28-68

INSTITUTE FOR RESEARCH ON POVERTY

PROBLEMS IN MAKING POLICY INFERENCES
FROM THE COLEMAN REPORT

Glen Cain

Harold W. Watts

DISCUSSION PAPERS

THE UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN

PROBLEMS IN MAKING POLICY INFERENCES FROM THE COLEMAN REPORT

by

Glen Cain and Harold W. Watts

The authors would like to acknowledge the helpful comments recieved from Arthur Goldberger, Robinson Hollister, M. J. Lefcowitz, Burton Weisbrod and Walter Williams. We are most grateful to Mrs. Felicity Skidmore for editorial help.

The research of the authors has been supported by funds granted to the Institute for Research on Poverty pursuant to the provisions of the Economic Opportunity Act of 1964.

ABSTRACT

This paper is a methodological critique of the Coleman Report (James S. Coleman *et al*. *Equality of Educational Opportunity*, U.S. Office of Education, 1966). The Cain-Watts criticism is directed towards the statistical methods used, *not* to the report's substantive findings.

The principal theme of Cain and Watts is that the analytical part of the Coleman Report has such serious methodological shortcomings that it offers little guidance for policy decisions.

They show, first, that the specification of the theoretical model is inadequate--without a theoretical framework to provide order and a rationale for the large number of variables, there is no way to interpret the statistical results. They show, second, that in those instances where the Coleman Report makes clear the justification for the use of a variable in the regression model, the criterion used to assess or evaluate the statistical performance of the variable (namely, its effect on the coefficient of determination, or R^2, of the regression) is inappropriate.

Cain and Watts then go on to suggest a more meaningful approach to the problem of measuring determinants of educational achievement for policy purposes. They show (1) how the role of a variable in affecting objectives *can* be interpretable in the context of a carefully specified, theoretically justified model; and (2) that when such a model is in the form of a regression equation, an appropriately scaled regression coefficient is the most useful statistic to measure the importance of the variable for the purposes of policy action.

Problems in Making Policy Inferences from the Coleman Report.

Glen Cain and Harold W. Watts

I. Introduction

The aim of the Coleman Report [1] is twofold--(a) to describe certain aspects of our educational system, and (b) to analyze the way it is related to educational achievement--with the objective of prescribing policies to change the system. In its purely descriptive aspects, the Coleman Report presents a very dismal picture of the effectiveness of our educational system in securing equal opportunities for all our citizens. Looking at educational outcomes for children from different backgrounds one finds wide discrepancies which the American dream has assumed capable of elimination through the public school system. These discrepancies have been authoritatively established in the Coleman Report, and the indictment and challenge they present are a crucial contribution. Although we take a critical view of this Report, nothing in our subsequent commentary can detract from the importance of the findings regarding the inequalities in the education of children of different races, ethnic groups, and socio-economic classes.

Our criticism of the Report is directed toward its analysis, mainly found in Chapter 3, in which an implicit theory of the determinants of educational achievement is posited, tested, and used to point up prescriptive policy implications. The principal theme of our discussion is that the analytical part of the Coleman Report has

Second, in those instances where a theoretical justification for the use of a variable in the regression model is clear, the criterion used in the Coleman Report to assess or evaluate the statistical performance of the variable is inappropriate. Instead of providing information about the quantitative effect of a variable in altering educational achievement--information which would enable the reader to assess the feasibility and costliness of operating on the variable--the Report provides information about a statistical measure of the variable's performance (namely, its effect on the coefficient of determination, or R^2, of the regression), which gives no clear guidance for translating the statistical findings into policy action.

The remainder of the paper is organized around the development of these points. In Section II we begin by commenting briefly on the descriptive content of the Report, the social problem it reveals, and the policy objectives in response to the problem. In Section III, the core of the paper, we discuss the nature of a statistical-theoretical model necessary to handle any analysis of the determinants of educational achievement and illustrate the discussion with a hypothetical, simplified example. The purpose of this example is to indicate a relevant set of questions in terms of the objectives of social policy, and to suggest how the results from testing the statistical model should be translated into terms suitable for policy decisions. We should emphasize, however, that the example is hypothetical. The most serious gap concerning educational policy, particularly compensatory education, remains that of an inadequate theory, and we cannot fill that gap.

In Section IV of the paper we do, however, discuss a few of the many specific variables which are found in the Coleman Report to at least illustrate the points made in our hypothetical example and methodological discussion.

Our comments will be seen to take a predominantly negative tone. We admit that we are, at this time, so disenchanted with the theoretical and statistical methodology of the Coleman Report that we are pessimistic about what can be learned from it that will be useful for policy purposes.[2]

II. Policy Objectives Underlying the Coleman Report

A statement of a desirable or at least acceptable objective for social policy is provided by Coleman himself.

> From the perspective of society, it assumes that what is important is not to "equalize the schools" in some formal sense, but to insure that children from all groups come into adult society so equipped as to insure their full participation in this society.
>
> Another way of putting this is to say that the schools are successful only insofar as they reduce the dependence of a child's opportunities upon his social origins. We can think of a set of conditional probabilities: the probability of being prepared for a given occupation or for a given college at the end of high school, conditional upon the child's social origins. The effectiveness of the schools consists, in part, of making the conditional probabilities less conditional--that is, less dependent upon social origins. Thus, equality of educational opportunity implies, not merely "equal" schools, but equally effective schools, whose influences will overcome the differences in starting point of children from different social groups. [8, p. 72]

[2]Those who have witnessed the reception given the Report since its publication will recognize that our view of its statistical methodology stands in sharp contrast to the praise generally accorded this aspect of the Report. (See, for example, [3], [5], [6].) This reception has, indeed, increased the urgency we feel about pointing its weaknesses.

The task of translating the objective of equality of educational opportunity into operational terms, however, is a difficult one. The problem is, first, that the objective rests on a proposition which can be assumed but is not proven; and second, that the assessment of progress toward that objective requires measuring instruments that have yet to be perfected. These points may be spelled out as follows:

1. We can (and shall) assume that the average level of innate ability to perform in school is relatively similar across racial groups and, with only slightly less confidence, across economic classes. Lower economic classes do, of course, receive poorer health care, including pre-natal and infant care, and the relation health has to learning capacity--revealed most starkly among those mental retardates whose affliction can be traced to poor health care--suggests that innate ability or learning capacity at school ages may well be lower on average for the poorer groups. Despite this sort of relationship between class and ability, we might still accept the assumption that, say, the median levels of ability are roughly similar across race and class groups.[4]

2. When we seek to raise the educational achievement levels of individual students, we have in mind building the achievement on a given base of innate ability, and then testing to determine whether the achievement is at some desired level, given the ability factor. A serious obstacle to this approach is that our current measuring instruments are clearly not able to discriminate between ability factors and achievement factors.[5]

---

[4] The median is relatively insensitive to the location of the tails of the distribution--a fact that increases the acceptability of our working assumption. We set aside the question of how the dispersion or the distribution of innate abilities compares across groups.

