DOCUMENT RESUME

ED 126 610 EA 008 553

AUTHOR Irvine, David J.
TITLE Studying School Processes Through the Analysis of School District Data.
PUB DATE 22 Apr 76
NOTE 16p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Academic Achievement; Attendance; Educational Finance; Elementary Education; *Expenditures; Models; *School Districts; *Statistical Analysis; Statistical Data; Teachers

ABSTRACT

For years the Bureau of School Programs Evaluation of the New York State Education Department has studied relationships between school processes and student achievement. The primary strategy has been (1) to use multiple regression analysis to obtain relationships, (2) to use school district average scores as the data for analysis, and (3) to eliminate, to the extent possible, differences between districts that may be attributed to socioeconomic conditions and that are not controllable by the schools. This paper briefly describes a series of analyses carried out to study relationships between a number of school processes and student achievement on third- and sixth-grade mathematics and reading tests. Process variables fell under the general headings of teacher, financial, and student. Several teacher variables were found to relate to achievement, but no consistent pattern of relationships was found between expenditure variables and achievement. The positive results for several teacher variables support logic and conventional wisdom. The higher the salary and the higher the percent of certified teachers and teachers with graduate credit, the higher district achievement is likely to be. These analyses do not reveal whether cause and effect are inherent in these relationships. (Author/IRT)

# STUDYING SCHOOL PROCESSES THROUGH THE ANALYSIS OF SCHOOL DISTRICT DATA

David J. Irvine, Chief
Bureau of School Programs Evaluation
New York State Education Department
Albany, New York 12234

Prepared for presentation as part of a symposium: Austin D. Swanson (Chairman), "Searching for School Effects Through Conventional Multivariate Analysis and Through the Study of School Outliers," Annual Meeting of American Educational Research Association, San Francisco, California, April 22, 1976.

# STUDYING SCHOOL PROCESSES THROUGH THE ANALYSIS OF SCHOOL DISTRICT DATA

David J. Irvine*

The Bureau of School Programs Evaluation of the New York State Education Department has for a number of years studied relationships between school processes and achievement. The primary strategy has been: (1) to use multiple regression analysis to obtain relationships; (2) to use school district average scores as the data for analysis; and (3) to eliminate, to the extent possible, differences between districts which may be attributed to socioeconomic conditions and other variables not controllable by the schools. Thus far, the criteria have been restricted to reading and mathematics test scores.

This paper briefly describes a series of analyses carried out to study relationships between a number of school processes and achievement. More detailed and technical information regarding specific analyses can be found in papers produced by the bureau. They are listed under References and are available on request. The present paper does not attempt to describe the technical aspects of individual studies.

For purposes of this paper, the term school processes is used to include those factors under the control of the school or school district.

---

*I wish to acknowledge the efforts of my colleagues in the Bureau of School Programs Evaluation who carried out the analyses described here and assisted in the preparation of this paper. They are Gerald H. Wohlferd, Guy D. Spath, Philip J. Pillsworth, and Gennaro DiGiovanni.

For example, expenditures are defined as processes even though they do not impinge directly on students.

## Data Sources

The New York State Education Department has an unusually large and varied pool of educational data which is updated each year through the Information Center on Education. Data of several types are collected from each school district. The data files include:

1. The Basic Educational Data System (BEDS), which contains data on each district, school, and professional staff member in the state.
2. The Pupil Evaluation Program (PEP), which each year tests virtually all students in grades three, six, and nine on reading and mathematics. Retarded students and non-English speaking students are exempt from taking the tests. Results are available for each school and each district.
3. Financial reports, which show income and expenditures of school districts.
4. U.S. Census data, which have been collected or estimated for each school district in the state.
5. School census and enrollment, which contains data for schools and districts.

## Criterion Variables

The achievement criteria used to assess the importance of school processes were district mean scores on the PEP tests:

Third-grade Reading
Third-grade Mathematics

Sixth-grade Reading
Sixth-grade Mathematics

Data for the 1971-72 and, in most cases, 1972-73 school years were examined.

