REGIONAL ANALYSIS STUDIES

FORECASTING PUBLIC EDUCATION EXPENDITURES

By Richard Kosobud

INTRODUCTION

I WANT TO DISCUSS a program of empirical research into forecasting or predicting selected government expenditures, and revenues. As a small step in the direction of carrying out such a program, I present later some results obtained from statistical estimation of forecasting models of public elementary and secondary education expenditures. The models are based upon the hypotheses that cost and investment relationships underlie current and capital expenditures. These results are obtained from existing time series and cross section data. However, the main fruits of the research I have in mind are likely to be obtained by the study of government budget estimates or spending intentions for specific functions and purposes. These intentions indices could be used as a tool for forecasting purposes, and as a variable for the investigation of government spending (and revenue) patterns similar to the uses of spending intentions indices obtained from surveys of consumers and businessmen. It seems to me that this is a new, promising area for application of survey research methods. I would like to develop these points in a general, preliminary fashion.

For many kinds of problems it seems convenient and useful to abstract from government activity. It may be assumed that government activity is independent of the values of variables in the private sectors of the economy. Such an assumption, for example, appears to lie behind input output models in which government spending is placed in the bill of goods or final demand side. The models are constructed as if no important feedback occurs from the private to the public economy. I am sure, however, that we could list quite a few situations in which government expenditures, and even more directly government revenues, depend upon private activity. Similarly, a full discussion of many problems calls for treating government activity in a less exogenous way.

(a) Local school officials may request estimates of school expenditures in the next period during which time they specify that the level of educational services offered to the student will remain about the same.

(b) Local government officials may request estimates of highway and road maintenance expenditures in the next period during which time they specify that maintenance standards will remain about constant.

(c) Public trust fund officials may request estimates of fund expenditures and revenues in the next period during which time they specify that eligibility criteria and benefit rates will remain unchanged (e. g., retirement and social security trust funds).

(d) Public officials may specify a preference ordering of government expenditures which depends upon the expected level of unemployment in the next period. In this last example, a complete reversal of the traditional model takes place in that budget expenditures become dependent upon variables in the private economy.

A considerable amount of research in this vein has taken place, of course. Post office, public utility, and other government activity has been incorporated into the intermediate product matrix of input output models. Econometric models have become much richer in government institutional detail.[1] Performance budgeting by government units is frequently based upon hypothesized cost relationships between expenditures and measures of output.

It would be worthwhile in this connection to draw up a list of budget items arranged in the order of their dependence upon private activity, including items which are likely to be closely related to population characteristics. Short run rather than long run dependence is of greatest interest. Heading the list in order of dependence upon private activity would be a group of tax and transfer items for which rates or benefit criteria remain unchanged during the next period. Receipts or expenditures depend in these circumstances mainly upon movements in the tax base or in the number of applicants and their characteristics. Intermediate in the list would come, I believe, a group of items which make up the services largely provided by state and local governments. These may be viewed, in many instances, as connected to measures of output by cost relations. Near the bottom of the list I would place a group of budget items which are classified as exogenous in the traditional way. I would put military and international economic assistance expenditures in this last group because of the influence upon them of world events. Also stabilization of farm income expenditures ought to be placed in this last group because of the influence of climate. An item may fall at one time in the dependent or endogenous group and at another time, when public policy with respect to the item is under review, in the exogenous group. Perhaps individual income tax collections during 1963 provide an example of an item which ought to be treated as exogenous during that period.

The middle group of budget items in our list includes many services provided by government which have been characterized in economic analysis as social goods because their consumption, it is argued, cannot be limited to in-

---

[1] An example is provided in "The Outlook for 1963, an Econometric Forecast by the University of Michigan Research Seminar in Quantitative Economics," November 15, 1962, equations 26 through 36, p. 6. (Mimeographed.)

dividuals. An interesting discussion of this subject has been initiated,[2] but the point I want to make is that the forecasting goal put forth in this study has little to do directly with the question of the desirable amounts of government expenditures and revenues; i. e., the optimum budget.[3] Since at any one time only a few budget items are under review or being discussed seriously in terms of social priorities, it may be very convenient to have estimates for the next period of other items. Certain normative criteria, for example that costs ought to be minimized for a given output, are hypothesized to underlie some forecasting models. But the issues involved in the question of the optimum budget are put to one side.

The accuracy of prediction of government budget items is not necessarily related to their dependence upon the private economy. Unemployment compensation benefits, public assistance, and other expenditures are dependent largely upon the volume and characterstics of unemployment, but predictions of the latter variable may be quite wide of the mark.[4] In the case of services provided by government, it seems clear that operating expenditures can be forecasted more accurately than capital expenditures although both may depend upon population characteristics. This conclusion is indicated in the results to be presented later in the study of public education expenditures. It may be noted that it is the short run forecasting accuracy that is of most interest.

As forecasting models of private economic activity continue to improve the question of tracing the implications of these forecasts for the public economy would seem to become more important. Even if the values of variables in the private economy are more accurately forecasted, predicting related government variables poses problems. In the case of tax or transfer items, the micro tax relation, say house owner and property tax payment, may be thought to hold exactly. However, in estimating the tax collections for even one local government unit, aggregation over many tax payers and property classifications will be required. Random events in the collection process among other disturbances are likely to enter the relationship. Consequently, there is much to be said for the introduction of probability ideas into tax and transfer functions. If highway and road maintenance expenditures can be treated as costs dependent upon measures of traffic characteristics, then these cost relationships can be expected to hold only in a stochastic sense. Other examples come readily to mind, all of which support formulation of these forecasting problems in probability terms.

Given the desirability of research into the area of the dependence of public budgets upon the private economy, it is possible to approach the task in several ways. For particular units of government time series data are available in the

[2] P. A. Samuelson, "The Pure Theory of Public Expenpitures," *The Review of Economics and Statistics* (November, 1954), pp. 387–389; and J. Margolis, "A Comment on the Pure Theory of Public Expenditures," *The Review of Economics and Statistics* (November, 1955), pp. 347–348.

[3] R. A. Musgrave, *The Theory of Public Finance* (New York: McGraw-Hill Book Company, Inc., 1959), p. 4.

[4] A forecast of increased unemployment which brings about effective public policies to increase employment is not an inaccurate forecast.

form of reported expenditures and revenues in budgets which are accessible public documents. Forecasting relationships also may be prepared from the budgets of a number of school districts, municipalities, and other local units at one point in time—a cross section. Government units may be chosen from a metropolitan area, or larger region. A problem that is being only partially solved by the passage of time is the development of data on the private economies of these areas. But a number of current studies in this vein indicate growing opportunities.[5] However, I want to raise the possibility of a slightly different approach to the same questions which has not yet been adequately exploited as far as I know.

TOWARD GOVERNMENT REALIZATION RELATIONS

Many government budgets contain estimates for the next period as well as past events. These estimates may be viewed as intentions to spend and as forecasts of receipts prepared in a manner that compares favorably with the formation of consumer and business spending intentions.[6] Let me reconsider some of the questions raised earlier in the light of this wealth of information on anticipated government activity. We may wish to think of a series of budgets for one government unit extending back in time, or a cross section of like government units in a metropolitan area, or other population.

Instead of listing budget items by their dependence, we may investigate systematically specific budget estimates and their realized values. It is possible by this comparison to answer a number of questions about the accuracy of budget estimates. Are operating expenditures for public education overestimated or underestimated? How do they compare in this respect to maintenance expenditures for roads? Do capital expenditures differ from current account expenditures in this respect? Are turning points—changes in the level of particular items over time—accurately estimated, or are decreases more accurately estimated than increases? A simple regression of budget values upon estimated values provides considerable information on this point.[7] These questions are posed in terms of the administrative budgets for which legislative approval has been secured. Evaluating differences between budgets proposed by executive agencies and budgets approved by legislative bodies would seem to pose different problems.

[5] H. E. Brazer, *City Expenditures in the United States*, Occasional Paper 66 (New York: National Bureau of Economic Research, 1959).

[6] For information on intentions data prepared by the Survey Research Center, see George Katona, *The Powerful Consumer* (New York: McGraw-Hill Book Company, Inc., 1960), pp. 81-83 and the literature cited therein. For the consumer intentions data prepared for the Federal Reserve Board see pp. 977-1003 of the *Federal Reserve Bulletin* (September, 1960). Among intentions indices prepared for business investment is the one prepared by the U. S. Dept. of Commerce and the Securities and Exchange Commission. See Murray Foss and Vito Natrella, "Ten Years Experience with Business Investment Anticipations," *Survey of Current Business* (January, 1957).

[7] H. Theil, *Economic Forecasts and Policy* (Amsterdam: North-Holland Publishing Company, 1958), pp. 64-76.

The accuracy analysis outlined above is largely a descriptive treatment of the relation of budget estimates to realized values. It provides a means of classifying and comparing budget items, but it does not provide any direct tests of hypothesized relationships between government variables and private activity. However, potentially useful information on this last matter is only a step away. A new variable may be formed, the difference between the estimated and the realized observation. This deviation variable may then be related to changes in variables in the private economy.

The advantages in the preparation of the deviation variable is that it may help achieve an important and potentially useful simplification of the budget making and decision making process.[8] Consider the following rough model. Budget estimates or intentions may be formed after taking into account the past, present, and anticipated values of a large number of variables, a quite complicated relationship. Actual budget expenditures (or receipts) may be the consequence of decisions which are influenced by a large number of past and present variables (the latter being realized values of anticipated variables). By subtracting from the actual value, the estimated value, we remove the influence of those variables which have not changed during the period. By relating the deviation variable to those variables which have moved in the private economy, we may obtain estimates of a less complicated structural relationship.[9] For example, the deviation between actual and estimated public school operating expenditures may be related to the change of enrollment during the period and to the change of other variables such as length of school term. The deviation between actual and estimated public school capital expenditures may be more properly related to the change of enrollment. These possibilities are indicated by the results to be presented in a later section.