[5] The problem of inadequate measuring instruments is emphasized in Mosteller, "Report of the Harvard SEEOR Group A, " May 11, 1967, in [4,pp. 7-8].

One way to cope with the measurement problem is to rely heavily on the assumption under point 1 above--the assumption of relative similarity in average abilities. On this basis, changes in factors (other than ability) which bring about educational achievement may be implemented, and the success of this effort may be tested by achievement scores that are correspondingly averaged over relatively large groups.

Such a focus on instruments of public policy to narrow the gaps between average levels of educational attainment across racial and economic groups has several implications:

1. The first priority is to develop a model in which the selection of variables is governed by a distinction between those variables amenable to policy manipulation and those that are not. The use of non-policy variables may be desirable for (a) stratifying the population if we think the policy variables have different effects on different groups, and (b) controlling for intervening effects which otherwise may bias the statistical measures of the effects of policy variables. Adding non-policy variables also serves to reduce residual variation (i.e., to increase the R^2); but with the current availability of large sample sizes this may not have a high priority, particularly since problems of interpreting the statistical results arise as more and more variables are added, some of which inevitably overlap into the role of a policy variable. The policy variables in the model of educational achievement must, of course, be relevant to the level of the decision makers for whom the

analysis is intended. An extreme example of a model that will <u>not</u> be useful in analyzing the national problem of educational inequality is one that applies to individual deviant cases and emphasizes long term clinical psychotherapy.

2. A possible conflict arises between the objective of narrowing the gap between groups and the objective of raising the overall average level of each group. Certainly there would be little support for a policy which <u>lowered</u> average levels of performance. If, however, our <u>prima facie</u> evidence leads us to the assumption that the lower economic groups and disadvantaged ethnic minority groups are performing well below their potential, then a policy which seeks to raise their performance levels may be both egalitarian <u>and</u> an efficient way to raise the overall average level of performance of all the groups combined. If we think of efficiency in terms of the cost of resources, then achieving this goal depends on (a) the identification of factors (other than ability) that affect educational performance and (b) the cost-effectiveness of operating on these factors for the disadvantaged groups compared to other groups. (We return to this issue of cost-effectiveness below.)

3. A similar conflict between (a) reducing dispersion and (b) raising the mean level, also exists <u>within</u> a group. (We should note at the outset that we must expect large

variances within groups relative to that between groups. Every ethnic and economic group, after all, includes imbeciles and geniuses, stable personality types and psychotics, hard working students and lazy students, and so on.) Now, a strategy of compensatory education aimed at a disadvantaged group might call for raising the mean level at the expense of widening the distribution. The acceptability of this outcome would have to be examined in the particular case, but it is difficult to believe that our society is likely to undertake any policies to cope with between-group differences that will widen (or indeed severely compress) existing within-group variance.

4. It may appear trivial to suggest that the variables which serve to represent educational achievement ought to be carefully chosen and justified. The Coleman Report has, however, fixed on one measure--test scores on verbal ability--to carry almost the entire burden of the published anlysis. There is no mention of parallel analyses of several measures or of the creation of a combined index. If tests of achievement differ one from another in any well-accepted sense, then theoretical considerations ought to dominate the choice of the most suitable "output" variable. Any remaining doubt could be resolved by analyzing each such variable that passes the test of theoretical pertinence. If they are all measuring the same thing

(each one imperfectly) then some, indeed almost any, linear combination of the several tests would be better than any one of them.

However, the authors seem to have postulated that one of the tests contained "it" or anyhow more of "it", and then performed the most remarkable feats of statistical augury to discover which one.[6] Perhaps other measures would have performed in the same way as the verbal ability test--we won't know until someone has tried them. But there is no indication that the choice was made on any relevant basis, and any unique properties of the measure that was used only adds to the concern about the interpretation of the findings.

III. A Suggested Approach to Measuring the Determinants of Educational Achievement

Two points about the analysis of specific variables as determinants of educational achievement are developed in this section: (1) The role of a variable in affecting objectives can only take on

[6] One justification for selecting verbal ability was that this variable possessed the largest relative inter-school variances. Another was that among the inter-student variances of test scores, school input variables accounted for more of the variance of verbal ability than of other test scores. It appears that what underlies these puzzling justifications is a preoccupation with "getting large R^2's," about which we will have a good deal to criticize in the next section. Suffice it to say here that the R^2 criterion is not relevant. What is relevant (but nowhere forthcoming in the Report) is a defense of such a verbal ability test as being a valid measure of educational achievement that is related, on the basis of a hypothesis concerning the determinants of educational achievement, to a specified set of school input variables. Instead, the fact that the verbal ability test is less likely to abe affected by the variation of school curricula and instruction than are some of the other tests is offered as further justification for settling upon the verbal ability test! (See pages 293 ff in the Report [1].)

meaning and be interpretable in the context of a carefully specified and theoretically justified model; (2) when we have such a model in the form of a regression equation, an appropriately scaled regression coefficient is our most useful statistic measuring the importance of the variable for the purposes of policy action. The second point, which is the more specific, is discussed first.

A. The Issue of the Significance and Importance of a Variable

In the analysis of the relation of school factors to achievement, the principle statistic offered in evidence by the Coleman Report is the percent of variance explained. As indicated in their methodological appendix, this is because the authors are interested in assessing the "strength" of various relationships, and they believe that the percent of variance explained provides the best general purpose indicator of "strength." It will be argued below that this measure of strength is totally inappropriate for the purposes of informed policy choices, and cannot provide relevant information for the policy maker.

Consider a general function expressing a relation between y and several x's, $y = f(x_1, x_2, \ldots x_k)$. What conceptual framework can be used to discuss the strength of the relation of y to, say, x_2? If we are limited to the information provided by the function $f(x_1, x_2, \ldots x_n)$, the partial derivative $\partial y/\partial x_2 = f'_2(x_1, x_2, \ldots x_n)$ is both simple and complete. In the case of linear functions, the partial derivative is a constant and expresses the change in y induced by a unit change in x_2.