Noncontrollable Variables

Several socioeconomic variables and other variables not controllable by the schools were used to control statistically for the effects of nonschool factors:

- Total Population of the District
- Percent Rural Population
- Percent of Children Living in a Father/Mother Family
- Percent of Owner Occupied Housing Units
- Percent of Population Living in Units with 1.01 Person or More Per Room
- Percent of Population Living in Units Lacking Some Plumbing
- State Aid Ratio
- Prior Achievement

Process Variables

The process variables studied came principally from the BEDS and financial files. They included:

Teacher variables:
- Percent Certified
- Percent with Graduate Credit
- Median Experience
- Percent Married
- Percent Male
- Median Salary

Financial variables:
- Total Expenditure
- Expenditure on Regular Day Instruction
- Expenditure on Teachers' Salaries

Expenditure on Central Administration
Expenditure on Principals' Salaries
Expenditure on Supervision
District Wealth (value of taxable property divided by enrollment)

Student variables:
Percent Attendance
Student Mobility

## Procedures

The analyses were carried out using multiple regression models, following procedures for generating and comparing full and restricted models. The full model for a given criterion includes a set of non-controllable variables plus the process variable of interest. The restricted model is the original equation without the process variable. The percent of variance of the criterion accounted for by the full model is compared to the percent accounted for by the restricted model. The difference is the portion of variance which can be uniquely attributed to the process variable. An F-test can be applied to test the significance of this unique variance in the following manner:

$$F = \frac{(R_1^2 - R_2^2)/df_1}{(1 - R_1^2)/df_2}$$

where: $R_1^2$ = the squared multiple correlation of the full model (i.e., the percent of variance on the criterion accounted for by all the variables).

$R_2^2$ = the squared multiple correlation of the restricted model.

$df_1$ = the number of linearly independent variables in the full model less the number in the restricted model.

$df_2$ = the number of cases or observations less the number of linearly independent variables in the full model.

Results

The teacher variables were studied using 736 school districts in New York State. This number represents about 99% of the districts in the state which operate schools at the third- and/or sixth-grade level.

The teacher variables are expressed either as percentages of the total teaching staff in the district or as medians for the teaching staff of the district.

Each full model was made up of a set of socioeconomic variables and, in some cases, prior mean achievement in addition to the teacher variable of interest. The restricted model was the same with the teacher variable deleted. An F-test was used to test the significance of the difference in $R^2$ between the two models. Results are presented in Table 1. Only those results for which the F-test was significant at the .01 level are included. This rather stringent test of significance was selected because the tendency of slight differences to be reported out of context in news media make it important not to report an excessive number of differences which could be attributable to chance or to error of measurement. In short, an attempt was made to reduce Type I errors with the resulting increase in Type II errors.

Table 1

UNIQUE CONTRIBUTIONS TO VARIANCE IN ACHIEVEMENT
ATTRIBUTABLE TO TEACHER VARIABLES
N = 736 Districts

| | Percent Certified | | | Percent with Graduate Credit | | | Median Salary | | |
|---|---|---|---|---|---|---|---|---|---|
| | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ |
| 1971 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | .4468 | .4552 | .0084 | .4468 | .4683 | .0170 | .4468 | .4709 | .0241 |
| Mathematics | .4909 | .5100 | .0191 | .4922 | .5254 | .0332 | .4922 | .5243 | .0321 |
| GRADE 6 | | | | | | | | | |
| Reading | .5066 | .5176 | .0110 | .5066 | .5391 | .0325 | .5066 | .5428 | .0362 |
| Mathematics | .3841 | .3898 | .0057 | .3840 | .3949 | .0109 | .3840 | .3983 | .0143 |
| 1972 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | NO ANALYSES DONE | | | .4229 | .4291 | .0062 | NO ANALYSES DONE | | |
| Mathematics | | | | .4517 | .4722 | .0204 | | | |
| GRADE 6 | | | | | | | | | |
| Reading | | | | .5592 | .5736 | .0144 | | | |
| Mathematics | | | | .4113 | .4214 | .0101 | | | |

Results are reported only for those variables for which the increase in $R^2$ was significant at the .01 level.

A negative relationship is indicated by a minus sign (-) following the increase in $R^2$.