However, research workers in this area are are unlikely to remain content with a mysterious variable like budget estimates or intentions which so far has been taken to be exogenous. For many budget items, the challenging task is to devise an adequate explanation of budget estimates or intentions for those items dependent upon private decisions. Thus we are carried more deeply into the area of study of budget decisions of public officials and their relation to the pr.ferences of constituents. It is just at this point also that it may be seen that such proposed research ought not be limited to copying budget estimates and reported values. The attitudes and preferences of public officials about future budget events are likely to prove of value in preparing intentions indices, and in investigating realization relationships.[10] Public officials may suggest variables which they have taken into account in drawing up the budget. They may also

---

[8] F. Modigliani and K. M. Cohen, "The significance and Use of Ex-Ante Data," in *Expectations, Uncertainty and Business Behavior*, ed. M. J. Bowman (New York: Social Science Research Council, 1958), pp. 51–164.

[9] The argument assumes that lagged or past values of a variable may be substituted for the values of anticipated variables, which is surely not correct. However, there are other means to obtain values of anticipated variables which would find a place in a more explicit discussion of the model.

[10] For example, direct estimates of anticipated variables could be obtained.

provide valuable explanations of budget intentions. Similarly, the attitudes and preference of constituents about future budget events provide an area for concurrent, important research.

Survery research methods seem ideally suited to preparation of these data. A panel of government units would seem to provide a good research design. Considering that there about 40,000 school systems, 40,000 municipalities, and 10,000 counties plus additional local government units, intricate sampling problems may be anticipated. If studies were designed to develop realization relationships for local governments in, for example, metropolitan areas, interesting sampling problems are still likely to be encountered. New and challenging interviewing problems may be anticipated, including the problem of locating and interviewing the relevant public officials.

If such research designs were attempted, what can we expect in terms of useful results? It may be noted that if intentions indices are prepared for specific budget items, they can be used directly for forecasting purposes regardless of whether or not the budget item is classified as endogenous or exogenous. The forecasting value of intentions in many instances, however, is likely to rest upon the relationships connecting the public economy to the private. Estimation of these relationships seems to me a strong motivation for conducting such research. Can we expect more information than is provided by an extrapolation of some simple linear trend? That is, it may be argued that most government series have exhibited a steady, upward trend.

If we compare the variation of certain government items with such volatile series as consumer durable spending and business investment spending in the period 1930 through 1960, excluding the war years, some tentative conclusions may be drawn. As a basis for comparison we calculate the coefficient of variation or the standard deviation of the series divided by the mean. Public education current account expenditures indicate the least variation with a value of .60. All state and local government expenditures combined indicate more variation with a value of .63. State and local government capital expenditures yield a coefficient of .68. Consumer durable spending reveals a coefficient of .73 and private business investment a coefficient of .88. Public education capital expenditures yield a coefficient value of 1.00. The coefficient of variation provides no information on turning points, however. Government current account expenditures reveal a ratio of turning points to observations of about .20 (time series observations have reversed direction about once in every five years on the average). Public education capital outlays reveal the highest ratio of turning points to observations or a ratio of .33. Consumer durable and business investment and all local and state government capital expenditures yield ratios of about .25.

PREDICTING PUBLIC EDUCATION EXPENDITURES

In the following discussion some exploratory results are presented which flow from treating current expenditures for public elementary and secondary education purposes as costs related to indices of the level of activity. Very experimental investment relations are estimated in which school capital con-

struction outlays depend upon enrollment variables. To the extent relationships satisfying theoretical and statistical criteria can be estimated, a basis for making conditional forecasts of public education expenditures is made available. Both time series and cross section data are utilized in these efforts. The theory of cost minimizing decisons of the firm provides a guide to the choice of variables, and of hypotheses.

The task of the decision making unit is considered to be that of minimizing total variable costs for a given well defined output subject to the constraint of the production relationship; that is, inputs are required in order to obtain any given positive output volume. The cost relationship may be derived as a consequence of rational action designed to achieve the aim of cost minimization. The adoption of the cost criterion may be stimulated by a search for profits, or by administrative prescription.[11]

It is assumed that the unit has access to full information about technical alternatives which may be depicted in a production relationship:

$$P(y_1, \ldots, y_n; \ x) = 0.$$ (1)

The y_i represent inputs of resources; and x represents an output variable. The production function specified implies technical maximization. For each level of output, a combination of inputs is utilized such that it is impossible to obtain the same output with lesser quantities of inputs.

It is assumed that there is a unique value for the output given fixed amounts of the y_i. The range of the function of interest to us is ever increasing with respect to increases in the y_i. Continuous partial derivatives of at least the second order are assumed to exist.

Total costs may be defined as any fixed costs, fixed in money terms, plus the amounts of variable factors acquired multiplied by their prices.

$$C = A + \sum_1^n p_i y_i.$$ (2)

C indicates total cost, A represents fixed costs, and the p_i stand for the prices of inputs, $i = 1, \ldots, n$. If the employer can acquire additional amounts of inputs at prevailing prices, the relation of cost to output variation will be independent of the p_i.

To reveal the consequences of cost minimizing behavior, a function may be formed from the cost definition and production function by introduction of a Lagrangean multiplier,[12]

$$Z = A + \sum_{i=1}^n p_1 y_1 - \lambda \, [p \, (y_i, \ldots, y; \ \bar{x}) = 0]$$ (3)

in which $\bar{x}$ represents a given output and the y_i are the input variables subject to control. Among the inputs or resources available for the time period under consideration are flows of teacher, equipment, and school plant services. School

---

[11] Herbert A. Simon, "New Developments in the Theory of the Firm," *The American Economic Review*, Vol. LII, No. 2 (May, 1962), p. 2.

[12] P.A. Samuelson, *Foundations of Economic Analysis* (Cambridge: Harvard University Press, 1958), p. 60.

officials may decide to utilize existing plant and equipment at varying levels, or acquire the services of new plant and equipment. It is important to know whether the flows pertain to a fixed plant being utilized at various levels, or whether they pertain to plants of varying size being utilized near capacity. In either case, however, it is the flow of input services to which attention is directed.

In the case of given output, a proper relative minimum is achieved by use of inputs to the point at which

$$\frac{\partial Z}{\partial y_i} = 0$$

or

$$p_i - \lambda \varepsilon_i = 0 \quad (i = 1, \ldots, n)$$

where ε_i indicates the marginal physical productivity of variable input i. The n equations obtained in this manner may be written

$$\frac{\varepsilon_1}{p_1} + \frac{\varepsilon_2}{p_2} + \ldots + \frac{\varepsilon_n}{p_n} = \frac{1}{\lambda} \tag{4}$$

If the ratio of marginal physical productivities of input to input price is everywhere the same, then it may be said that the marginal productivity of the last dollar of expenditure for inputs is equal in every use. There are sufficient equations to determine the use of each input in the model, and from these together with equation (1) λ may be determined.

The uniqueness and stability of the solution may be shown to depend upon second order equilibrium conditions or

$$\partial^2 Z > 0 .$$

This may be interpreted as indicating that in the neighborhood of a relative proper minimum, any substitution of one input for another will result in an increase in variable cost. If the varying input combinations which can be used to obtain a given output make up an isoquant convex to the origin, uniqueness is secured. The discussion of both points could be expressed in terms of the tangency of the convex isoquant to the lowest total variable cost line as in Figure 1.

The cost function we are seeking is implicit in equation set (4); however, it will be useful to take an additional step to indicate how variation in output may be accounted for in an explicit way. Ordinarily, decisions as to the proper volume of output and the right amounts of inputs are considered to be made by the same unit. In the case of school districts, it seems plausible to think of enrollment, considered to be one aspect of output, as beyond the influence of school officials. The problem for school officials is formulated to be that of minimizing variable costs for a given enrollment in an educational program of a given content.