It should be clear that a change in the unit of measurement will change the magnitude of such derivatives, and that any comparison among them must establish some basis for comparability between the units of measurement. In the context of an analysis of the relation of school factors to pupil achievement, it would seem evident that our interest lies in purposive manipulation of the x's in order to effect an improved performance in terms of y. We can, and should, ask for the expected change in y induced by spending some specific amount of money on working a change in x_2, say, as compared with the alternative of spending the same sum on x_3.

It must be emphasized that the use of budgetary cost is not necessarily the only basis of comparability. Any other basis which enables the policy-making authority to compare the effectiveness of unit changes in policy variables on the objectives of the policy will serve the same purpose. One can imagine a context in which time, or "political capital," or man-hours, might be the appropriate basis for comparing the effects of the x's. But unless some such basis is defined and its relevance to policy explained, the question of "strength" has no meaning. If we have defined a relevant basis and know the set of alternative changes in the x's, it is a relatively easy and straight-forward task to find out which one does the most for y.[7]

What basis of comparison among the x's is implied by the percent of variance explained--which is the indicator of the "strength" of a variable used in the Coleman Report? To answer this question we will consider the common case of a linear function, the only type of function investigated in the Report.

---

[7] As above, this discussion makes no reference to the first mentioned problem of deciding upon the "carefully specified and theoretically justified models." The problem of knowing what empirical variables belong in a model is, of course, a crucial one for posing the question to be addressed by the empirical work, and logically prior to the problem of estimation and inference.

The ordinary regression coefficients, b_i, for i = 1, 2, . . . k, represent the partial derivatives of y with respect to the several x's--where each x is measured in some conventional (perhaps arbitrary) unit. As indicated earlier, some adjustment of these derivatives is generally required in order to establish comparability. By using the percent of variance uniquely explained by x_i, call it ϕ_i, as the measure of strength, the authors have implicitly assumed that x's will be rendered comparable by measuring them in units corresponding to the orthogonal (or uncorrelated) part of their respective sample variances. It is easily shown that:

$$\phi_i = b_i^2 \frac{s_{x_i}^2}{s_y^2} (1 - R_{ai}^2) \quad \text{or}$$

$$\phi_i = \beta_i^2 (1 - R_{ai}^2)$$

where the s symbol refers to the sample standard deviations and R_{ai}^2 is the coefficient of multiple determination for the "auxiliary" regression of x_i on the other (k - 1)x's.[8] If there is only one x , i.e., k = 1, or if x_i is orthogonal to all other x's, the term involving R_{ai}^2 drops out and we have:

$$\phi_i = \beta_i^2 = \left[b_i \frac{s_{x_i}}{s_y} \right]^2 = \text{the squared Beta coefficient.}$$

---

[8] R_{ai}^2 is the same statistic as the C^2 referred to by Coleman in his reply [7] to the comment by Bowles and Levin [2]. Note, however that Coleman's definition of the "unique contribution" of a variable, which involves C^2, is in error unless the variable whose contribution is being assessed has a unit variance. ([7], pp. 241-242).

Thus ϕ_i represents the square of the regression coefficient which would have been obtained if

(1) each of the x's had been divided by its standard deviation discounted for its relation to other variables and

(2) y had been divided by its standard deviation.

I.e.,

$$\phi_i = \left(\frac{\partial y^*}{\partial x_i^*}\right)^2 ,$$

$$\text{where } y^* = \frac{y}{s_y} \text{ and } x_i^* = \frac{x_i}{s_{x_i}\sqrt{1 - R_{ai}^2}} .$$

It seems very difficult to find a reason why x's measured in terms of "dependency-discounted-deviations," or 3-D's, are comparable for any policy purpose. Is a 3-D increment of x_1 equally costly, equally feasible, or equally appealing to the Congress as an increment of x_2? Is there, indeed, any basis for arguing that these 3-D units form a relevant set of policy alternatives such that one would have the slightest interest in how the several variables rank according to ϕ_i?

It should be clear that measuring "strength" by the usual regression coefficients, or by the Beta coefficients, is in general no better than using ϕ. Whether the variables are scaled conventionally or by some equally arbitrary sample-generated unit, they will usually have to be re-adjusted to secure comparability in the context of a specific choice problem. (This task is usually simpler if the conventional scale hasn't been fiddled with, and it is more likely to be recognized as a necessary step in the analysis.[9]) Although the discussion above was in terms of single variables in a given function, analogous arguments hold for groups of variables or for the same variable in functions describing relations for different groups, regions, years, etc.

How did the choice of such an odd measure of "strength" come about? The most plausible explanation runs in terms of an all-too-common failure to distinguish "statistical significance" from policy or "substantive significance." In fact the F-ratio test statistic, which is commonly used to test the hypothesis that one or several coefficients in a linear function are equal to zero, is very simply related to Φ. In the one variable case, the F-ratio is strictly proportional to Φ:

$$F_{1,\ t-k-1} = \frac{\Phi_i\ (t-k-1)}{1-R^2} \quad ,$$

where t = sample size, and k = number of independent variables in the regression.

Where F is greater than some critical value, one commonly reports that the variable in question is significantly greater than zero at, say, the .05 level. All this means is that in order to maintain a belief that the variable in question has absolutely no effect, one must believe that the sample analyzed has surmounted odds of 20 to one

---

[9]Indeed, an important advantage of the ordinary regression coefficient, b_i, is that, since the units in which x_i are measured are customarily given, the effect of a unit change in x_i on y is, as a matter of course, translated by the user of the statistics into terms relevant for his decision context.

It has been suggested that publication of the regression coefficients produced by Coleman's research would lead to reckless and irresponsible interpretations ([7] p. 240). This must be because either the statistics themselves, or the users of them, are untrustworthy. If the problem lies with the statistics, it is hardly more responsible to publish statistics which are better behaved simply because they are definitionally limited to the positive numbers between 0 and 1, without revealing the more suspicious-looking joint products of the analysis. If the problem lies with the analysts, why give them any statistics at all?

by showing such a large apparent effect. Clearly, the greater is ϕ_i or F, the greater the statistical significance and the harder it becomes for a betting man to stick to the belief that the partial derivative is zero. This is surely a very restricted and specialized meaning of "significance," since it may bear no relation to the significance (i.e., importance) a variable has for policy purposes.