Table 1 (continued)

UNIQUE CONTRIBUTIONS TO VARIANCE IN ACHIEVEMENT ATTRIBUTABLE TO TEACHER VARIABLES
N = 736 Districts

| | Percent Married | | | Percent Male | | | Median Experience | | |
|---|---|---|---|---|---|---|---|---|---|
| | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase In $R^2$ |
| 1971 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| GRADE 6 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | .3840 | .3897 | .0057(-) | .3852 | .3973 | .0120 |
| 1972 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | .4229 | .4358 | .0128(-) | .4229 | .4321 | .0092(-) | NO ANALYSES DONE | | |
| Mathematics | .4517 | .4634 | .0117(-) | .4517 | .4581 | .0063(-) | | | |
| GRADE 6 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | | | |
| Mathematics | --- | --- | --- | --- | --- | --- | | | |

Results are reported only for those variables for which the increase ir $R^2$ was significant at the .01 level.

A negative relationship is indicated by a minus sign (-) following the increase in $R^2$.

As can be observed from the table, significant results were obtained over all of the achievement variables for Percent Certified, Percent with Graduate Credit, and Median Salary. In addition, significant results were obtained in two different years for Percent with Graduate Credit. Percent Certified for 1972 was not studied because the proportion of uncertified teachers in the state had dropped substantially, drastically reducing the amount of variation observed in that variable in 1972. Percent Married and Percent Male produced inconsistent results in the two years studied. Median Experience was found to be related only to sixth-grade Mathematics. (DiGiovanni, March 1975, April 1975a, April 1975b.)

Expenditure variables were studied using 705 school districts. Financial data are collected for New York City as a single district, whereas BEDS and PEP data are collected for each of the 31 community school districts. For this reason, New York City was omitted from the analyses using expenditure variables.

The analyses were carried out in much the same way as the analyses using teacher variables. For each expenditure variable, the full model (which included the expenditure variable of interest as well as variables outside the control of the school district) was compared with a restricted model (from which the expenditure variable was deleted). An F-test was used to test the significance of the difference in $R^2$ between the full and restricted models. Results are presented in Table 2. Only those results for which the F-test was significant at the .01 level are included.

Table 2

UNIQUE CONTRIBUTIONS TO VARIANCE IN ACHIEVEMENT
ATTRIBUTABLE TO EXPENDITURE VARIABLES
N = 705 Districts*

| | Total Expenditure per Pupil | | | Expenditure on Regular Day Instruction per Pupil | | | Expenditure on Teachers per Pupil | | |
|---|---|---|---|---|---|---|---|---|---|
| | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ |
| 1971 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | .3651 | .3795 | .0144 | .3651 | .3781 | .0130 | .3651 | .3773 | .0122 |
| GRADE 6 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1972 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | .3503 | .3633 | .0130 | --- | --- | --- | --- | --- | --- |
| GRADE 6 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | --- | --- | --- | --- | --- | --- |

*Thirty-one New York City community school districts omitted.

Results are reported only for those variables for which the increase in $R^2$ was significant at the .01 level.

A negative relationship is indicated by a minus sign (-) following the increase in $R^2$.

Table 2 (continued)

UNIQUE CONTRIBUTIONS TO VARIANCE IN ACHIEVEMENT
ATTRIBUTABLE TO EXPENDITURE VARIABLES
N = 705 Districts*

| | Expenditure on Central Administration/Pupil | | | Expenditure on Principals/Pupil | | | Expenditure on Supervision/Pupil | | |
|---|---|---|---|---|---|---|---|---|---|
| | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ | Restricted Model $R^2$ | Full Model $R^2$ | Increase in $R^2$ |
| 1971 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | .3651 | .3711 | .0060 | --- | --- | --- |
| GRADE 6 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1972 | | | | | | | | | |
| GRADE 3 | | | | | | | | | |
| Reading | .2830 | .2925 | .0095(-) | --- | --- | --- | --- | --- | --- |
| Mathematics | .3503 | .3685 | .0182(-) | --- | --- | --- | --- | --- | --- |
| GRADE 6 | | | | | | | | | |
| Reading | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mathematics | --- | --- | --- | .4287 | .4360 | .0073(-) | --- | --- | --- |

*Thirty-one New York City community school districts omitted.

Results are reported only for those variables for which the increase in $R^2$ was significant at the .01 level.