A solution to the optimum amount of each input may be derived which can be shown to be a function of input prices and the given output.

$$y_i = y_i (\bar{x}; \ p_i, \ldots, p_n) .$$

These expressions for the y_i may be substituted in the cost definition (2) to obtain
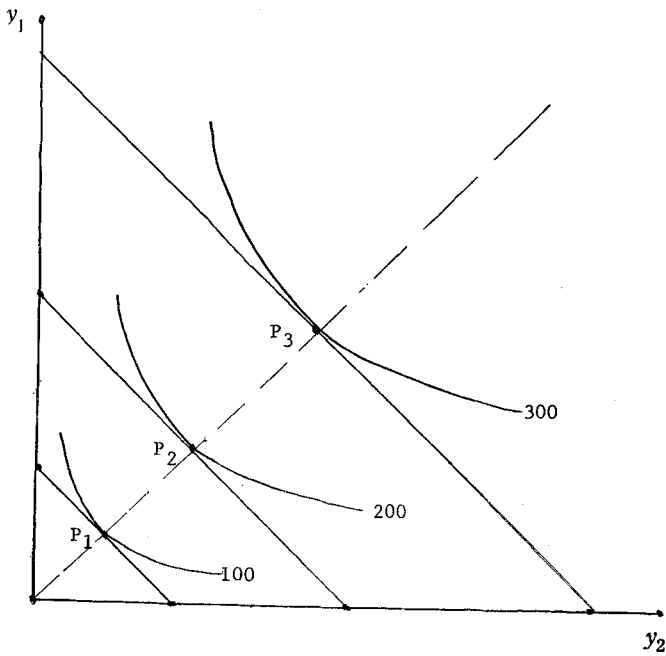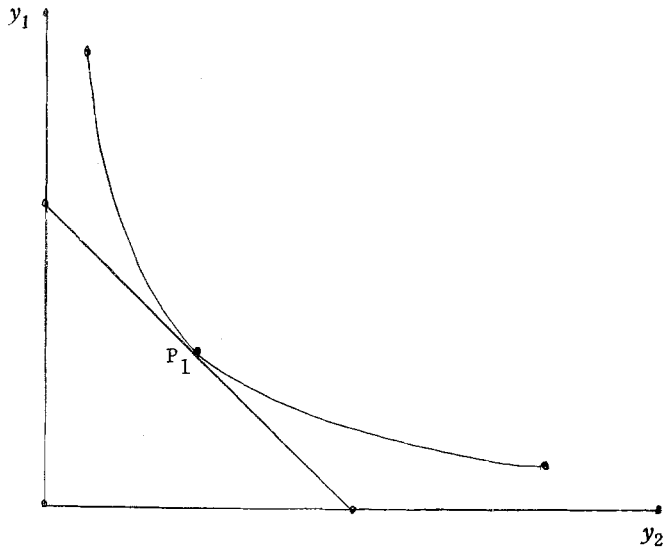
FIGURE 1. Cost Minimization Behavior for a Given Output (Above),
and for Varying Output (Below).

$$C = A \sum_{i=1}^{n} p_i y_i (\bar{x}; \ p_i, \ldots, p_n).$$

Permitting output to vary yields equilibrium solutions given by

$$C = A + C(x_i; \ p_i, \ldots, p_n), \tag{5}$$

the total cost function. If input prices remain constant, the shape of the total cost function will depend upon fixed costs and the production function, as is well known. If the total variable cost function is estimated, fixed costs play no role.

Few restrictions upon the data are contained in the model specified.[13] Marginal costs may be increasing, decreasing, or constant in the short run, as well as exhibiting all three characteristics over certain ranges of output. A smooth, continuous production function has been assumed, but this may be considered a question to be settled by empirical work. Implied in this last issue is the technical possibility of substitution of one input for another, and the adaptability of fixed factors to variable inputs. Discontinuities can be introduced without altering the discussion appreciably.[14] If factor prices vary with increasing output, marginal costs will depend upon factor prices as well as technical considerations.

As the data with which we shall work appear to come closer to providing information on the long run cost curve, it is of interest to consider what can be said beforehand about this relationship. All inputs or resources may be considered variable. The question of returns to scale becomes important in the long run situation.

It does not seem convincing to prejudge the issue of whether long run average costs ought to rise, or fall without considering the evidence in each instance. Also it would seem necessary in the long run case to evaluate whether factor prices and technical possibilities have changed and introduced extraneous elements in the long run cost curve. The idea of a well defined, unchanging commodity has been implicit in the discussion so far. In the case of the educational service, the problem of definition and measurement and the question of changes in the content of the eductional service seem especially challenging.

APPLICATION OF THE COST MODEL TO THE PUBLIC EDUCATION SECTOR

The crucial work can now begin of relating variables in the model to available data. The appropriate micro unit is not at all obvious—ought it to be school plant or school district? A quick glance at some of the descriptive data on the public education sector is sufficient to yield the impression that several trends have been underway in the size of newly constructed school plants, and in the content of the educational program. These points have a strong bearing on the results.

---

[13] J. Johnston, *Statistical Cost Analysis* (New York: McGraw-Hill Book Co., Inc., 1960), p. 24.

[14] R. Dorfman, *Application of Linear Programming to the Theory of the Firm* (Berkeley: University of California, 1951), p. 86.

It may be pointed out that the school official's life is too rich and compli-
cated even to be remotely described by cost minimization behavior. There is
little need to debate the point; it seems worthwhile to try such a model to see
how it works. Our ambition is to describe a limited range of decisions. Should
a cost model developed for private enterprise be applied to public activity? In
general, the application of cost minimization behavior to a public activity
instead of a private one doesn't seem to raise any large difficulties, that I can
see. The cost criterion function can become the object of policy, and published
policies can be examined for a statement of aims.[15] Another possible objection
to an effort to forecast public education expenditures on the basis of a cost
model may arise from the belief that other variables pertaining to education
(changes in the educational program and factor prices, for example) are not
treated explicitly in the model. In other words, variables such as enrollment
which are considered exogenous in the cost model ought to be treated as en-
dogenous in a more fully developed model. Additional comments about extension
of the cost model along this line are given in the last section.

In an attempt to obtain a degree of homogeneity of school systems, the
study has been limited to expenditures made for elementary and secondary
education purposes. About 85 per cent of the population in the age group five
to seventeen is enrolled in public schools of these levels. About two-thirds of
all expenditures for public and private educational purposes, which was reported
to total 25 billion dollars in 1960, are made in this sector.[16]

The local school district may have a number of school plants in its juris-
diction so that there is no necessary correspondence between the number of
school plants and the number of school districts. The question arises as to
what micro-unit is the appropriate one for a cost analysis. The public education
sector contains the big city system with many plants as well as the small
rural school system with one building and one teacher. Within a multi-plant
system, administrative devices such as school plant redistricting, and trans-
portation of pupils are available to school officials so that the enrollment in a
number of school plants can be varied simultaneously. In this sense, the school
district makes up a plausible microunit for cost study. Most of the expenditure
data are reported on a school district basis.

As time series data for the United States will be utilized for one set of
estimates, it is relevant to consider several trends in the size and number of
school districts and plants. Consolidations have brought about a reduction in
the number of school districts from about 130,000 in 1930 to about 60,000 in
1956. Similarly the number of school buildings has decreased during the period

---

[15] School district policies giving explicit statements of the aim of minimizing costs for
a given enrollment are easy to find. That school district officials have relevant data is a
conclusion that can be reached after looking through, for example, U. S. Office of Educa-
tion, *Financial Accounting for Local and State School Systems* (Washington: G.P.O., 1960).

[16] U. S. Bureau of the Census, *Statistical Abstract of the United States, 1961* (Washing-
ton: G. P. O., 1961), p. 118; and U. S. Department of Commerce, *Survey of Current
Business*, July 1961, Table 15 and Table 23.

by about half to 120,000.[17] Enrollment on the other hand has increased from almost 26 million to about 33 million children. The new schools which have been built have had a higher average enrollment capacity than existing or replaced facilities.

Expenditure and related data by states will be used for purposes of cross section estimation. It is important, therefore, to look into the relation of enrollment size of schools by total enrollment in the state. The association
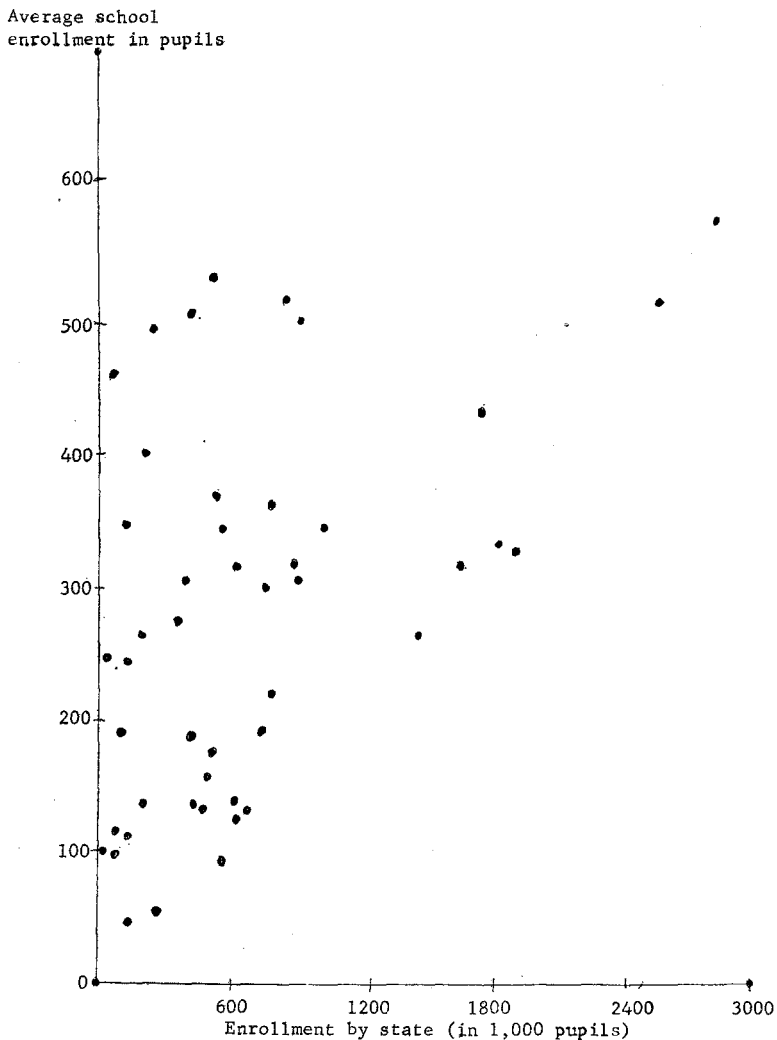

FIGURE 2. Scatter of Average Enrollment per School and Enrollment by State.