The F-test (or related t-test) of the "net" or "partial" coefficients is not, of course, affected by the order of introduction of the variables into a stepwise calculation of the regression. But, the effect of a variable or set of variables (however "effect" is measured) will show up as different in the case where another set of variables is "held constant," from the case where there is no control over that other set. The only exception is when the variables to be controlled are uncorrelated with the set being examined, but this situation is present so rarely in non-experimental data that it can be dismissed. An extensive controversy concerning the order of variables has appeared in the literature, [2], [7], [9], [10]. But neither critic nor defender has presented an adequate theoretical framework within which the objects of their dispute become worth arguing about.

When there is a legitimate interest in testing the zero-effect hypothesis, nothing else will quite do. There is an entirely unwarranted tendency, however, to use the F-statistic (or its cousin ϕ) to indicate the more relevant kind of policy significance. To take a homely example, one might suppose that height and sugar consumption are both related to an individual's weight (among other things of course). In most contexts height would explain more variance than sugar consumption, but who is interested in knowing by how much? No one, certainly, who seeks to embark on a weight control program. Anyone who would seriously entertain the hypothesis that weight does not depend on height has more blind objectivity than most of us--but such a person is the only one who should care about the relative size of that test statistic. It is easy to imagine an interest in a test on the "sugar effect", but why say that it is less important or significant or strong, just because it explains less variance?

When ϕ is properly interpreted as a test statistic, one must keep two things in mind. (1) Its relevance is limited to the zero-effect null hypothesis and (2) that, as in all hypothesis tests, the power of the test is as important as the level of significance. A body of data may be unable to reject the hypothesis that some coefficient is zero, and be equally consistent with a hypothesis embodying a miraculously high effect. Alternatively, a very powerful test might reject the zero-effect hypothesis, and also reject a hypothesis that the effect is large enough to warrant any further interest in a variable.

A second possible defense for the practice of evaluating variables by ϕ_i lies in its similarity to the Beta coefficient. The use of such "standardized" regression weights is usually predicated on an assumption (rarely made explicit) that the sample standard deviations used for adjusting the regression coefficients indicate a relatively fixed range of variation for the several variables. There is, in other words, some notion of "normal" limits of variation which are related somehow to the variation actually found in a population. If some x shows little variation in a representative sample drawn from an interesting population--the argument goes--then we must reduce its coefficient in order to achieve comparability with the coefficient of another x that has a larger variance.

The use of ϕ_i for comparing the effects of variables can be interpreted as the result of following this same logic farther into the labyrinth of least-squares regression algebra. Specifically (as seen by the formulas on p. 10), the standardization involved in ϕ_i is in general sensitive to the sample variances <u>and inter-correlations for all the x's in the regression</u>. Such a standardization is of interest only if one feels that the entire joint distribution of regressors is both fixed in the population and well represented by the sample.

There are many contexts, particularly in the natural processes studied in the physical sciences, when the persistence of specified sizes of the variances and correlations among some of the variables may be a warranted assumption (this should be spelled out and supported, however). But it is patently absurd to postulate such invariance for variables that <u>can</u> be affected, directly or indirectly, by the policy alternatives that have motivated the analysis.

The use of Beta coefficients (standardized only for variance) is subject to the same sort of criticism--they retain their meaning only so long there is no intervention by man or nature to change the variances used for standardization. But where β_i is only crippled as a guide to policy, ϕ_i is totally disabled. The latter maintains its relevance as a description of a relationship only if we stand aside and wring our hands.

B. A Hypothetical Numerical Example

A number of the points discussed above can be grasped most readily by a review of a simple numerical example. Suppose that the relation between a suitable measure of school outcomes (y), and indexes of school quality (x_1) and non-school background and environment (x_2), is as follows:

$$y = 1 + x_1 + 2.0x_2 + u$$

The constant term reflects an arbitrary choice of origin for the outcome measure, and we assume that x_1 and x_2 are scales with zero means and unit variances. (These scalings merely simplify the numerical calculations and interpretations of the example.) The final term, u, is an unobserved disturbance term which must, in part, reflect measurement errors in y and other relevant factors such as "native ability" (whether genetic or irreversibly determined at some earlier time). This disturbance is defined to have a zero mean and to be uncorrelated with x_1 and x_2. (Assuming that x_1 and x_2 are uncorrelated with u, either singly or in a linear combination, permits us to accept the regression coefficients as unbiased measures of the effects of x_1 and x_2.) Its variance has here been set at unity quite arbitrarily.

Now consider several alternative situations which reflect different policies with regard to the allocation of the composite bundle of factors which determine school quality, x_1. For greater simplicity we will not consider allocations that change the variation of x_1 over schools. Only the degree, and sign, of the correlation between x_1 and x_2 (ρ_{12}) will be changed. To make the policy more concrete (and more obviously hypothetical), suppose that all schools have wheels so that a fixed population of schools of various qualities can be moved around to serve an equal number of communities. A zero correlation between x_1 and x_2 ($\rho 12 = 0$) would result from a random assignment of schools to communities. It would be changed to a positive value by moving some of the better schools from "bad" communities (as measured by x_2) to "good" ones, and vice versa. Similarly, $\rho 12$ would become negative if the bad communities swapped their bad schools for good ones from the good communities. $\rho 12$ would approach 1.0 if the "best" school served the "best" community, the second best school the second community and so on.

Any alteration in the way input variables are combined will change the distribution of the outcomes; for instance, a change in the variance of y is a <u>necessary</u> result of a change in the correlation between x_1 and x_2, given our specification of constant variances of x_1 and x_2 and constant effects (b's) of x_1 and x_2. Table 1 shows the consequences for several parameters when the correlation between x_1 and x_2 takes on several different values.

