredient.**

Let us tentatively assume that the quality of a child's education is causally related to four variates--the formal school education process, the informal home and environmental educational process, motivation toward learning, and native ability. In order to pursue the goal of examining the formal school process, it is necessary to isolate the effects of informal home and environmental process upon the child's

*Except for beginning salaries perhaps. Beginning salary is a relatively good salary variable also.

**See, for example, (2), p. 4.
learning as much as possible. The child's motivation to learn is a rather complicated factor in the learning process, being a function (at least) of home environment, influence of peers, and influence of school environment. And finally the success of formal education presupposes the ability of pupils to absorb knowledge if only they make the attempt. The crucial question for the analysis of formal education is whether enough of the non-school effects can be captured such that the findings for the school effects are interpretable.

Since differences in native ability are either randomly distributed or associated with socio-economic background, it seems unnecessary to use intelligence test score as an explanatory variable for differences in native ability. If native ability is associated with socio-economic status, the proper explanatory variable is a socio-economic status variable.*

* This represents a departure from the author's thinking when the earlier paper was written.

Not using I.Q. scores eliminates the problem of simultaneity which was present in the earlier paper because intelligence and achievement tests to a large extent measure the same thing. Because of this, both the magnitude and significance of the estimates for the expenditure-performance relationship were underestimated in the prior paper. Space does not permit a complete listing of the correct estimates. The following are the values for the t-statistic for the urban school districts for sixth-grade composite score and gain in composite score from grades 4 to 6.

<table>
<thead>
<tr>
<th>Population</th>
<th>Sixth Grade</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional</td>
<td>-0.41</td>
<td>1.22</td>
</tr>
<tr>
<td>Managers, officials</td>
<td>2.26</td>
<td>1.75</td>
</tr>
<tr>
<td>Clerks</td>
<td>0.87</td>
<td>1.38</td>
</tr>
<tr>
<td>Skilled, semi-skilled</td>
<td>0.71</td>
<td>2.30</td>
</tr>
<tr>
<td>Unskilled</td>
<td>-0.08</td>
<td>0.70</td>
</tr>
</tbody>
</table>

The model used had performance a function of expenditure, size, growth, and occupation index. As in the author's earlier paper, the middle SES groups show performance more related to expenditure than the extremes.
Pupil motivation is difficult to account for with the data at our disposal. It might be assumed that the motivation imparted by the school is captured by achievement test scores themselves and therefore no further notice needs be taken of that aspect. We have no good variable for capturing the influence of the pupil's peer groups, although average occupational background of the pupils in the school is probably a fair proxy. The other part of motivation, as well as actual learning in the home, is a function of the home environment and a good socio-economic status variable might be expected to account for these influences.

Problems in Going from Theoretical Model to Empirical Model

A number of problems stand between conception and implementation of the scheme just described. They are of two types having to do with model specification and statistical application.

Specification Problems

Several specification problems exist. First it is necessary to specify the proper unit of production. Perhaps most proper in theory would be the pupil's experience in the individual classroom. The

*As Bowles suggests, (2) p. 9, there is an element of simultaneity introduced because of the relationship between achievement performance and motivation. Thus high achievement creates confidence which creates increased motivation which in turn produces higher achievement, etc. As Bowles also suggests, since we are merely interested in performance levels, and not particularly in the process by which the pupil arrives at high levels of performance, this phenomenon is probably not germane to the present discussion.
smallest administrative unit in public schools is the individual school building under the administration of a single principal, while the smallest completely independent administrative unit in the public schools is the school district. Richard Turner has shown that as a result of uniform hiring policies teacher characteristics were quite homogeneous in a sample of school districts in Indiana.* On the other hand, we are aware that in some of the larger school districts quality of teaching and administrative personnel is not homogeneous between schools because of the ability of teachers, and to a lesser extent administrators, to transfer.** Thus teachers with greater seniority can be expected to transfer from ghetto schools to silk-stocking or upper-middle-class schools in most instances. We have achievement performance data both by school building and by school district and have summarized it on that basis. However, input data by school building is virtually impossible to obtain. An aggregation problem exists, therefore, which is difficult to overcome satisfactorily.

Although the Turner findings offer some consolation with respect to this difficulty, the school districts were examined for heterogeneity more directly. Differences in the occupational backgrounds of the children in different schools within each school district were carefully examined, the hypothesis being that if the pupil characteristics vary greatly from school to school there is much greater likelihood that teacher characteristics differ also, since it is because of such differences that teachers usually ask for transfers within the school.

*See (22) Chapters 5 and 6.
**See (15) Chapters 4, 6, and 7.
district. Upon analyzing the homogeneity of school districts in the sample, seven were judged to be excessively heterogeneous and those districts were discarded from the analysis.*

Another important issue concerning these data concerns whether the index of parent's occupation accounts for socio-economic background differences on the part of school districts and students in a manner which is reasonably satisfactory.** Two problems are involved. First, is the occupation index itself reliable as an indicator of socio-economic status? Second, if the influence of socio-economic background is removed with a stratification scheme which utilizes the five categories used in the analysis, are the stratification intervals too wide to be helpful?