17 U. S. Office of Education, *Trends in Significant Facts on School Finance, 1929-30 to 1953-54* (Washington: G. P. O., 1957), Chart 41, p. 47.

revealed in Figure 2 is not as clear cut as one might wish, but a tendency for larger schools to be located in states with larger enrollments is visible. Among the aggregation problems such data may contain is the one brought about by the increased proportion of eligible children who have attended high school both over time and in the states with larger enrollments. It may be added in this connection that a simple summation over many types of inputs and over the school year characterize the macro variables.

A number of other variables have been prepared in order to test for the

TABLE 1. Variables Related To Public Education Expenditures*

Description of Series	Available in	
	Time Series Form (for school years ending in the even numbered years)	Cross Section Form
Independent (right hand) variables		
School enrollment, October school year (July 1, t-1 through June 30, t)	×	×
Per cent high school enrollment to total enrollment		×
Average daily attendance during the school year		×
Net national product, calendar year, 1954 dollars	×	
Personal disposable income, calender year in 1954 dollars	×	×
Population: the number of people as of July, t	×	×
Population with high school degree (per cent)	×	×
Number of public elementary and secondary schools	×	
Number of public schools with one teacher	×	
Time: the 1931–1932 school year is assigned no. 1	×	
Value of school property, January 1956		×
Price index for outlays for instruction	×	×
Price index for capital outlays	×	
Price index for operation and maintenance outlays	×	
Region $\begin{cases} 1 = \text{states with standard metropolitan areas in them} \\ 0 = \text{states without such areas} \end{cases}$		×
Length of school term, in days	×	
Dependent (left hand) variables		
Number of books circulated		×
Gross capital outlays (on land, equipment, and buildings) during the school year in current dollars	×	×
Operating outlays during the school year in current dollars	⎫	×
Maintenance outlays during the school year in current dollars	⎬ ×	×
Administrative outlays during the school year in current dollars	⎭	×
Instruction outlays during the school year in current dollars	×	×

* Sources—See Appendix 1.

A brief description of the variables is presented in Table 1. Sources for these data are listed in Appendix 1.

An important institutional feature of the education sector is the legislative requirement that children must be enrolled in some school up to the minimum school leaving age. This rule would imply that school officials must provide for all children who are enrolled. It will be assumed that school officials adjust passively to changes in enrollment which they may anticipate more or less accurately, but which they cannot influence.

To what extent can enrollment be accepted as the index of the level of service of service of public education—the output of the cost relation? In an interesting approach to estimating school expenditures using engineering type data, Isard and Coughlin observed that "... no operational measure of level of service is available."[18] It seems a workable compromise, perhaps, to start with the number enrolled, and to take into account such visible trends in the educational program as the increased length of the school term, and the increased proportion of children attending high school. By taking these trends into account we hope to correct for shifts in the cost curve due to quality changes in the educational service. An example of the problems caused by such changes is depicted on the left of Figure 3. If three paired observations
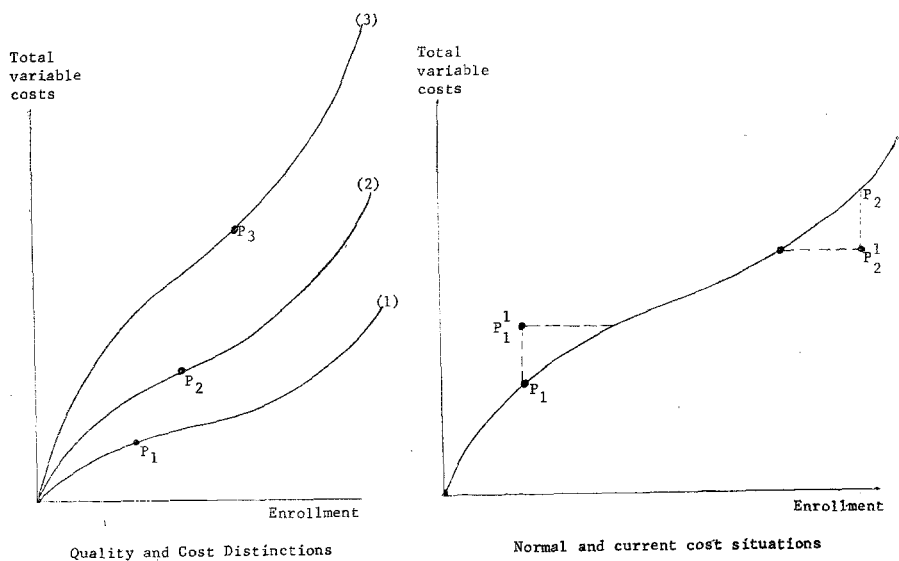

Quality and Cost Distinctions   Normal and current cost situations

FIGURE 3.

were available on costs and erollment, p_1, p_2, p_3, an enrroneous impression of the cost relation would be obtained from a least squares regression line. It is assumed that linear, additive corrections for shifts in the cost curve will be sufficiently accurate approximations for the purpose of this study.

18 W. Isard and R. Coughlin, *Municipal Costs and Revenues* (Wellesley: Chandler-Davis Co., 1959), p. 77.

influence upon the cost function of some of the trends which we have discussed.

Ideally, school accounting records ought to yield expenditures for homogeneous groups of inputs recorded over a unit of time during which enrollment and the educational program do not vary much. Observations for most variables pertain to the school year. Variation in enrollment and expenditures could occur during a shorter time period, say from half year to half year. To that extent, bias may be introduced from aggregation over time. The fact that observations are paired on a school year rather than calendar year basis introduces no fresh difficulty. A closer look at the data available on expenditures indicates that cost categories are broad and contain mixed elements. Instruction expenditures include salaries of teachers, supervisors, and principles, and expenditures for supplies of instruction such as books. School operation expenditures include those for fuel, water, light, power, as well as salaries and wages of operating personnel. Maintenance expenditures include repair of buildings and replacement of equipment and furniture.[19] The correspondence of these categories to the idea of a flow of labor and capital services and raw materials seems approximate. A reasonable initial step seemed to be to estimate a cost relationship for each of the expenditure categories separately.

No readily available price indices were available for the expenditure categories, consequently several were constructed from materials at hand. An index for school operation and maintenance outlays was prepared from a weighted combination of (a) an average annual earnings index of employees in government enterprises, (b) a price index for fuel, power, and lighting materials, and (c) a price index of metals.[20] These three indices were weighted respectively,

$$p_t = .78p_{1t} + .18p_{2t} + .04p_{3t} .$$

An index for instruction salaries was prepared from average annual salary data.[21] The implicit deflator for public construction available in the national income and product accounts was utilized to deflate capital outlay. A noteworthy characteristic of these price indices is that they moved very closely together during the sample period so that the amount of substitution of one input for another due to relative price changes may have been small.

An elementary but basic point concerns the amount of variation exhibited by enrollment, expenditures, and other variables over time and over the cross section of states. The scatter of cross section points in Figure 2 indicates considerable variation in enrollment size of public schools. As for time series data, a coefficient of variation (the standard deviation of the distribution divided by the mean) provides a rough test of the variation of variables independent of the unit of measurement. It is also of interest to calculate the number of turning points for each variable over time. For purposes of comparison, the same

---

[19] U. S. Office of Eduction, *Financial Accounting for Local and State School Systems* (Washington: G. P. O., 1960), p. 28 et seq.

[20] U. S. Bureau of the Census, *Historical Statistics of the United States, op. cit.*, Series D-707, E-31, and E-36.

[21] U. S. Office of Education, *Trends in Significant Facts on School Finance* (Washington: G. P. O., 1957), p. 60.

calculations are made for durable expenditures of consumers and private in-vestment expenditures. War years are excluded in these calculations although they were included in the estimated cost relations. As the number of obser-vations varies among the variables, the number of turning points is presented in the numerator of a ratio, the denominator of which contains the number of observations. The results are presented in Table 2.

The least variation over time is exhibited by public school enrollment; the highest by public education capital expenditures. Consumer durable and private investment expenditures are ordinarily considered to be volatile series, and they rank higher than instruction and operation and maintenance outlays on this score. Yet, differences do not seem striking. Other state and local government expenditures appear only slightly less variable during the sample period. Con-sidering the opportunities for turning points, public education capital and private investment expenditures rank highest on this count. Enrollment, and current account public school expenditures rank lowest. The fact that the enrollment variable exhibited less variation than current account expenditures may indicate that quality considerations have influenced the cost function.
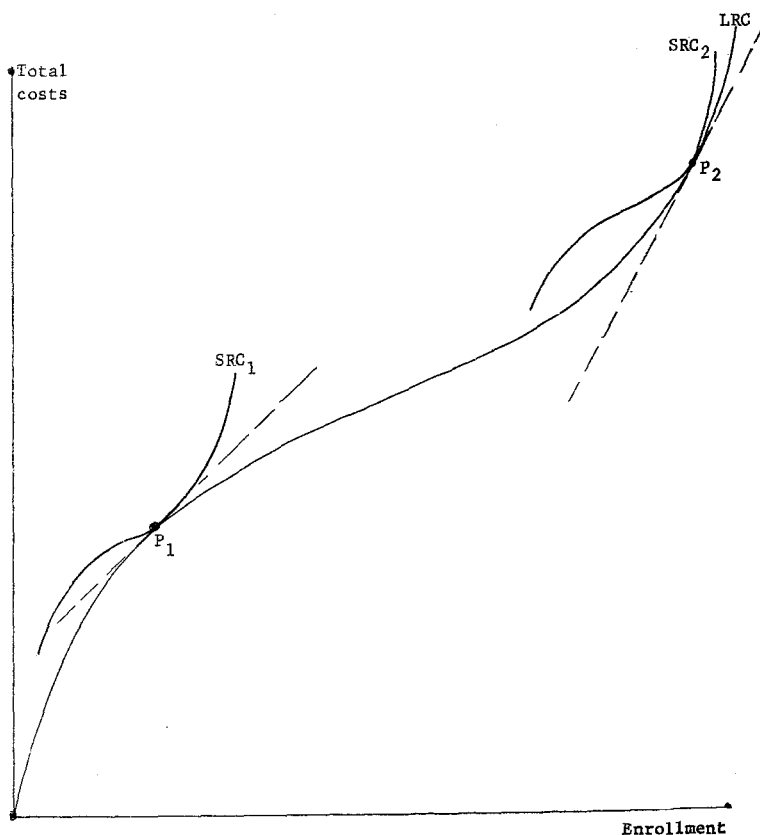

FIGURE 4. Linear Approximations to a Nonlinear Cost Relation.