In column IV one finds the simple case when x_1 and x_2 are uncorrelated--schools have been assigned to communities at random. The variance of y (σ_y^2) is equal to 6.0, and this partitions nicely into a component due to school differences with variance 1.0, another component due to community differences with variance 4.0, and a third due to the combination of factors accounted for implicitly by the disturbance term with variance 1.0. The two variables, x_1 and x_2, together account for 5/6 of the variance--1/6 for x_1 and 2/3 for x_2--as shown in the

Table I

Consequences of Varying Correlation Between Regressor Variables in a Simplified Regression Model

Model: $y = 1.0 + 1.0x_1 + 2.0x_2 + u$

$\sigma x_1^2 = \sigma x_2^2 = \sigma_u^2 = 1.0$

$\rho ux_1 = \rho ux_2 = 0.0$

Row Number	Parameters	I	II	III	IV	V	VI	VII
1	ρ_{12}	1.00	.90	.50	-0-	-.50	-.90	-1.00
2	ρ^2_{12}	1.00	.81	.25	-0-	.25	.81	1.00
3	σ^2_y	10.0	9.6	8.0	6.0	4.0	2.4	2.0
4	$\rho^2_{yx_1}$	.900	.712	.500	.167	-0-	.267[a]	.500[a]
5	$\rho^2_{yx_2}$	.900	.876	.782	.666	.563	.505	.500
6	$R^2 y.x_1x_2$	.900	.896	.875	.833	.750	.583	.500
7	ϕ_1	-0-	.020	.093	.167	.187	.078	-0-
8	ϕ_2	-0-	.184	.375	.666	.750	.316	-0-
9	$R^2 yx_1.x_2$	-0-	.160	.429	.500	.429	.160	-0-
10	$R^2 yx_2.x_1$	-0-	.432	.750	.800	.750	.432	-0-
11	β_1	.312	.327	.354	.408	.500	.645	.707
12	β_2	.624	.654	.708	.816	1.000	1.29	1.114
13	b_1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
14	b_2	2.00	2.00	2.00	2.00	2.00	2.00	2.00

[a]The squared simple correlation coefficiencies shown here are squares of negative values for ρyx_1. All other values for ρyx_1 and ρyx_2 in the table are positive.

entries for the simple squared correlations ($\rho^2_{yx_1}, \rho^2_{yx_2}$) and the squared multiple correlation, $R^2_{y.x_1x_2}$.

Because x_1 and x_2 are uncorrelated (orthogonal), the incremental fraction of explained variation that is obtained when x_1, say, is added to the regression ($\phi_1 = R^2_{y.x_1x_2} - \rho^2_{yx_2}$) is equal to the fraction explained when x_1 is used alone ($\rho^2_{yx_1}$). The same is true for the increment due to x_2.

The squared partial correlations are obtained by dividing the increment due to, say, x_1 by the fraction of variance left unexplained by x_2:

$$\rho^2_{yx_1.x_2} = \frac{\phi_1}{1 - \rho^2_{yx_2}}$$

$$= \frac{R^2_{y.x_1x_2} - \rho^2_{yx_2}}{1 - \rho^2_{yx_2}}$$

The Beta coefficients, β_i, are simply the partial regression coefficients divided by the standard deviation of y, σ_y, and multiplied by the unitary standard deviation of x_i. The partial regression coefficients shown in the last two rows are constant, of course, because the populations have been generated by maintaining that assumption. (Column I and VII, where x_1 and x_2 are perfectly correlated, are limiting cases--the multiple regressions would be impossible to carry out with data generated from these cases.)

The values of the various parameters listed in the columns of this table must be regarded as "population" values. A limited sample drawn at random from one of these populations could produce estimates of these parameters which would differ from the "true" values by sampling errors of the usual sort.

If the allocation of x_1 is changed from a random one by matching "good" schools with "good" communities, the correlation between x_1 and x_2 becomes positive. Moving toward the left from column IV in the table,one finds first that the variance of y gets larger. This is intuitively explained by thinking of the schools as re-inforcing and intensifying the inequality found in the environments. The simple correlations shown in the third and fourth rows both increase as the two variables become increasingly good substitutes for each other, and the multiple correlation goes up because the constant amount of unexplained variance (from u) becomes a smaller part of the whole variance of y.

The incremental explanatory power or "unique contribution" (measured by ϕ_i) declines as ρ_{12} increases from zero, and ϕ reaches zero in the limit where $\rho_{12} = 1$. The partial correlations display basically the same pattern. Both are transparent consequences of the increasing interchangeability of x_1 and x_2--as their correlation increases, having both adds very little new information. Finally, the Beta coefficients decline as a consequence of increases in variance of y. Any deeper meaning of this change must be supplied by those who have a penchant for using this scaling convention.

Consider now the consequences of allocating relatively more "good" schools to the "bad" locations and vice versa. As ρ_{12} falls from zero to negative values one finds the variance of y falling also. (See columns IV to VII.) Here the schools compensate for or suppress the inequality produced by unequal backgrounds.

The squared simple correlations, $\rho^2_{yx_i}$, both fall initially; $\rho^2_{yx_1}$ going to zero at $\rho_{12} = -0.5$.[10] The variance explained by x_2 falls steadily until at the limit it explains only half of the (smaller) variance of y. Beyond $\rho_{12} = -0.5$ (in columns VI and VII) the simple correlation of x_1 with y becomes negative, and in the limit it is simply a mirror-image of x_2 and thus has the same squared correlation.

The squared multiple correlation falls as the "unexplained" component of the variance becomes relatively more important. The net or unique contributions, ϕ_i, are seen to reach a peak at $\rho_{12} = -0.5$ and then to fall once more to zero as x_1 and x_2 become more identical. The partial correlations are seen to fall quite symmetrically on both sides of column IV where $\rho_{12} = 0$.

Finally, the smaller variance in y brings about an increase in the Beta coefficients. By this measure the effects of both x_1 and x_2 become more and more powerful; by contrast, the regression coefficients measuring their effects remain unchanged at their assigned values.

Now consider a not-entirely-hypothetical society which has shown some tendency to place its "best" schools in the "best" places and to direct its "best" efforts toward its "best" pupils. This produces an ρ_{12} somewhere between 0.5 and 0.9--like Cols. II or III. An educational survey might very well find that background and environment

[10]Intuitively, when $\rho_{12} = -0.5$ we can think of the positive contribution of x_1 to explaining variation in y being exactly negated because of the negative correlation between x_1 and x_2. As the negative correlation between x_1 and x_2 gets larger in absolute value than -0.5, the true positive effect of x_1 is more than offset in the simple relation between x_1 and y (when x_2 is not held constant).

is 4-10 times as strong as school quality if it looks at the relative size of the ϕ_i. Such an analysis would make less extreme, but no more relevant, statements if it compared the b's or β's. If, however, the survey is large enough to get decent estimates of the b's, its authors might note that school quality <u>does</u> make a difference, and reason that moving some schools could change β_{12}, and move the society's educational process toward one described by Columns V or VI. Such a re-allocation would substantially reduce the inequality of outcomes and attenuate the correlation of outcomes with social origins; and it would seem to be a proper sort of alternative for an educational survey to consider.