The answer to the first question would seem to be affirmative. Kahl and Davis (11) made a factor analysis of the relationships of five indexes of socio-economic status to a number of other variables denoting socio-economic characteristics and found that the occupation indexes

*In the original data there were eight socio-economic occupation categories. For the purpose of summarizing the data by school, each individual pupil's classification from one to eight was taken and averaged into a mean for each school building. Then we arbitrarily assumed that if an individual school building in a school district was within one-half of one category of the school district mean that it showed homogeneity in the school district, while if it was within three-tenths of one classification from the mean it would show high homogeneity. Seven districts had less than 60 percent of their schools within one category of the district mean and were therefore discarded from the analysis. Of the seven districts that were not used, two were large cities, two medium-sized, and three were small cities. As a result of doing this, plus a slight underrepresentation of large school districts to begin with, large school districts are underrepresented in the findings that follow.

It should be emphasized that a very large percentage of the school districts reported upon in this paper were highly homogeneous from school building to school building in terms of father occupation characteristics and that about half of the districts had only one elementary school building.

**The occupational index used is an earlier version of that generally used by the U.S. Census. (7) There is very little difference between the earlier and later versions of the index.
(one of which being the one used here) were the most highly correlated with most characteristics. Baer (9) found that the occupation index, while not as good as either education or a more complex index (Index of Status Characteristics), was nevertheless a good SES indicator. A reading of the sociological literature gives the impressions that occupation is well regarded as a status indicator. Also, from all the author has been able to learn, it appears that the information itself was collected with great care by most of the school districts.

Medsker and Trent, in a study of college attendance, state: "Of the several indices which reflect the socio-economic and family background of any particular group within the general population, the occupation of the father is considered by many to be the best single index." (17)

It is probably fair to say that comments in some of the economic literature have been overcritical of the use of occupation in favor of income. For example, Burkhead in his study of high schools in Chicago and Atlanta, commenting upon the author's doctoral dissertation: "The study depends on occupation of parents as the single measure of socio-economic status; this is unfortunately not a reliable measure of relative family income among different communities." (5)

It is probably only natural that economists would over-stress income as a socio-economic status variable. It would be helpful that those of us so inclined consider the following statement by Kahl and Davis (in explaining their factor analysis findings, p. 322): "Income stands in... isolation. It has an equal loading on both common factors, and not a very high one at that. Why is the amount of family income a poor measure of socio-economic status? Observation suggests that the core of status is a culturally defined, group-shared style of life, and income is a necessary but not a sufficient condition thereof. Values intervene between the receipt of a paycheck and its expenditure in consumption. A satisfied blue-collar worker and an ambitious clerk may have the same income but a different mode of living. The former is likely to have a bigger house in a cheaper neighborhood, to spend more on automobiles, to save less, and to have working-class friends and beliefs. There is a great deal of overlap and variability at precisely this point of the stratification hierarchy, and it is at this point that we have to arbitrarily de-economize our variables. Income is probably a good index at the extremes, but weakens as one approaches the great 'common man' group at the middle of our system."

It is precisely in those situations which Kahl and Davis cite in the middle class in which income fails where an index of occupation would be more successful in denoting socio-economic status.
which participated in the study, with the teachers carefully questioning pupils as a double-check.

The question of category width is a difficult one. The U.S. Census occupational groupings are obviously too wide to hold constant all, or even nearly all, environmental differences, although Duncan's findings (20) showing large differences in status characteristics between the U.S. Census categories would suggest that an important part of such differences is accounted for. As discussed above, stratification is only one of two ways in which environmental differences are accounted for, the use of an \_\_\_\_ index for the school district being the other.

**Statistical Problems**

It was necessary, finally, to deal with two important statistical problems in formulating the multiple regression model. First, and less

---

* A principal critic of the stratification scheme used by the author is Bowles: (2) "Available evidence suggests that while this technique (stratification) is certainly useful in reducing the multicollinearity among the explanatory variables, it is a thoroughly inadequate representation of non-school effects on learning." Bowles cites a study (19) in which variables measuring home environment and parent-child interaction explained large percentages of variance in achievement scores for black and white "lower class" urban children in the South. Bowles concludes that "The predictive power of dimensions of home environment within narrowly defined social strata suggests that an analysis using no other control for social environment will be subject to serious specification bias." While Bowles' basic point is valid, surely he overstates his case. Duncan's findings offer one set of counter-evidence. Also, the study Bowles cites utilized observations of individual children while in the New York data group means are utilized. This is an important difference and there is some question as to whether it is possible to generalize from one to the other.