ESTIMATED COST RELATIONS FROM TIME SERIES DATA

For purposes of estimation it is assumed that cost relationships do not hold exactly but stochastically. Equations linear in the variables are estimated by application of single equation least squares methods. If the long run cost curve exhibits curvature as depicted in Figure 4, the sign of the constant term in an equation linear in coefficients and variables may serve as piece of evidence indicating nonlinearity.

TABLE 2. Coefficient of Variation and Turning Points for Selected Variables.

Variable[a]	Coefficient of Variation	Turning Points (numerator) Years in Sample (denominator)
Public education enrollment	.41	2/12
Public education instruction exp. (undeflated)	.60	2/12
Public education operation and maintenance exp. (undeflated)	.60	2/12
State and local government total exp. (undeflated)	.63	5/24
State government capital exp. (undeflated)	.68	2/14
Local government capital exp. (undeflated)	.68	4/14
Consumer durable exp. (undeflated)	.73	6/24
Private investment exp. (undeflated)	.88	7/24
Public education capital exp. (undeflated)	1.00	4/12

[a] Public education data use in the calculations are from the U. S. Bureau of the Census, *Historical Statistics of the United States* (Washington: G. P. O., 1960), Series H–256, H–257, and H–259. State and local capital expenditures are from the U. S. Bureau of the Census, *Historical Summary of Governmental Finances* (Washington: G. P. O., 1959), Table 6. Consumer, private investment, and state and local government expenditures are from U. S. Department of Commerce, *U. S. Income and Output* (Washington: G. P. O., 1958), Table I-1, p. 118.

The single equation least squares estimator of the regression coefficients will yield estimates with desirable consistency and efficiency properties only under certain conditions. A major aim of this study is to search for significant variables, for which purposes the estimator chosen may be suitable. Consistency of estimates depends upon the independence of right hand variables and the disturbance term. The assumption of independence may not be a poor approximation in the cost relations as we have previously argued that decisions to enroll children in school are made by families, while decisions to spend are made, in the last instance, by school officials. As for another independence assumption, that disturbance terms are distributed randomly over time, a serial correlation statistic may be estimated for purposes of evaluation.

All variables except time have been prepared in index number form with 1954 equal to 100. The regression coefficients may therefore be taken as approxi-

TARLE 3. Linear Cost Relationships Estimated by the Single Equation Least Squares Method: Public Elementary and Secondary Schools.

Left Hand or Dependent Variable	Constant term	Enroll-ment	Length of school term	Per cent of population with high school degree	Time	Number of schools	Number of rural schools	$\bar{R}^2$	Serial correlation number
Deflated instruction outlays	21.379 (13.170)	.918 (.069)	.029 (.088)			−.148 (.016)		.971	
Deflated instruction outlays	−14.210 11.979	.854 (.134)		.234 (.103)	.202 (.188)			.885	
Undeflated instruction outlays	−192.407 (52.991)	1.798 .594		.793 (.458)	1.070 (.833)			.769	
Deflated operation and maintenance outlays	−143.115 (27.000)	1.276 (.198)	.906 (.213)		.904 (.173)			.941	1.836
Deflated operation and maintenance outlays	−148.656 (23.774)	1.272 (.178)	1.328 (.203)				−.082 (.013)	.922	1.385
Deflated operation and maintenance outlays	−122.915 (40.263)	1.272 (.295)	.481 (.332)	.427 (.148)				.797	
Undeflated operation and maintenance outlays	−307.677 (86.670)	1.909 (.641)	1.019 (.724)	1.072 (.323)				.757	

mations to elasticity coefficients[22] A number of equations have been selected for presentation in Table 3 so that each equation may be read across the row.

The cost relation approach to explaining public school expenditures yields, in general, high correlation coefficients all of which are significant according to the conventional F test. The importance of the enrollment variable in explaining expenditures of all types emerges clearly. The stability of the regression coefficient of the enrollment variable in the various equations is noteworthy. However, significant contributions to an explanation of total expenditure variation are made by other variables. Their role according to one interpretation is to allow for shifts in the cost curve due to changes in the content or quality of the educational service received by the student.

The length of the school term increased appreciably during the sample. When this variable is introduced in the cost relation explaining instruction expenditures, the hypothesis that its associated coefficient is zero cannot be rejected. However, the role of length of term is significant in the equations with operation and maintenance expenditures. A variable "per cent of population with high school degree" was introduced in index number form. It was significant in both deflated instruction and deflated operation and maintenance relationships. In this case, the elasticity of expenditures may be taken to be less than unity. A better and more direct approach to this issue, however, would be to prepare a variable measuring the number or proportion of children attending high school. Such a variable was tested in the cross section study. If cost relations for elementary and secondary education have different slopes, an alternative procedure would be to estimate them separately and sum predictions for each subsector.

If it is believed that the quality of education tends to be better in the larger public schools and districts, then, an increase in the proportion of enrollment in these schools (which has occurred) would imply in this model an increase in the quality of the educational service. Several variables were prepared to test for this shift: one measures the number of school plants; and the other measures the number of rural schools. The results reveal that both variables take on negative signs in the deflated instruction and operation and maintenance equations. As these variable declined in value markedly during the sample period, the signs are in accord with an expectation of an inverse relation between the number of small schools and the quality of the educational program.

An implication of the cost model outlined at the start is that the optimum combination of inputs which results in cost minimization for a given output will not be changed if all input prices are altered by the same proportion. General movements in price levels will introduce variations in money expenditures but not into the flow of inputs. This line of reasoning suggests that deflated cost relations ought to have higher correlation coefficients than undeflated cost relations as extraneous price fluctuations are removed. This was in fact the outcome as may be seen by inspection of Table 3. In an actual forecasting application, estimates of input prices would have to be prepared independently.

---

[22] The approximation is valid only in the neighborhood of 1954.

Introduction of a time variable is sometimes interpreted as an indication that shifts in the cost function have occurred regularly during the sample period. This possibility may be rejected in the case of the deflated instruction equation, but it is consistent with the result in the deflated operation and maintenance equation. Even in the latter case we may wish to interpret the result as a sign that undetected variables have exerted an influence. A better way to explore the latter question is to test disturbance terms for serial correlation.

Failure of the disturbance terms to be randomly distributed over time will impair our confidence in the estimated standard error of the regression coefficients even in large samples. The evidence of positive serial correlation of disturbance terms may be interpreted as a sign that variables which properly belong in the cost equation have been left out. This last possibility is of special interest in this study because of the difficulties of measuring the quality of the educational service. It may be seen in the last column of Table 3 that the disturbances terms are more randomly distributed over time in the first equation with deflated operation and maintenance expenditures than in the second in which positive serial correlation is present according to a one tail test[23] As the first equation contains the variable time and therefore positive serial correlation may have been removed in an unsatisfactory way, the conclusion is drawn that not all variables influencing costs have been specified.

If the cost curve exhibits curvature, the sign, and significance, of the constant term may be interpreted as providing information about the range of the curve estimated in the sample period; i. e., whether increasing or decreasing marginal costs prevail. The negative signs suggest increasing marginal costs; however, the constant terms in the case of deflated instruction expenditures may not differ from zero. Several equations were fitted in which higer powers of the enrollment variable were introduced. The signs attached to these variables seemed implausible. Several equations in logarithmic form were fitted with no improvement over equations with linear variables. A cautious attitude with respect to nonlinearity seems defensible. A nonlinear macro relation may be the result of improperly aggregated variables. It is possible that the situation is one in which linear macro rolatious have different slopes; say, for example, linear high school and elementary school cost relations.

ESTIMATES OF COST RELATIONS FROM CROSS SECTION DATA

The cross section data to be exploited at this juncture provide opportunities to extend, and to compare the results obtained from time series data. The

---

[23] The serial correlation ratio was estimated as

$$\delta^2 = \frac{\sum (\widehat{U}_t - U_{t-1})^2}{\sum \widehat{U}_t^2} \cdot \frac{N}{N-1} ,$$

where $E(U) = 0$, and $\widehat{U}_t$ equals calculated expenditures minus estimated. The hypothesis that the residuals are randomly distributed over time cannot be rejected in the first mentioned equation, but can be rejected in the second at the 5 per cent level of significance. H. Thiel and A. L. Nagar, "Testing the Independence of Regression Disturbances," *Journal of the American Statistical Association*, Vol. LVI (December 1961), Table 2, p. 802.

intention is to compare the significance of variables and the magnitude of parameters as if the time series and cross section situations provided independent information about the same underlying relationship.