It must be heavily underscored that, in terms of the model reviewed above, comparisons of the relative explanatory strengths of the two variables x_1 and x_2, whether one uses simple, partial or multiple correlation coefficients, unique contributions or regression weights, adjusted or not, are pointless. If one is concerned with assessing the possible effects of <u>educational</u> policy, comparisons of any kind with the effect of "control" (i.e., non-policy) variables are silly. Moreover, they are all, except for the unadjusted regression coefficients, dependent upon the particular policies pursued when the data was collected. Their use runs the risk of declaring a policy feeble simply because historically it was not vigorously applied.

In the example shown in Table 1 the "best" allocation to achieve equality calls for a perfect negative correlation between x_1 and x_2. By this allocation the variance of y is reduced to a minimum (=2). It should be noted that educational policy might also change the mean and/or the variance of x_1. With these added degrees of freedom it would be possible, in principle, to eradicate all gross association of y with x_2, and--as an added option--reduce the variance of y to the absolute minimum introduced by the unobservable variable u.

C. The Need for a Theoretically Justified Model Relevant to the Policy Context

In general terms one may view the survey of Equality of Educational Opportunity as providing information on the joint distribution of a large number of variables. The analytical effort should be directed toward answering questions about how new or altered policies (more particularly educational policies) would change various characteristics of that joint distribution either directly or indirectly. To do this, one must have a consistent and complete set of specifications concerning: (1) which characteristics of the joint distribution are constant, (2) which can be changed directly by specific activities (policies), and (3) which ones must therefore be determined by the assumed structure and prescribed policy.

This set of specifications is commonly termed a theory or model. In the Coleman Report there is no explicit discussion of a consistent theory of this sort. Some theory, of course, must underlie any sort of policy prescription. It is not that one can choose to draw conclusions from the objective facts alone without the aid of any theory, but rather that if one leaves the theory implicit, ambiguous and obscure, possibly nonsensical or even self-contradicting premises go unnoticed.

The theoretical structure of the simple model discussed above asserts that the functional relation between y and x_1, x_2, and u can be approximated satisfactorily by a linear and additive function, with coefficients that would remain fixed under policies designed to change the distribution of x_1 and/or x_2. Similarly it is assumed that the mean and the variance of the disturbance variable, u, will be unaffected by policies aimed at affecting y via x_1 or x_2. The objective of policy is taken to be some optimal combination of high average level of outcomes (mean of y), minimal inequality (variance of y)--at least as the variance or inequality is affected by inter-group differences--and easy class mobility (correlation of y and x_2).

The tools of educational policy are taken to be measures that would shift the mean of x_1, compress or expand its variability, and/or revise the correlation between x_1 and x_2. If one wishes to consider social policy more broadly, similar alternatives for changing the distribution of x_2 would be available. Within the structure so far specified it is possible to deduce the effects on the marginal and conditional distribution of y for any particular change in the x_1 or x_2 distributions. If no further restrictions or relevant information is added, it is clear that any particular goal in terms of the basic objective can be achieved by a wide range of different manipulations of the x_1 and x_2 distributions. The question of relative strength, in the sense of ability to manipulate y, can now be seen to be meaningless--remembering that the scaling of x_1 and x_2 was arbitrary to begin with. Each of them can be used to achieve the objective so long as unlimited freedom is available for changing the mean, variance and correlation. If x_2 is not manipulable by educational policy, on the other hand, who cares how effective it might be if it were?

Consider, however, a very simplified situation in which the objective is to close a substantial gap between the mean value of y for Negroes and the mean for whites. Assume that the function above holds for Negroes, and that one's policy choices are limited to changing--at most--the mean value of x_1 and x_2 for Negroes. Which policy or conbination of them one chooses will depend on further information about the costs of each alternative.

Costs may be in terms of dollars, time, political consensus or all three. But each must be made explicit. If only dollar costs are considered and the cost of a unit change in the mean of x_1 costs \$5 billion and a unit change in x_2 costs \$25 billion, then it is clear that one can change the mean of y most cheaply by operating on x_1.

Indeed, one might, for purposes of policy analysis, scale the variables available for manipulation so that a unit change in x_1 is an equally costly (or time-consuming or consensus-using) alternative to a change in x_2. If an "Iso-chunk" of x_1 is defined to be a \$1 billion worth, each one must be a fifth as large as the original unit costing \$5 billion--hence its coefficient must be 0.2 (i.e., the old $b_1 = 1$ coefficient multiplied by its new unit of measure, 0.2). Similarly, an Iso-chunk of x_2 is only 4% of an original unit, and hence its coefficient must be 0.08.

Several variations on the "Iso-chunk" idea can be specified. Take as given the relation between "output," y, and "inputs," x_1 and x_2:

$$y = a + b_1x_1 + b_2x_2 + u \qquad (1)$$

Suppose first that the "costs" of alternative mixes of x_1 and x_2, in terms of any scarce item one finds important, are given by:

(2) $$C = c_1x_1 + c_2x_2 \quad .$$

One may now rewrite equation (1) in terms of "Iso-chunks" which correspond to the amount of x_i gotten by using one unit of whatever "cost" consists of--dollars, man-hours, class-hours:

$$x_1' = c_1x_1 \quad ,$$

and

$$x_2' = c_2x_2 \quad .$$

Thus, "Iso-chunks" (read dollars or hours) of C spent in changing x_i can be substituted in (1) for the x_i:

(3) $$y = a + B_1x_1' + B_2x_2' + u \quad ,$$

$$\text{where } B_i = \frac{b_i}{c_i} \quad .$$

We may call these B_i "bait coefficients"--derived from Israeli pronunciation of the Hebrew name for the corresponding alphabetic character.[11]

The bait coefficients give quite direct answers as to which use of the scarce item C yields the largest increment in y. To the extent that relations (1) and (2) adequately reflect the way the world works, one could confidently proceed to add to the existing educational process by directing all available C into the x_i for which B_i is the largest.

---

[11] This felicitous terminology is gratefully accorded to Professor Arthur S. Goldberger.

Unfortunately, one does not usually have that much confidence in a couple of simple linear relations. Commonly, relation (1) will be estimated on the basis of a limited sample, and one's confidence in extrapolations beyond the range of observed conbinations of x_1 and x_2 deteriorates rapidly. Moreover, one would rarely encounter a "cost function" as simple as the one in (2)--usually there will be diminishing returns causing marginal costs to rise beyond some point. Bait coefficients derived as above ought, therefore, to be interpreted as reflecting, at best, the relative effectiveness of variables in that vicinity of the data over which a linear approximation is deemed to be "sufficiently accurate," taking into account reservations about both relation (1) and relation (2).