There is nothing in Bowles' quotation that tells us how wide the group of "lower class" urban children was in the Peterson and Ne Bord work. The quotation from Kahl and Davis in the note above would suggest, moreover, that an occupational stratification scheme is at its worst in removing significant within-group socio-economic effects for low-status children. Bowles' supposed (and possibly erroneous) demonstration that stratification is inadequate for the lowest category carries no inference with respect to the other categories.
troublesome, are the differences in the expected variance of error terms from smaller to larger school districts because a different number of pupil scores were summarized in each. This difficulty was overcome with the use of a weighting scheme. The second problem exists because of the great multicollinearity which is present between most of the school and community variables germane to our analysis, a problem which already has been widely recognized in the educational production function literature. Two procedures were used to deal with this difficulty. The first of these was factor analysis, which is a helpful technique for exploring relationships between groups of collinear variables. Second, considerable experimentation was done in introducing different combinations of variables into multiple regression equations in order of contribution to the coefficient of multiple determination.

The factor analysis yielded a three-factor rotation which was disappointing in that many of the important school variables were closely associated together in the first factor, which of course merely reimplies the researcher with the essential collinearity of the data. The other two factors were more identifiable, however. The second

---

*Aitken's Generalized Least Squares. See (8) and (10) Appendix B.
** See, e.g., (16), (18), (21).
*** For examples, see (3) and (12).
† Variables loading upon the first factor, with factor loadings in parentheses, were: teacher-pupil ratio (.62); expenditure per pupil on books and supplies (.63); expenditure per pupil on health services (.66); average salary of teachers in the top salary decile (.94); average teacher experience (.53); value of school district property per pupil (.55); salary of the superintendent of schools (.79); mean teacher salary (.92); expenditure per pupil on principals and supervisors (.46); mean salary of principals (.81). The author has no good explanation for the negative loading for teacher experience, for which the expected sign is positive.
factor consisted of variables indicating resources going toward school administration-supervision, while the third included school district property variables.

The procedure of examining various combinations of variables in order of contribution to $R^2$, while highly heuristic, nonetheless yields important statistical insights which allow the researcher at least to discard variables which never contribute explanatory power to the model. If there are also reasons for the researcher to think that such variables are theoretically unimportant, he may eliminate them from the estimating model with a minimum of danger. The following list of variables were admitted as being potentially important explanatory variables for the school model.

- Index of occupation
- Teacher-pupil ratio
- Average teacher salary
- Average salary of teachers in the top salary decile
- Average salary of principals
- Expenditure per pupil on books and supplies
- Expenditure per pupil on supervisory personnel
- Value of school equipment per pupil
- Value of school buildings per pupil
- Value of total school property per pupil

*Factors, variables, and loadings were: Factor II--principal to teacher ratio (.95); principal-student ratio (.96); average teacher experience (.43), expenditure per pupil on principals and supervisors (.37). Factor III--special staff personnel to pupil ratio (.32); number of pupils per classroom (.47); value of school property per pupil (.65); value of school district buildings per pupil (.85); value of school district furniture and equipment per pupil (.42).

A test developed by Bartlett was used to choose the proper number of factors in the rotation (see (1), p. 79).
The Multiregression Model

The list just given was the basis for the explanatory equation used in the analysis. Variables for value and supervision were suggested from the factor analysis, and expenditure per pupil for principals and supervisors was included as the supervision variable. Since all of the preliminary analysis showed the three value variables to be highly correlated, only one (that for total value) was retained. The same held true for the three salary variables; it would seem that "salary policy" is a school characteristic which suffers from being divided much further. As discussed above, the salary of top teacher decile variable is thought to be superior on theoretical grounds and it also turned out to be superior in explanatory power, and therefore that variable was chosen to represent salary pol... All other variables on the list were retained in the explanatory model.

In summarizing the data, averages were computed for pupil scores on two of the five individual tests—language and arithmetic—as well as for the composite score for all five tests in the Iowa Basic Skills battery, which are work-study skills, vocabulary, and reading, besides the two just mentioned. The language and arithmetic scores were used separately because it was thought that they represent better than the other three effects which are relatively more taught in the school than at home.

*Most educators think, and those of us who attended public schools probably would agree, that arithmetic skills are taught much less in the home than language skills. But of the two language-related tests in the Iowa battery, language is a test which is directed toward the mechanics of correct expression, rather than vocabulary, and our thought was that mechanics might be relatively more a function of school than home.
To summarize the estimating model, the following multiple regression model was fitted to average pupil performance in five groupings according to occupation of principal family breadwinner for the language, arithmetic, and composite scores on the Iowa Test of Basic Skills:

\[ Y = b_1 + b_2 O + b_3 T + b_4 E_b + b_5 S_{10} + b_6 V + b_7 E_s + U \]

where:

- \( Y \) = Average achievement score of pupils in the relevant grade and occupational grouping
- \( O \) = Index of average occupation of breadwinners of pupils in grade 5
- \( T \) = Number of teachers per 1000 pupils
- \( E_b \) = Expenditure per pupil on books and supplies
- \( S_{10} \) = Average salary of teachers in the top salary decile
- \( V \) = Value of school district owned property per pupil
- \( E_s \) = Expenditure per pupil on principals and supervisors

If we recall that average teacher salary and salary of the top teacher salary decile are highly correlated, the five school inputs in this estimating equation represent most of the resources used in public schools. The index of occupation is used to capture peer-group effects and other aggregate socio-economic school effects which are not related to individual family effects which, it is hoped, are captured through the stratification scheme.