A negative relationship is indicated by a minus sign (-) following the increase in $R^2$.

Few consistent relationships were found. Total Expenditure per Pupil was related to Third-Grade Mathematics in both 1971 and 1972. Several other expenditure variables were related to Third-Grade Mathematics in either 1971 or 1972, but the relationships did not hold up over both years. In all, expenditure variables were found to be significantly related to mathematics achievement in seven out of twenty-four possible relationships. Two of the significant relationships were negative. Only one significant relationship was found between expenditure variables and reading achievement out of a possible twenty-four. (Spath, January 1975.)

A major problem in studying expenditures is the gross nature of the variables. When data are aggregated at the district level, specific effects are unlikely to be observed on a specific area of achievement. The findings might be different if the data indicated how much was spent for a particular set of pupils at a given grade level for a particular set of instructional objectives. Unfortunately, these kinds of data were not readily available for the analyses described here.

A somewhat more elaborate procedure was used in studying the relationships of rate of attendance to achievement. In addition to testing for linear relationships, curvilinear relationships and interactions between rate of attendance and several socioeconomic variables were investigated to determine their contribution to the variance of third-grade and sixth-grade reading and mathematics. In no instance, was a significant relationship, linear or curvilinear, found between attendance and achievement when socioeconomic factors were controlled. Nor were any of the interaction terms found to

contribute significantly to explaining variance of any of the achievement criteria. Apparently, rate of attendance is largely a function of socioeconomic factors and, once those factors are accounted for, rate of attendance does not vary with achievement in any systematic way. (Spath, August 1974.)

A separate analysis was carried out to study the relation of student mobility to achievement. Stepwise multiple regression analysis was used rather than a model-building approach. Third-grade data, collected in 1972 from 85 districts, were analyzed. Results of the analysis failed to reveal any relationships between mobility and achievement which reached the .01 level of significance. (Spath, January 1975b.)

Conclusions

Several teacher variables were found to relate to achievement. No consistent pattern of relationships was found between expenditure variables and achievement. Neither rate of attendance nor student mobility was found to be related to achievement when socioeconomic factors were accounted for.

The positive results for several teacher variables support logic and the conventional wisdom. The higher the salaries and the percent of certified teachers and teachers with graduate credit, the higher district achievement is likely to be. These analyses do not reveal whether cause and effect are inherent in these relationships.

Nor do they indicate, if a causal relationship is present, the mechanism by which the teacher variables influence achievement.

The lack of established causality and the small (though significant) proportion of variance accounted for by the teacher variables studied here suggest caution in applying the results. If school district policy calls for hiring or rewarding teachers who are fully certified and/or who have earned graduate credit, these analyses provide some support; but that policy, we presume, is also supported by logic, experience, and perhaps other research findings. Applying these results to individual cases would be especially questionable without further evidence.

The analyses described above exemplify some of the kinds of studies which can be done using the district as the unit of observation. Some limitations of working with district data are also evident. It was because of the limits of this approach that the State Education Department stimulated additional studies which are being described in this symposium.

REFERENCES

DiGiovanni, Gennaro. The Relationship Between Percent of Certified Teachers and Student Achievement in New York School Districts. Albany, New York: State Education Department, Bureau of School Programs Evaluation. March 1975.

DiGiovanni, Gennaro. Personal and Professional Teacher Characteristics and Student Achievement in New York State. Albany, New York: State Education Department, Bureau of School Programs Evaluation. April 1975a.

DiGiovanni, Gennaro. Teacher Salaries and Student Achievement in New York State. Albany, New York: State Education Department, Bureau of School Programs Evaluation. April 1975b.

Spath, Guy D. Does Attendance Rate Make a Difference on Reading and Arithmetic Achievement? Albany, New York: State Education Department, Bureau of School Programs Evaluation. August 1974.

Spath, Guy D. Identifying Expenditure Variables Which Relate to Student Performance at the District Level. Albany, New York: State Education Department, Bureau of School Programs Evaluation. January 1975a.

Spath, Guy D. Does Student Mobility Affect School District Results on Reading and Math Achievement? Albany, New York: State Education Department, Bureau of School Programs Evaluation. January 1975b.