IV. Interpreting Specific Variables in The Coleman Report

The absence of any explicit theory of educational achievement is the chief source of the difficulty in interpreting the statistical results of the Coleman Report. We can illustrate the problem by discussing some of the variables used in the Report.

A. Attitudinal Characteristics of the Student

One remarkable finding of the report's analysis is the high partial correlation of fate control/personal efficacy variables with the verbal ability score used as a measure of educational outcomes.[12] The relation was particularly strong (by the Report's criterion) among minority group children. Without a theory, however, we cannot answer the following types of questions:

---

[12] A number of questions in the survey attempted to measure the student's sense of control over his environment and his sense of fatalism.

(a) Is this variable itself merely a reflection of (perhaps "caused by") educational achievement? One can easily imagine situations in which educational accomplishment would instill confidence in a youngster and produce a high score on the measure of this variable.

(b) Is this variable important only because it is related to various objective factors about the student's family, community, and school environments, which are not fully measured in the model, and which "really" explain both school performance and the fatalism score? This set of relations would again be quite plausible on *a priori* grounds.

Moreover, all of the evidence offered in the Report seems consistent with the interpretation that these attitudinal variables are just another means of measuring the joint output of school and non-school processes impinging on a child's development. Their strong simple and partial correlations with verbal ability--in the absence of any model which suggests that these fate control/personal efficacy variables should be regarded as a separately manipulable cause of achievement in school--should immediately evoke the suspicion that these variables are in fact another measure of educational outcomes. The report explicitly notes that the simple correlations of verbal ability and the fate control variable are similar to the intercorrelation among the achievement variables. If the investigators had thrown one of the other achievement variables such as reading comprehension into a regression "explaining" verbal ability they would, no doubt, have observed another striking similarity with the fate control variables but in that case they might have perceived the tautology themselves.

Under situations (a) and (b) above, we can say no more than the following. <u>Either</u> changes in the variable, "control over one's fate," are unattainable unless performance on the other objective variables is changed; <u>or</u>, if some change in the score could be induced (by, say, counseling), there is no reason to believe educational performance would change.

(c) What if--contrary to (a) and (b)-- the fatalism variable is a personality trait that does have a separate influence on educational achievement? We still need to know how policy can change the trait to make use of our finding. Clearly these attitudes may be quite congruent with an objective assessment of the situation children find themselves in. If so, the school may be severely limited in its ability to reorient such attitudes (one may have to re-introduce prayer). It may be, of course, that school achievement cannot be improved without an improvement in these attitudes, and moreover, that no feasible change within the common day-school framework can affect much change in the attitudes. A verdict of helplessness may have to be passed on the schools, but the evidence in the report supports it, <u>neither</u> by adding to our knowledge of the causal relation, <u>nor</u> by indicating a low payoff from interventions within that relation.

B. Characteristics of the Student's Peer Group

In a review of the Report's findings, Harry C. Bredemeier notes: "More important than all school characteristics and teacher quality for Negro students is the degree to which the <u>other</u> students in their schools have the following characteristics: Their families own encyclopedias, they do not transfer much, their attendance is regular, they plan to go to college, and they spend rather much time on homework," ([11], p. 21) He notes in a footnote,"I assume no one will infer from this that the 'solution' is to put encyclopedias in everyone's home."

But, is such an inference less satisfactory than making no inference? Is it any more naive than the presentation of the vague theoretical framework that permits us almost no grounds for saying how we should interpret the "significant positive coefficient" of the encyclopedia variable? We can illustrate the difficulties of the "no theory" position with the following attempt to supply an interpretation:

> "Encyclopedia ownership is a variable that indicates an intellectual atmosphere in the home conducive to schooling, and/or a measure of affluence that is not fully captured in other measures (of affluence) in the model, and/or a measure of parental attention or affection that contributes to the students' emotional stability and, thereby, to school performance--any or all of which factors creates the positive peer group influence."

Presumably, this interpretation is "more sophisticated" than the inference Bredemeier noted. But is it more helpful? Indeed, what our hypothetical theory has told us up to now is that: (1) if it is intellectual atmosphere that underlies the relation, the variable has probably no policy significance since we do not know much about changing intellectual atmosphere. If we thought we did know something about how to make the change, we would need to know the specification of the relation between encyclopedias and intellectual atmosphere. (2) If it is affluence that underlies the relation, then we need to ask our theory to translate a unit of encyclopedias to a unit of wealth (or income flow) so that we know how much of a change in income will be necessary to yield the changes in educational performance.

We could continue these "if" questions almost indefinitely, but let us summarize the function of our hypothetical theory by saying that it has forced us to consider the possible tortured interpretations we have to make or preposterous policy actions we might have to follow as a consequence of such cavalier inclusion of *ad hoc* variables in our model.

C. Environmental Characteristics

The Coleman Report stressed that the influence of the regional and urban location of the school and the socio-economic status of the student body in the school were highly important in explaining school performance. A theoretical proposition underlying the authors' interpretation of this finding was that the environment is exogenous and "causally prior" to such factors as school resources, so that an appropriate procedure was to enter the former variables, note the contribution to R^2, and then add the school resource variables and observe their additional contribution to R^2. Other demurrers to the procedure, quite apart from the issue of the R^2 criterion, may be mentioned.

If families select their residence on the basis of the quality of school, residence is neither exogenous to the process nor causally prior to the school resources variable. Particularly with regard to the racial composition of the school, the phenomenon of selective migration may be confounding the results. For example, if a large percentage of whites in a school or a large percentage of high socio-economic groups appear to have a positive effect on the educational performances of Negroes or low SES groups, we should consider the hypothesis that the latter families have strong "tastes" for a high quality education for their children and have moved to a district where the school has a favorable reputation. The observed positive effect of the environment on the educational achievement of disadvantaged groups may therefore be overstated, since some of the effect stems from the unmeasured personal traits of the families, and it is further possible that some effect is attributable to the beneficial resources of the school.

More generally, any variation in school performance that is attributed to the urban-rural or regional location of the school cannot be "set aside" so that the remaining variation can be examined in detail. What theory of educational achievement justifies "urbanness," "Southernness," etc. as causal factors, except insofar as these traits are related to such specific variables as the family characteristics and quality of schools found in these areas? There is a real danger that such location variables serve only to attenuate the influence of other variables of interest when such other variables are unmeasured or measured with a large error component.