Finally, three variations of the model were used with respect to level of score as opposed to gain in score. These are further discussed with the findings below.
Findings

The most important of the fitted regression equations, from the analysis, are presented in Tables 1-3. The discussion of these findings is best organized by topic.

1. Urban and non-urban school districts

In the author's earlier paper it was found that educational expenditure levels are much less related to performance in urban and village school districts and in small school districts (less than 2000 ADA) in general. In this analysis it was decided to separate districts into urban and non-urban groupings and fit the model to each.* Upon doing this it was found that only in the urban districts are the school inputs important; the only significant variable in the village-rural districts is the occupation index.

A number of hypotheses came to mind as to why the village and rural districts exhibit such random behavior. Their small number might provide part of the explanation. This is especially true if we consider each group separately.** Also they are widespread geographically and often in non-competing teacher market areas.*** It is feasible,

* There were too few observations in either the rural or village categories to provide enough degrees of freedom for taking them separately.
** There are 12 rural and 15 village districts in the sample.
*** Turner's analysis of teacher characteristics in Indiana schools suggests that school districts are somewhat at the mercy of the area in which they are found in obtaining teachers, with the most important characteristics being availability of professional employment for teachers' husbands and cultural attractions in general. These conditions vary much more widely for rural and village districts since teachers can to a large extent commute within most metropolitan or highly urbanized areas.
<table>
<thead>
<tr>
<th>Socio-Economic Group (Occupation of Family breadwinner)</th>
<th>Intercept</th>
<th>Index of Occupation</th>
<th>Teacher-Pupil Ratio</th>
<th>Per Pupil Expenditure on Books and Supplies</th>
<th>Teacher Salary Top Decile</th>
<th>Value of School Property Per Pupil</th>
<th>Per Pupil Expenditure on Principals and Supervisors</th>
<th>Fourth Grade Score</th>
<th>N</th>
<th>m</th>
<th>s</th>
<th>n</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Professional Persons</td>
<td>2.26 (0.21)</td>
<td>1.127</td>
<td>-0.012</td>
<td>-0.0065</td>
<td>0.013</td>
<td>-0.0065</td>
<td>0.0017</td>
<td>4.4</td>
<td>2.64</td>
<td>0.36</td>
<td>0.708</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Proprietors, Managers, officials</td>
<td>2.17 (0.17)</td>
<td>1.244</td>
<td>-0.018</td>
<td>-0.020</td>
<td>0.0067</td>
<td>-0.0098</td>
<td>0.0036</td>
<td>4.4</td>
<td>2.34</td>
<td>0.31</td>
<td>0.444</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Clerks and kindred Workers</td>
<td>2.29 (0.18)</td>
<td>0.104</td>
<td>-0.024</td>
<td>-0.007</td>
<td>0.0087</td>
<td>0.0067</td>
<td>0.0013</td>
<td>4.3</td>
<td>2.15</td>
<td>0.20</td>
<td>0.296</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 and 5 Skilled and Semi-Skilled Workers</td>
<td>2.01 (0.20)</td>
<td>0.368</td>
<td>-0.018</td>
<td>-0.017</td>
<td>0.0054</td>
<td>-0.0031</td>
<td>0.0068</td>
<td>4.3</td>
<td>2.17</td>
<td>0.33</td>
<td>0.370</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Unskilled Workers</td>
<td>1.92 (0.20)</td>
<td>0.905</td>
<td>-0.015</td>
<td>-0.018</td>
<td>0.0049</td>
<td>-0.0032</td>
<td>0.0058</td>
<td>4.3</td>
<td>1.64</td>
<td>0.37</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Pupils</td>
<td>1.91 (0.16)</td>
<td>2.355</td>
<td>-0.021</td>
<td>-0.0071</td>
<td>0.0050</td>
<td>-0.0040</td>
<td>0.0050</td>
<td>4.5</td>
<td>1.38</td>
<td>0.21</td>
<td>0.551</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Statistical Significance:** * indicates significance at the ten percent level, ** indicates significance at the five percent level, and *** indicates significance at the one percent level.

*For each SES group, the first line represents the fitted equation for the gain in average performance from grade 4 to grade 6 while the second is the equation fitted to grade 6 score where grade 4 score is one of the explanatory variables. Figures under the coefficients of net regression are the standard values of t, while the figure under the intercept is the standard error of the estimate.