Cross section data are considered to provide information on the long run cost relationship.   The units of observation for these data are states for the school year 1957–58.   It may be assumed tentatively that short run influences tend to exert themselves in like fashion upon the school systems so that the data reflect long run adjustments.   As pointed our earlier, there is evidence that the average size of school system, and school plant, increased as the number of pupils in public education increased in the state.   State data may be thought to reflect some representative or average school system.
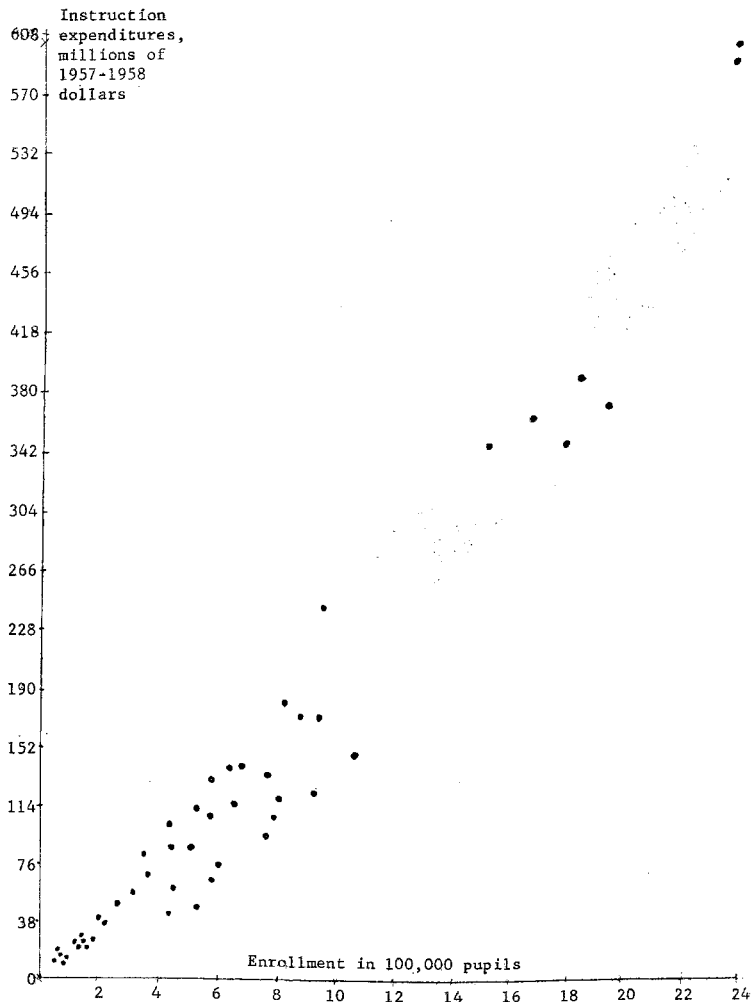

FIGURE 5.   Scatter of Enrollment and Instruction Expenditures, by States, for the School Year 1957-1958.
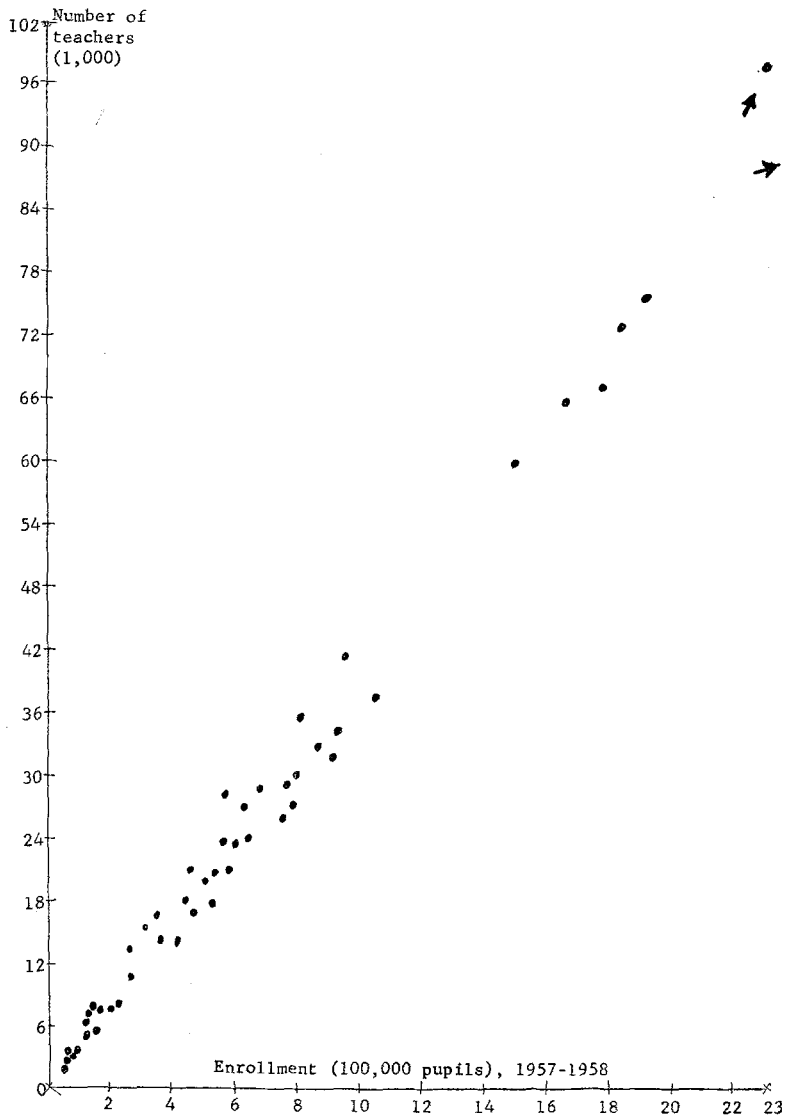
FIGURE 6. Scatter of Number of Teachers by Enrollment, by State.

It has sometimes been assumed that some variables are held fixed in the cross section situation; for example, the prices of inputs may be assumed to be the same for all systems. A simple test of the assumption that prices of resources are the same to all school systems regardless of their location by state is provided in Figures 5 and 6 in which first the number of teachers is plotted against enrollment, and second, instructional outlays are plotted against enrollment. Roughly, more unexplained variation is indicated in the latter. Variation in the salaries of comparably trained teachers by state may be influences

at work which could account for these differences. Although the assumption of constant price cannot be made aggressively in the cross section situation for these reasons, it appears to be worth while to explore further the consequences of such an assumption.

A number of major variables in the cross section study are identical in concept to those in the time series work. Expenditures are measured in dollars and enrollment in the number of children, however. The major differences occur in the variables which represent quality or level of education. Data to represent length of term, number of schools, and population with high school degree are replaced by data on number of books circulated and a dummy variable which separates states containing some part of a standard metropolitan area from those which do not. The former is offered as a measure which may correspond to the proportion of population with high school degree, and the latter may also reflect the higher level of educational service available in in metropolitan areas. Also a variable measuring the per cent of students enrolled in high school is introduced.[24]

In some cost relationships, a size of plant variable has been placed in the cost equation in order to eliminate the variation in unit costs attributable to economies or diseconomies of scale. The stock of capital measure introduced is not considered such a size of plant measure. It is considered an indicator of variation in the quality of the educational program, rough as it is.

A general appraisal of cross section results presented in Table 4 reveals a strong resemblance to the results from equations fitted to time series data. Correlation coefficients tend to be similar in magnitude to those of time series equations with deflated expenditures, which is the most appropriate comparisor if the assumption is correct that school officials are confronted with the same input prices. Again, the enrollment variable dominates the cost relation with significant contributions to costs being made by other variables. In order to make the regression coefficients of the enrollment variable more comparable to time series coefficients, they have been converted to elasticity coefficients and presented in the next to last column.

The first equation is the result of a regression of total costs upon enrollment and a regional variable. The latter takes on the value 1 if a standard metropolitan area is located in the state and 0 if none is located in the state. The average size of school facility and district is larger is the metropolitan areas and a possible interpretation of the significant coefficient attached to this varable is that a different, and higher quality, educational program is offered in these areas. In this case, the elasticity of response of total costs to enrollment is 1.14; however, no confidence interval has been estimated for the coefficient.

Considering instruction expenditures, it is apparent from equation (2), that the enrollment variable by itself is capable of accounting for a large share of the variation of instruction expenditures during the sample period. In equation (3), we introduce a regional dummy and a variable representing the number of books circulated in the state. The influence of the former cannot be said to be significant with much confidence. The latter is much more significant. It

---

[24] Sources for these data are presented in the Appendix.

TABLE 4. Cost Relations Estimated by Least Squares Method: Cross Section Data by States: Public Elementary and Secondary Education 1957-1958 School year (in 1957-1958 dollars)

Equation number	Cost variable	Constant term	Enrollment	Enrollment squared	Region	Value of school plant	Books circulated	Elasticity Coefficient	$\bar{R}^2$
1	Total costs	−38.109 (9.842)	314.515 (13.345)		42.605 (16.379)			1.14	.947
2	Instruction	−3253.582 (912.414)	252.440 (96.710)					1.23	.935
3	Instruction	−2350.567 (775.564)	172.495 (17.320)		8.577 (13.077)		4.310 (.938)	.84	.959
4	Operation	−325.889 (122.515)	42.954 (5.379)			2.428 (4.202)		1.57	.943
5	Operation	−5.493 (1.683)	30.818 (2.045)		9.612 (2.765)				.892
6	Maintenance	−154.425 (63.804)	13.481 (.676)					1.10	.894
7	Maintenance	−96.263 (45.432)	7.120 (1.014)		1.808 (.766)	.319 (.050)		.63	.952
8	Administration	−292.663 (100.350)	17.447 (1.064)						.849
9	Instruction	1656.680 (817.618)	104.528 (18.561)	6.052 (.717)					.974
10	Operation	175.022 (232.702)	15.278 (5.283)	.782 (.204)					.894
11	Maintenance	53.706 (80.864)	7.211 (1.836)	.256 (.071)					.914

was introduced in the hope that it might be a consequence of a higher quality educational program. The elasticity of response of instruction expenditures to enrollment (.84) is remarkably close to the comparable time series regression coefficient (.85). For the operating cost relation in equation (4), a variable representing the value of school plant was introduced in the expectation that it would reflect differences in quality. It cannot be said to play a significant role according to our results. In equation (5), a variable representing region is introduced and it is significant. There remains the difficulty in the case of this variable of distinguishing quality from price aspects.