D. Teacher Quality

One type of variable that belongs in the category of school resources over which we have some degree of policy control is "teacher quality"--itself a composite concept made up of several variables. The conclusion in the Report about teacher quality appears to strike a rare optimistic note regarding the beneficial influence school resources can have in compensatory educational efforts. The Report states on page 317 that "a given investment in upgrading teacher quality will have the most effect on achievement in underprivileged areas." Surely, the theoretical justification for this variable should be quite firm. Moreover, the wording of the Report's conclusion exactly fits the criterion we have requested for assessing each variable.

Unfortunately, the statistical evidence in support of the finding the authors present concerns "variance explained": "Given the fact that no school factors (excluding student body composition) account for much variation in achievement, teachers' characteristics account for more than any other." And, "by the 12th grade, teacher variables account for more than nine percent of the variance among Negro students, two percent among white students" (page 325). It is perhaps superfluous to mention again that this ranking of importance of a variable in terms of variance explained does not tell us what the "bait coefficients" are, nor permit us to derive them; therefore, the conclusion about a "given investment in upgrading teacher quality" for underprivileged areas is not supported. If, for example, the variance of verbal ability was large among teachers of Negro students and the educational achievement scores had a relatively small variance, the high partial correlation coefficient of this variable would be consistent with a small value for the bait coefficient--even setting aside cost considerations. (See the formulas on pp. 12 and 26 of this paper.)

The full complement of variables representing teacher characteristics is, itself, not very reassuring. In section 2 of the Coleman Report we learn that Negro children are being taught by teachers who (to a significantly greater extent than teachers of white children):

(1) have low verbal scores

(2) have been born and educated in a county where they are now teaching

(3) rate their students as low either on motivation or achievement

(4) lack desire to teach high ability students.

We submit that only item (1) in this list has any clear connection to teaching quality as it links up to educational achievement, but, again, this issue can only be resolved by an explicit theory which would justify the proper linkages and which would precede empirical testing.

E . School Resources

Perhaps the single category of variables most susceptible to policy manipulation is that of school resources. Unfortunately, the variables used to measure shcool resources are very much like the "encyclopedias in the home " we discussed above. It is difficult to know whether, for example, library books or laboratories are supposed to represent their own effects, per se, or whether they are supposed to represent a more extensive collection of items under the rubric of school facilities (or some other concept of school characteristics).

One can argue for either interpretation. On the reasonable assumption that libraries and laboratories are and would be closely linked to an underlying specification of the usage of these facilities, we could treat libraries and labs as proxies for the "usage" concepts, which in turn can be plausibly linked to educational performance. Given this, the reader might further surmise that the two variables must be standing solely for their own effects, for otherwise the authors would have included the other items.

If, on the other hand, it is naive to assume that facilities present are facilities used, and if it would have been overly burdensome to include all relevant items in the survey, then we can more readily accept the argument that the included variables are meant to be representative of some different and/or larger collection. If so, we need to ask: (a) what are these other variables; and (b) what is the specification (i.e., regression equation) by which they are linked to the other variables. This really breaks up into two other questions: how accurate is the representation (i.e., how strongly are they correlated), and what is the quantitative magnitude of the relation (i.e., what are the regression coefficients linking the full set of variables to the proxy variable)?[13]

The sort of questions we have been posing serves to illustrate the two weaknesses we have noted previously. First and foremost is the absence of an explicit underlying theory with which to interpret the "facts" reported in the statistical work of the Coleman Report. If the questions we have raised are overly demanding of the state of theoretical knowledge about the educational processes, we can only say that this shaky base should be made explicit. Perhaps researchers will be led to work with a more simplified model that can be well specified and interpreted--better this than a complex model that defies interpretation.

[13]The complexity of this specification need not be exaggerated. There are many decision contexts in which proxy variables may represent a bundle of heterogeneous components, and it may not be worthwhile or expedient for the decision maker to distinguish among the components to determine their separate measures of effectiveness. What is necessary, however, is some translation of a unit of the proxy variable into a unit of the larger bundle (along with, eventually, some measure of the costs of the larger bundle).

the statistical and methodological techniques in the Report should be viewed as reinforcing the challenge to the "educational establishment"[14] to provide evidence on the effectiveness of their programs, especially compensatory education programs. Nor should any research into the determinants of educational achievement overlook the potential contribution that may stem, however indirectly, from the simple improvement in economic status of the student or his family or the families of his fellow students.

---

[14] The term was used by Daniel P. Moynihan [12] in the context of his criticism that "educationists"--administrators, teachers, research personnel--have shirked their responsibilities to evaluate their performance and have attempted to use "technical" criticism of the Coleman Report as an excuse for continued inaction.

REFERENCES

[1] James S. Coleman et al. Equality of Educational Opportunity. Washington: U.S. Office of Education, 1966.

[2] Samuel S. Bowles and Henry M. Levin. "The Determinants of Scholastic Achievement--An Appraisal of Some Recent Evidence," Journal of Human Resources, 3 (Winter 1968).

[3] William H. Sewell, "Review," American Sociological Review, vol. 32, no. 3, June, 1967, pp. 475-479.

[4] Report of the Harvard Faculty Seminar on the Equal Educational Opportunity Report, Group A, drafted by Frederick Mosteller, May, 1967.

[5] Reviews by Leonard A. Marscuilo and Harold W. Pfautz, in the "Review Symposium," American Sociological Review, vol. 32, no.3. June, 1967, pp. 479-483.

[6] Robert C. Nichols, "Schools and the Disadvantaged," Science, vol. 154, December 9, 1966, pp. 1312-14.

[7] James S. Coleman, "Equality of Educational Opportunity: Reply to Bowles and Levin," Journal of Human Resources, 3 (Spring 1968).

[8] James S. Coleman, "Equal Schools or Equal Students?," The Public Interest, vol. 1, Summer, 1966, pp. 70-75.

[9] Marshall S. Smith, "Comments on Bowles and Levin," Journal of Human Resources, vol. III., no. 3, Summer, 1968, pp. 384-389.

[10] Samuel S. Bowles and Henry M. Levin, "More on Multicollinearity and the Effectiveness of Schools," Journal of Human Resources, vol. III, no. 3, Summer, 1968, pp. 393-400.

[11] Harry C. Bredemeier, "Schools and Student Growth," The Urban Review, April, 1968, pp. 21-27.

[12] Daniel P. Moynihan, "Sources of Resistance to the Coleman Report," Harvard Education Review, Vol. 38, no. 1, Winter, 1968, pp. 23-36.