**Corrected for degrees of freedom lost.

*For the school districts represented in the "All Pupils" grouping.
<table>
<thead>
<tr>
<th>Socio-Economic Group (Occupation of Family Breadwinner)</th>
<th>Intercept</th>
<th>Index of Occupation</th>
<th>Teacher-Pupil Ratio</th>
<th>Per Pupil Expenditure on Books and Supplies</th>
<th>Teacher Salary Top Decile</th>
<th>Value of School Property Per Pupil</th>
<th>Per Pupil Expenditure on Principals and Supervisors</th>
<th>Fourth Grade Score</th>
<th>N</th>
<th>w</th>
<th>r</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Professional/Teachers</td>
<td>2.24</td>
<td>0.23</td>
<td>0.20</td>
<td>-0.0098</td>
<td>0.064</td>
<td>0.063</td>
<td>0.013</td>
<td>44</td>
<td>2.39</td>
<td>0.34</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.28)</td>
<td>(1.52)*</td>
<td>(1.84)</td>
<td>(1.08)</td>
<td>(0.80)</td>
<td>(5.50)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Proprietors, Managers, Officials</td>
<td>4.13</td>
<td>0.04</td>
<td>0.07</td>
<td>-0.0089</td>
<td>0.074</td>
<td>0.053</td>
<td>-0.011</td>
<td>44</td>
<td>7.41</td>
<td>0.32</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.41)</td>
<td>(1.66)</td>
<td>(1.83)</td>
<td>(1.34)</td>
<td>(0.71)</td>
<td>(4.78)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Clerks and Kindred Workers</td>
<td>1.84</td>
<td>0.12</td>
<td>0.02</td>
<td>-0.028</td>
<td>0.13</td>
<td>0.013</td>
<td>0.0035</td>
<td>44</td>
<td>2.24</td>
<td>0.28</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.93)</td>
<td>(2.06)*</td>
<td>(2.52)</td>
<td>(2.12)**</td>
<td>(1.61)</td>
<td>(5.31)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 and 5 Skilled and Semi-Skilled Workers</td>
<td>2.33</td>
<td>0.26</td>
<td>0.23</td>
<td>-0.016</td>
<td>0.059</td>
<td>0.035</td>
<td>0.015</td>
<td>44</td>
<td>2.17</td>
<td>0.21</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(2.30)*</td>
<td>(2.35)*</td>
<td>(1.75)</td>
<td>(1.47)</td>
<td>(0.62)</td>
<td>(6.80)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Unskilled workers and Servants</td>
<td>1.02</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.0029</td>
<td>0.13</td>
<td>0.062</td>
<td>-0.0062</td>
<td>44</td>
<td>2.00</td>
<td>0.25</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.98)</td>
<td>(1.04)</td>
<td>(2.57)</td>
<td>(2.61)</td>
<td>(0.85)</td>
<td>(2.35)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Pupils</td>
<td>2.16</td>
<td>0.26</td>
<td>0.13</td>
<td>-0.015</td>
<td>0.059</td>
<td>-0.0041</td>
<td>0.0026</td>
<td>44</td>
<td>2.29</td>
<td>0.21</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(1.26)</td>
<td>(1.70)</td>
<td>(1.31)</td>
<td>(1.98)</td>
<td>(0.73)</td>
<td>(9.71)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistical Significance: * indicates significance at the ten percent level, # indicates significance at the five percent level, and ** indicates significance at the one percent level.

For each SES group, the first line represents the fitted equation for the gain in average performance from grade 4 to grade 6 while the second is the equation fitted to grade 6 score where grade 4 score is one of the explanatory variables. Figures under the coefficients of net regression are the standard values of t, while the figure under the intercept is the standard error of the estimate. Corrected for degrees of freedom lost.
<table>
<thead>
<tr>
<th>Socio-Economic Group (Occupation of Family Breadwinner)</th>
<th>Intercept</th>
<th>Index of Occupation</th>
<th>Teacher-Pupil Ratio</th>
<th>Per Pupil Expenditure on Books and Supplies</th>
<th>Mean Salary of Teachers in Top Salary Decile</th>
<th>Value of School Property Per Pupil</th>
<th>Per Pupil Expenditure on Principals and Supervisors</th>
<th>N</th>
<th>m</th>
<th>s</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Professional Persons</td>
<td>7.39</td>
<td>.301</td>
<td>-.0083</td>
<td>.0052</td>
<td>.0031</td>
<td>-.0063</td>
<td>.010</td>
<td>44</td>
<td>8.19</td>
<td>0.36</td>
<td>.388</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(2.08)*</td>
<td>(0.66)</td>
<td>(0.39)</td>
<td>(0.47)</td>
<td>(0.67)</td>
<td>(3.51)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Proprietors, Managers, Officials</td>
<td>6.30</td>
<td>.372</td>
<td>.0016</td>
<td>.0077</td>
<td>-.0013</td>
<td>.016</td>
<td>.0045</td>
<td>44</td>
<td>7.62</td>
<td>0.40</td>
<td>.636</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(2.62)*</td>
<td>(1.33)</td>
<td>(0.61)</td>
<td>(0.23)</td>
<td>(2.22)*</td>
<td>(1.79)†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Clerks and Kindred Workers</td>
<td>7.14</td>
<td>.583</td>
<td>-.036</td>
<td>-.020</td>
<td>.0077</td>
<td>.014</td>
<td>.0035</td>
<td>44</td>
<td>7.32</td>
<td>0.37</td>
<td>.437</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(3.38)**</td>
<td>(2.40)*</td>
<td>(1.46)</td>
<td>(1.25)</td>
<td>(1.64)</td>
<td>(1.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 and 5 Skilled and Semi-Skilled Workers</td>
<td>7.59</td>
<td>.543</td>
<td>-.054</td>
<td>.00029</td>
<td>.0081</td>
<td>.0026</td>
<td>.0047</td>
<td>43</td>
<td>6.98</td>
<td>0.37</td>
<td>.357</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(2.58)*</td>
<td>(3.31)**</td>
<td>(0.02)</td>
<td>(1.22)</td>
<td>(0.26)</td>
<td>(1.32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Unskilled Workers and Servants</td>
<td>6.72</td>
<td>.95</td>
<td>-.039</td>
<td>-.0075</td>
<td>-.0015</td>
<td>-.0022</td>
<td>.013</td>
<td>43</td>
<td>6.68</td>
<td>0.44</td>
<td>.435</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(4.03)**</td>
<td>(2.15)*</td>
<td>(0.51)</td>
<td>(0.17)</td>
<td>(0.21)</td>
<td>(2.63)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Pupils</td>
<td>6.46</td>
<td>.796</td>
<td>-.026</td>
<td>.021</td>
<td>.0052</td>
<td>-.0022</td>
<td>.0038</td>
<td>46</td>
<td>7.45</td>
<td>0.46</td>
<td>.767</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(6.52)**</td>
<td>(2.58)*</td>
<td>(1.99)†</td>
<td>(1.04)</td>
<td>(0.33)</td>
<td>(1.59)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistical Significance: † indicates significance at the ten percent level, * indicates significance at the five percent level, and ** indicates significance at the one percent level.