In the case of maintenance expenditures in equations 5 and 6, a considerable reduction in the importance of the enrollment variable occurs as additional variables are introduced into the equation. Both the regional dummy and the value of school plant appear significant. According to our interpretation both are indicators of a quality shift. A direct comparison of the response of operating and maintenance costs to enrollment between cross section and time series results cannot be made because the two cost categories were grouped in the latter study. If cross section elasticity coefficients for each cost type are weighted by the respective regression coefficient, an estimate of 1.4 is obtained for combined expenditures. This may be compared to a regression coefficient of 1.3 in the case of the equation fitted to time series data.

It may be noted in equation 8 that administration expenditures are also highly correlated with enrollment.

A variable measuring the ratio of high school students to all students was prepared with the following result:

$$\text{total costs} = -173.76 + 304.10 \text{ enrollment} + 5.32 \text{ per cent}$$
$$(91.96)(13.92)(3.92)$$
$$\text{high school students} + 36.49 \text{ region.}\quad \bar{R}^2 = .938.$$
$$(19.70)$$

The sign of the coefficient of the variable is correct; but it cannot be said to be different from zero with great confidence. A more satisfactory treatment of this problem may be to estimate cost relationships separately for elementary schools and for high schools.

Can anything be said about the shape of the cost function? If the inflected total cost curve is maintained to be the underlying cost relationship, then the negative and significant constant terms in all the cross section equations indicate that marginal costs were increasing during the sample period. In view of these results, it seemed worthwhile to experiment with equations in which the enrollment variable was permitted to exert a greater influence. These efforts are presented in equation (9), (10), and (11) for the component cost categories. In each case, the coefficient of the squared enrollment term is significant, and its introduction results in a change of the constant term. While the evidence for curvature seems more clear cut in the case of the cross section study, judgment seems best suspended until more disaggregated cost relations are estimated.

It has been argued that cross section studies of decision making behavior are liable to the regression fallacy.[25] The argument may be put, perhaps, in

the following form: decision makers may have certain anticipations as to the level of sales, or enrollment. On the basis of these anticipations, assignments of costs to time periods are made. If the actual situation turns out differently than anticipated, a number of cost elements will not have been correctly assigned. To return to Figure 3 on the right, the situation portrayed at point p_2^1 is an enrollment larger than anticipated, and total reported costs are lower than they should be. If there is a systematic tendency for school systems with low enrollment to be in the situation of point p_1^1 instead of p_1, and systems with large enrollment to be at p_2^1 instead of p_2, multiple regression estimation methods may yield, if a linear model is adopted, an intercept too large, and slope too low. A bias toward linearity is also indicated. The existence of a regression fallacy would bias the cross section results. Such systematic biases in the responses which the data are to measure ought not to appear in the time series data.[26] These short run uncertain events may be assumed to cancel out. A test of the presence of "regression fallacy," and of the usefulness of this concept in general is obtained by comparison of time series and cross section estimates of the same relationship. The close agreement of elasticity coefficients in our results is evidence that the "regression fallacy" is not of importance in this case.

PUBLIC EDUCATION INVESTMENT RELATIONS

It has been calculated that the coefficient of variation was higher for school capital outlays than for cost outlays during the sample period. In fact, capital outlays for schools exhibited more variability on these counts than consumer durable and private investment expenditures (all in current dollar series). This suggests that uncertainty is likely to play a larger role in the investment decision. Several possible interpretations of capital outlays were worked out for empirical testing, all in a very experimental manner as there appear to be few guideposts in this area.

A possible consequence of the institutional provision for public education up to the minimum school leaving age is that school districts may be considered to stand ready to provide school plant for each pupil in some constant amount. This rule for investment data may be written

$$\frac{K_t}{E_t^{?}} = a_o v_t \tag{6}$$

in which K represents capital stock, E represents enrollment, a_o represents a constant, v a disturbance term. If a_o is a constant, then,

$$\frac{K_{(t-1)}}{E_{(t-1)}} = a_o v_{t-1} \tag{7}$$

which may be subtracted from (6) and rewritten

$$K_t - K_{(t-1)} = a_o[E_t v_t - E_{t-1} v_{t-1}]. \tag{8}$$

[25] G. H. Borts, "The Estimation of Rail Cost Functions," *Econometrica*, Vol. XXVIII, No. 1 (January 1960), p. 112.

[26] *Ibid.*

The left hand side of relationship (8) may be defined to be net investment. Changes in enrollment are taken as independent of the disturbance term. This relationship could be estimated in logarithmic form.[27] Additional variables may be introduced plausibly. It may be reasoned that local school officials will not decide to build new schools until old ones are filled to capacity. If the investment data are, in fact, gross rather than net, some estimate of replacement or depreciation would seem to be necessary. These refinements could not be introduced into the model because of the lack of suitable macro data. It was possible to prepare a stock of capital proxy on a cross section basis.

An alternative formulation of the investment decision is possible. School officials may repond to changes in enrollment in formulating investment plans (which may be presented to the community, state governments, etc.). These plans are based, not only upon the recent past, but also upon previous changes in enrollment. However, events in prior years are given diminished weight in drawing up plans. This interpretation may be written as

$$I_t = C_o + C_1(E_t - E_{t-1}) + C_1 k[E_{(t-1)} - E_{(t-2)}]$$
$$+ C_2 k^2 [E_{(t-2)} - E_{(t-3)}] + \ldots + C_m k^m [\ldots] + U_t , \qquad (9)$$

where $k < 1$. This equation may be simplified for purposes of estimation, as is well known, by multiplying the prior year's relation by k and subtracting the result from the original so as to obtain

$$I_t = d_o + d_1[E_t - E_{(t-1)}] + kI_{t-1} + V_t \qquad (10)$$

where V_t is a composite error term.

Estimated equations embodying these and other more pragmatic efforts are presented in Tables 5 (time series data) and 6 (cross section). The dependent variable is gross capital outlay for public schools which is deflated in 1954 dollars in the case of the time series variable. Other variables are familiar from the discussion of cost relations. The value of school property, as reported by school district officials on varying valuation bases, is introduced into the cross section relation in the expectation that it would remove from gross investment that portion which replaces used up school plant.

Estimated investment relations exhibit lower correlation coefficients than estimated cost relations. Serial correlation ratios also indicate a less random distribution over time of the residual terms (Table 5). These general results are interpreted to mean that undetected variables exerted an influence during the sample period, and our specification of investment relations is incomplete. However, the rationalizations of the investment decision which were attempted yielded better results than a pragmatic regression of capital outlays upon enrollment.

In equation (1) of Table 5, the coefficient of a total enrollment variable is significant. The sign of the variable "number of rural schools" is negative

---

[27] The error term is considered a multiplicative disturbance which means that it is thought that disturbances tend to increase in absolute amount as the exogenous variable increases, but not in relative amount.

TABLE 5. Public School Investment Equations: Time Series Data.
Index Numbers; 1954 = 100, Constant Dollars

Equation number		Constant	Enrollment	No. of rural schools	Change in enrollment	Population with high school degree	Capital outlay $(t-2)$	Enrollment $(t-1)$	No. of school plants	Length of term	$\bar{R}^2$	Serial Corr. No.
1	I	−203.847 (134.015)	3.063 (1.474)	−.119 (.072)							.332	.648
2	I	31.660 (4.431)			7.692 (1.123)						.780	1.030
3	I	.0018 (.0027)			7.295 (1.413)	.169 (.344)					.765	.991
4	I	15.809 (6.624)			5.166 (1.030)		.494 (.175)				.861	1.360
5	I	−120.797 (67.387)			6.898 (1.030)			1.735 (.766)			.836	1.150
6	I	−140.232 (74.149)			6.390 (1.272)	.210 (.295)		1.770 (.785)			.829	1.180
7	I	−237.981 (112.738)	3.300 .588						−.316 (.138)	.321 (.757)	.804	

TABLE 6. School Investment Equations Estimated by Single Equation
Least Squares Methods
(1957-1958 school year, in current dollars)

Equation number		Constant	Enrollment	Region	Books circulated	Change in enrollment $t - (t-2)$	Enrollment $(t-2)$	Value of school property	$\bar{R}^2$
1	I	−13.262 (7.966)	38.754 (17.790)	−8.302 (13.431)	4.766 (.964)				.844
2	I	.350 (6.740)				1196.941 (81.400)			.821
3	I	−.707 (7.089)		7.122 (13.653)		1170.220 (96.726)			.818
4	I	−12.492 (7.908)				854.859 (147.817)	45.919 (16.991)		.825
5	I	−1895.179 (778.395)	1.824 (26.704)					159.58 (34.176)	.836
6	I	−1526.060 (580.258)				605.747 (119.037)		92.709 (15.918)	.873

which was expected as this variable was introduced to measure new construction brought about by the shift in enrollment from rural to urban schools. However, the coefficient cannot be considered to be different from zero with much confidence. The low correlation coefficient and presence of positive serial correlation indicate the deficiencies of this pragmatic equation which resembles the cost relation.