*a: Corrected for degrees of freedom lost.
also, that the smallest districts are shaped much more by personality attitudes of individual administrators and teachers.

Because of the lack of relationships for the village and rural school districts, the findings presented in the tables are given only for the 46 urban school districts, and the conclusions developed in the remainder of this paper are relevant only to urban school districts.

2. School and socio-economic effects relative to the output measure chosen

As discussed above, it was hypothesized that the language and arithmetic test scores would be more highly related to school inputs (and by implication, less related to the index of occupation) than the composite score. While the language and arithmetic results were similar, school inputs do not appear to be more related to them than to the composite score in any consistent manner (indeed, the reverse is more true), and so this part of the hypothesis must be rejected. It is true, however, that the index of occupation is in general more highly related to composite score than to the other two scores, and this stands in partial support of the hypothesis. Because the language and arithmetic findings were similar, only those for arithmetic are given in the tables.

3. Use of level of achievement versus use of gain in achievement

Three variations of the model were used, and methodological importance attaches to whether the results obtained from each form were similar. In the first variant, the level of sixth-grade performance was used. In the second variant, gain in performance between grades 4 and 6 was used as the performance measure, while in the third, sixth-

* The reader should recall that this refers to the sixth-grade scores of pupils also present in grade 4.
grade level was used, but the fourth-grade score was introduced as one of the explanatory variables. Of these three, the third variant is theoretically the most satisfactory if there is much pupil mobility, since it overcomes the mobility problem—which using sixth-grade level by itself does not—and also allows for interaction effects between the fourth-grade score and other variables, which the straight subtraction does not. When there is little pupil mobility between schools, the first variant is best, since it admits six years of school effects, not just two. Tables 1 and 2 give fitted equations for the composite and arithmetic scores using the second and third variants, while Table 3 gives findings for the composite score using the first variant.

As the reader can see, there are no major differences in the results using either variant, a finding of methodological importance since future researchers, if these findings are generalizable, will be able to use level scores as a surrogate for gain. This is only generalizable to children who stay in school for the two years prior to the one studied, however, since this is how these data were summarized.* Of statistical interest is the fact that introducing grade 4 score as an explanatory variable does not explain so much of the variance that the other variables become insignificant. Indeed, the school inputs appear more significant in the two gain models than they do in explaining the level of sixth-grade performance.

4. Positive and negative school effects

Turning now to some of the more substantive findings, the school input variables divide themselves rather distinctly into two groups.

* However, use of scores for all sixth-graders instead of those who were present also in grade 4 does not greatly change the results.
In the first group are teacher-pupil ratio and expenditure per pupil on books and supplies, and these variables are negatively related to pupil performance, often at advanced levels of statistical significance. The second group includes the school inputs which, based on these data, we must conclude are those which are important inputs to quality education. No one of these variables is uniformly important in all pupil populations although one of them always is. For two populations for the composite score (managers-officials, skilled workers) all three variables are significant together.

The findings most in need of explanation are of course those for the two negative variables. Of these, the books variable—with an average expenditure of only about three percent of current expenditures ($14 per pupil)—is relatively insignificant in terms of resource use. It is puzzling nonetheless. Perhaps the figure would more properly have been averaged over a period of years rather than taken from just one year. A possible explanation might also be that school districts without the wherewithal to maintain high quality otherwise compensate somewhat by spending more on books and supplies.

Perhaps the most interesting finding in these results is that with respect to the consistent and significant negative relationship of teacher-pupil ratio to performance in these fitted equations and this despite the fact that the first order correlation between teacher-pupil ratio and achievement performance is weakly positive. (This is true for the books variable also.) Perhaps the most logical explanation for such a finding is provided by some research done by Vincent, et al. several years ago. (23). In a study of 132 school districts, these authors conclude that in all but the poorest and richest school districts
teacher-pupil ratio and salary policy are competing resources and that, when confronted with the hard choice between them, school administrators opt for salary at the expense of teacher-pupil ratio. The findings here would support such an interpretation. Note that this is not the same thing as saying that larger class size is better per se, which is nonsense. It does seem to say that within limits paying higher salaries buys more quality than changes in pupil-teacher ratio.*

Turning to the positive school variables, the consistently most important is expenditure per pupil on principals and supervisors, with the salary variable second. These relationships, if anything, emphasize the importance of resources spent for supervision (this is especially true for the extremes in socio-economic background, see below). This finding is similar to one by Turner, who found in an intensive study of teachers in Indiana that only districts with well-developed supervisory staffs were able to affect teacher behavior in desired ways. (22)

5. School effects versus socio-economic effects

If the index of occupation is at all good as a measure of the socio-economic "atmosphere" of schools (as it seems to be, heuristically, judging from its pervasively strong relationship to achievement performance, even in rural and village districts where nothing else seems to matter), then these findings seem to be telling us that both school and

*The words "within limits" in the previous sentence cannot be over-emphasized. In these New York data the range of teachers per one thousand pupils, which includes most school districts, is from 35 to 50 which implies class sizes in the range from 22 to 30.

There are always a few reported teachers who do not meet classes full time and so the figures do not convert exactly. Indeed, if reporting is not uniform, teachers per thousand would not convert to class size at all and therefore we must be careful to interpret the present variable properly. Accounting procedures and reports for school personnel in New York are quite precise, however, and any errors introduced because of reporting discrepancies are in my opinion quite small.
community variables are important. There is not much support here (at least in the urban districts) for those cynics who (sometimes because of a misreading of the Coleman report) feel that schools can do nothing to override the effects of environment. The importance of school effects are especially established by the gain models, since they—by introducing prior performance in one of two ways—probably underestimate the contribution of school, since prior performance levels which are due both to school and environment are summarily subtracted out.

6. School inputs and socio-economic background

Since the model uses a stratification scheme for social class according to occupation, there exists interesting potential for examining whether school inputs are differentially related to the performance of pupils from different social classes. Several generalizations are possible. Most striking is the fact that the supervision variable is very highly related to the performance of pupils from the highest and lowest occupational backgrounds, especially the former. (This difference is greater for arithmetic performance than composite performance, however.) But despite the importance of the supervision variable for the social class extremes, it is possible to conclude that the three positive school variables are in general much more consistently related to the performance of children in the middle of the socio-economic spectrum. This is especially true with the gain findings. Except for supervision, these schools seem better

---

* See (6).

** Admittedly, this theoretical expectation does not conform well to the fact that school variables in the gain models appear more significant. Perhaps mobility considerations for the pupils (entering the school district in grades 2 and 3 with beginning training from other districts) does indeed make for the difference.
"tuned to the wavelengths" of middle-class children than those from the highest and lowest backgrounds. This finding, especially as related to the lowest group, is consistent with what educational observers in recent years have been saying ever more insistently.

Another interesting finding is that the occupational index becomes consistently more significant as occupation level goes from high to low. The implication here is plain; peer group and other socio-economic school influences are indeed important for children from low socio-economic backgrounds.

7. Findings for all pupils taken together

The bottom fitted equation in each table is that for all pupils taken together, which was included for completeness despite the fact that the model demands stratification. Since some socio-economic effects are not removed by stratification, we might reasonably hypothesize that the occupation index is much more highly related to the scores for all pupils than for those in stratified groupings. With the composite score this is what happens as the occupation index displays levels of statistical significance which are considerably greater than those for the stratified populations. With arithmetic this is not as true on the other hand and, if not a chance occurrence, this would indicate that arithmetic is indeed more related to school and less related to socio-economic factors.

Summary of Findings

The findings just discussed can be briefly summarized into the following points.

1. The explanatory model lacks explanatory power in the 27 rural and village districts in the study.
2. The arithmetic measure is less sensitive to the socio-economic variable, but is not more highly related to the positive school inputs than the composite measure.

3. The models which use score level and those which use gain in score from grade 4 to grade 6 give similar results. School inputs are somewhat more related to gain in score than level of score.

4. Teacher-pupil ratio is consistently related to pupil performance negatively. One explanation for this is that educators within limits sacrifice class size for salary level.