In equations (2) through (6) are presented the results of variations of the accelerator, and accelerator with distributed lag models which were outlined at the beginning of the discussion. The change of enrollment variable emerges as a dominant, stable factor in the investment relation. In equation (4), the distributed lag version of this relation yields the best results in terms of correlation coefficient and serial correlation number. However, the values of the serial correlation numbers lead us to reject the view that disturbance terms are independently distributed over time; this rejection is according to a one tail test for positive serial correlation. Efforts to distinguish high school from elementary school outlays with the variable measuring per cent of the population with high school degree did not work well.

Investment relations prepared on the basis of cross section data are not inconsistent with time series results with respect to the significance of variables entering into these relationships. The change in enrollment variable appears to be the most important. Introduction of the variable measuring the total stock of existing school plant brings about an important modification if we are prepared to accept this variable as an index of the demand for capital replacement. In equation (7) of Table 6, the value of school property variable may be interpreted to represent the demand for replacement of existing worn out plant. If the stock of capital variable is replaced by a lagged enrollment variable as in equation (4), it can be seen that the enrollment variable is not a completely adequate substitute for the stock of capital variable.

A regional variable costructed as in the case of cost relations does not perform well. In equation (1) a variable representing quality of education (books circulated) is introduced, and exerts a positive influence. However, such a roundabout index of quality seems unsatisfactory. An experiment was performed in which a more direct measure of the per cent of students in high school was inroduced.

$$\text{Investment}(t) = -61.31 + 45.35 \text{ enrollment}(t) + 860.01 \text{ change in}$$
$$(58.71) \quad (17.06) \qquad\qquad (148.42)$$
$$\text{enrollment}(t) + 2.09 \text{ per cent enrolled in high school}(t). \quad \bar{R}^2 = .84.$$
$$(2.49)$$

The coefficient of the variable has the correct sign, but it cannot be held to be different from zero with much confidence.

In one respect, the results differ as between time series and cross section equations. An elasticity coefficient prepared from cross section data indicates that about a one per cent increase in the change in enrollment is associated with a one per cent increase in investment outlays. This elasticity coefficient in the case of the time series relations is in the neighborhood of .25. This comparison suggests that the observable time series data may in fact reflect

short run considerations. When these influences are held more or less constant in a cross section study, the response of public education to changes in enrollment approaches the long run coefficient. An important consequence of this difference is that utilization of the cross section coefficient in time series forecasts without modification would greatly exaggerate short run fluctuations. The sample time period has a number of aspects which may have made long run investment adjustments difficult to carry out. It may be noted that the great depression, the war and immediate postwar period have been considered a period of extremely restricted public education investment plans and restricted responses to changing circumstances in the opinion of a number of public educators.[28]

CONCLUSIONS AND FURTHER WORK

To what extent has a basis been prepared for making conditional forecasts of expenditures in the public education sector? Cost and investment relations satisfying statistical and theoretical criteria have been estimated from time series and cross section data. However, many questions remain. The macro relations which have been estimated may not correspond in desirable ways with micro cost and investment relations. More ideal data for purposes of aggregation may be obtainable from individual school district and school plant units. Micro variables for which observations are available over periods of time, or cross section studies of a number of units would appear to hold out promise. Distinctions between high school and elementary school units could
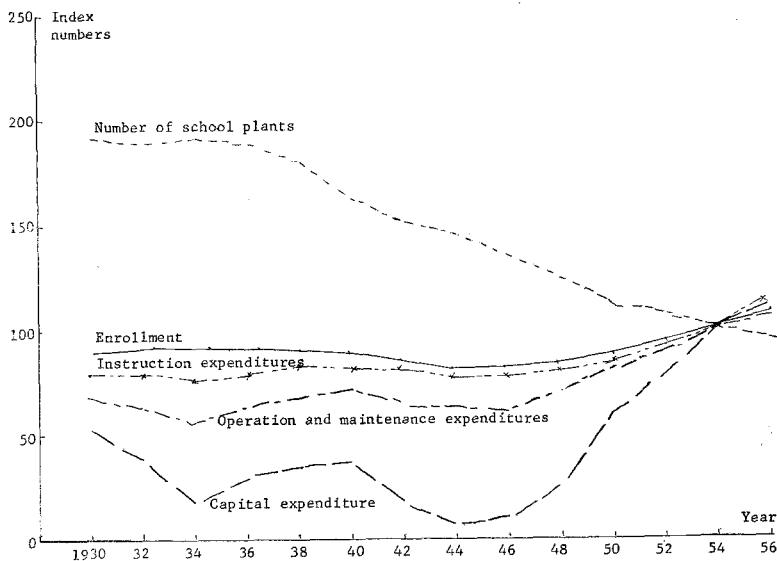
FIGURE 7. Public Elementary and Secondary Education Data. Index Numbers, 1954 equals 100. In 1954 Dollars.

[28] Comment by D. Willard Zahn, Dean, Department of Education, Temple University.

be made in these studies, and tests carried out to ascertain, for example, if cost relations have the same slopes. A more satisfactory treatment of the problem of changes in the content of the educational program ought to be possible at these levels. An unsatisfactory part of the macro studies which have been carried out is the inclusion of war year observations in both cost and investment relations estimated from time series data. It can be seen from Figure 7 that some unexplained variation was introduced in this way, especially in the investment relations.

While expenditure variables have been treated as endogenous, a number of other variables, for example the indices of changes in the educational program, have been introduced as if they were exogenous variables. In an actual forecasting application, future values of these variables would have to be estimated. A more significant question is raised by this point, however. A sector model ought to incorporate explanations of variables such as length of term and per cent of students in high school, among others. Attention has been paid to only one aspect of public education activity—the cost side, so to speak. On the other, or demand side, it would appear that income variables could be introduced as explanations of changes in the content of the educational program. As local school units receive revenues for purposes of providing an educational program from bond proceeds and tax rates on real property values, it would seem plausible to relate tax rates and bond proceeds to income measures in order to see if these variables can help explain changes in the educational program received by the population.

APPENDIX 1. DATA AND SOURCES

Cross Section Variables as Listed in Table 1

a) Population by state of 7/1/57, and 7/1/55. These observations exclude members of the armed forces stationed overseas. Sources: U. S. Office of Education, *Statistical Summary of State School Systems*, 1957–1958 (Washington: G. P. O., 1960), Table 2, p. 8.

b) Enrollment in public elementary and secondary day schools. A number which indicates at the end of the year all the different pupils enrolled in the state. The total, therefore, includes a double counting of pupils who are enrolled during the year in two or more states. The U. S. Office of Education estimated net enrollment U. S. A. in 1957–1958, to be almost a million pupils less than the stated total. The source is the one cited in a), page 2, and Table 1, page 6.

c) Personal income by state for the calendar year 1957 as reported in current dollars, U. S. Office of Business Economics, U. S. *Income and Output* (Washington: G. P. O., 1958), Table I-8, pp, 156–57.

d) Value of school property June 30, 1958, in current dollars as reported in U. S. Office of Education, *Statistics of State School Systems, 1955–1956* (Washington: G. P. O., 1959), Table 39, p. 106. In footnote 1 of this table it is stated that: "while the principal basis for determining value is original cost plus cost of all additions and alterations, other bases used probably

include insurance and replacement cost." Estimates of this variable were prepared for eight states on the basis of the average property value per pupil of the forty states for which information is available.

e) The number of books circulated by state is taken from U. S. Office of Education, *Statistics of Public Libraries, 1955-1956* (Washington: G. P. O., 1959), Table 12.

f) "Region" is a dummy variable which takes on the value 1 in states in which standard metropolitan areas are mainly located. For all other states the value is 0. *1* was entered for the states of Washington, Oregon, California, Wisconsin, Illinois, Missouri, Ohio, Michigan, Florida, New York, Maryland, Pennsylvania, New Jersey, Connecticut, Rhode Island, and Massachusetts.

g) Maintenance outlays are in current dollars as reported in Table 9, page 13 of the source for cross section data cited in a).

h) Operation, instruction, administration, and capital outlays are in current dollars as reported in Table 9, page 13 of the source cited for a).

Time Series Variables as Listed in Table 1

i) Enrollment for school years ending in the even numbered year, series H224, U. S., Bureau of the Census, *Historical Statistics of the United States, Colonial Times to 1957* (Washington: G. P. O., 1960), p. 207.

j) Net national product in 1954 dollars per capita. These series were prepared from the source cited in c), Table I-2, page 119, and subsequent issues of the U. S. Department of Commerce, *Survey of Current Business.*

k) Population data for the U. S. are from the source cited in i), Series A 1, page 7.

l) Number of schools (and school plants) from the source cited in i), Series H-253, page 208.

m) Length of school term (measured in days) from the source cited in i), Series H-230, page 207.

n) Per cent population with high school degree is measured as the per cent of population 17 years old who are high school graduates. Data from the source cited in i), Series H-233, page 207.

o) Number of schools with one teacher is prepared in index form to represent a measure of rural and small town facilites which have been associated with poor quality education. Data from the source cited in i), Series H-245, page 208.

p) Operating and maintenance expenditures in current dollars by states is from the source cited in i), Series H-257, page 209.

q) Instruction expenditures in current dollars by states is from the source cited in i), Series H-256, page 209.

r) Gross capital expenditures in current dollars by states is from the source cited in i), Series H-259, page 209.