The most consistently important school variable is expenditure on supervision, although the salary variable is as important for the middle-class socio-economic groups.

5. Both school inputs and socio-economic factors were found to be highly related to pupil performance.

6. Expenditure on supervisory personnel is most highly related to pupils from the highest and lowest occupational backgrounds—especially the highest. Salary and value of school district property are more highly related to the performance of children from middle-class homes.

The socio-economic index was most related to the performance of children from lower socio-economic homes.
Conclusions

In assessing the importance of the findings which are presented in Tables 1, 2, and 3, perhaps the first question to be asked concerns how far we have gone toward establishing an educational production function. The answer is, of course, "not very far." What has been demonstrated is a set of relationships within a group of 46 urban school districts during a three-year interval. We have no assurance that these districts were behaving such as to optimize the product (achievement performance) measured and, even more importantly, even if they were so optimizing, they may not have been successful at using the best methods for doing so. It might be possible to argue (in analogous fashion to students of the firm who aver that successful firms optimize profit whether consciously or not) that successful school districts optimize cognitive gains. Indeed, it is possible to speculate even more dangerously and argue that cognitive success is highly related to success in non-cognitive areas. While this might be reasonable for some groups, it doesn't seem obvious that it would be for all of them (this is a proposition which can be carefully tested and such testing is overdue). If there are joint products, in other words, and if they are competitive instead of complementary, then something more than cognitive test scores will be necessary to measure output. It should be added, however, that most policy-makers in the field of compensatory education, where deviation between cognitive and non-cognitive objectives is presumably greatest, are relying increasingly upon achievement test results as their criterion of success. A partial reason for this, in turn, may be that the testing art is much more poorly developed in the non-cognitive areas.
But there are more difficulties. The data used were aggregated such that there were virtually no detailed personnel characteristics. Also, without more of a theoretical guide from learning theory, and given highly colinear school and community variables, the empirical estimating model must necessarily be regarded as only proximate. Third, the correct unit for some of the characteristics studied is the individual school building or the classroom. Use of school district aggregates for such variables is similar to using an out-of-focus telescope.

On the positive side it should be noted that most school inputs were represented more or less completely by the estimating model and also that the data were more than merely class-sectional since a two-year longitudinal follow-up was involved. This longitudinal aspect is one of the real strengths in this data set. And, despite the problems concerning what schoolmen optimize, it nevertheless remains that this analysis has shown characteristics which go with success as measured by the most widely used measure of educational performance in existence. There are, to be sure, difficult questions raised by aggregation but these are not regarded as extreme since the 46 urban school districts represented in the findings either had only one elementary school or were reasonably homogeneous with respect to average occupation of the

*Since teachers and administrators comprise most of the meaningful educational inputs, much more detail is needed with respect to teacher and administrator characteristics. Work presently being done by Hanushek and by Levin and Michelson, which incorporates teacher performance on cognitive tests as one of the inputs, is a step in this direction. But in order to gain this kind of information the researcher usually must sacrifice some analytical breadth. Thus both Hanushek and Levin-Michelson are working with only one school district and this is harmful if there are significant between-district differences in teacher characteristics. Despite this drawback, both studies show much promise, however.
pupils' breadwinners. On balance, it is felt that these findings are important but, until replicated in other studies, should only be regarded as suggestive. *

Finally, a word needs to be said concerning policy implications of the findings in this paper. Three points are germane to this question. First, these data are relatively old and schools have changed rapidly in the past few years. Second, it would be most dangerous to extrapolate beyond the range of observed experience in the data. Third, the findings above give no clue concerning which combinations of inputs might have to be increased together to obtain desired results. For these reasons, especially the last, plus problems with the cardinality of the output measure, it was thought improper to assign efficiency (in the sense of number of months achievement performance associated with different amounts of expenditure on each input) and this has not been done. Such precision awaits better future work.

*There is undoubtedly a limit to the developing of educational production functions from cross-section analysis. More scientific would be the careful longitudinal testing of various different school inputs with carefully designed control groups. The cost of doing this, while great, would be worthwhile, and, given the fact that only 4 percent of the educational GNP is now spent on research and development, reasonable.
ACKNOWLEDGMENTS

The author wishes to thank Ulrich Ernst and H. Kirk Roberts for capable research assistance. The author wishes to acknowledge receiving helpful comments from the following persons on earlier drafts of this paper: Samuel Bowles, Lawrence Brown, James Coleman, Ulrich Ernst, Thomas Glennan, Jr., Jeffrey Green, Eric Hanushek, Robert Harrissott, Henry Levin, John Pincus, H. Kirk Roberts, and Richard L. Turner. Any mistakes are the responsibility of the author.
BIBLIOGRAPHY


5. Burkhead, Jesse; Thomas G. Fox, and John W. Holland, Input and Output in Large-City High Schools, (Syracuse: Syracuse University Press, 1967).


24. The Iowa Test of Basic Skills (Boston: Haughton-Mifflin).