STATISTICAL MODELS FOR ASSESSING COLLEGE IMPACTS

John A. Creager, American Council on Education

Studying the impact of college attendance on students should supply information for making useful and noncatastrophic decisions. We need to know the structure of the system of higher education, the flow of students through various kinds of colleges, the effects of different matches between students and colleges upon a wide variety of outcomes of the educational process, and the effects of changes in system parameters on such outcomes.

Complex research strategies are required to study the highly complex and heterogeneous system of higher education. Moreover, such research cannot possibly be the bailiwick of any special discipline; multidisciplinary attack is required both substantively and methodologically. It is therefore important that we have means, such as this symposium, for communication and mutual criticism of what we are doing and thinking.

Recently, Feldman (1970) reviewed the present state of the art of studying college impact. He has also suggested how to choose methods and models by noting differences in the investigative purpose and correlative features of the models that might make one more or less appropriate than another. The new directions in studying college impact will probably build upon selection, elaboration, and combination of the models noted by Feldman rather than upon the development of some radically new approach to the whole business.

In deciding which proposed model to use for studying college impact, one must consider the problem of multicollinearity. Our measurements of input, of the demography of students and colleges, environments or treatments, and outcomes simply are not orthogonal to each other. Moreover, multicollinearity is more severe than is indicated by observed correlations, which are attenuated by random error of measurement.

Some, but not all, of the multicollinearity of a system results from our inability to assign students at random to colleges, thus forcing us in practice to use the natural and quasi-experimental designs, especially the longitudinal study of well-defined cohort groups. Some investigators have tended to regard this matter of nonrandom input as a nuisance: i.e., something to be corrected for. This view is not necessarily wrong, but it needs to be supplemented by analyzing the multicollinearity itself. Otherwise, we will continue to be confounded in our judgments of the relative importance and interplay of variables purporting to measure some aspects of the system of higher education. Such attention to the nature of nonrandom input and to other sources of multicollinearity in the system should produce information useful in its own right and possibly alert us to matters requiring caution in our conclusions and recommendations.

Since the literature of psychology, sociology, and education includes extensive discussion of multicollinearity and of nonrandom input problems, we can dismiss from further discussion such models as analysis of variance and analysis of covariance, in either their univariate or multivariate forms, and most partial correlation techniques. As Feldman noted, these techniques are usually inappropriate and misleading for studying college impact. Either they simply fail to cope adequately with nonrandom input, often removing college effects along with the confounding input, or they presume some inappropriate distribution of multicollinear sources of variation among the input, treatment, and output variances. We are left then with three methodologies: (1) regression analysis, with subsequent partitioning of explained variance; (2) path analysis, with its representation and testing of causal theory; and (3) stochastic approaches to understanding the flow of students through the system.

Those of us involved with regression methods in studying college impact are not solely interested in prediction, useful as that may be. Regression analysis can give us much information about the status quo of the system and about those relationships in the system that can be used to formulate empirically based theory. This view assumes that certain realities of the educational system are reflected in regression analyses. Moreover, generalized regression procedures are applicable to ordinal and nominal, as well as interval data and to the examination of nonlinear relations. Thus, we have the powerful practical tool we need to explore the domain of higher education; i.e., to break through the folklore with hard data. The problem is one of interpreting the results in terms of sets of variables of interest. Bottenberg and Ward (1963) proposed omitting a given set of variables and observing the effect on the predictive efficiency of the system. What such a procedure indicates is just how much the variable-set in question adds to what is already present in the system to account for variance in the dependent variable. This is often useful to know, especially if certain variables have already been identified as optimal and irreplaceable for inclusion in analysis. Repeated application of this approach to various subsets of variables in the full regression model leads to the uniqueness-commonality model, which is essentially a multiple-part correlation method. I question this procedure for studying college impact because it can stumble on the multicollinearity problem in subtle and potentially misleading ways.

The closely related two-step regression model is also a multiple part-correlation procedure, one which reverses the Bottenberg-Ward procedure by first computing a reduced regression model and then building up a fuller model. Most investigators would agree with Feldman's summary of the criticisms of this procedure, especially with regard to its assignment of input-treatment relations to input. However, in one study, Astin (1968) reversed the order of the stepwise entry of input and treatment sets of variables to see whether his substantive conclusions were vitiated by the alleged error: that is, he assigned the input-treatment dependencies to the treatments. This does, indeed, give some protection against error in concluding that input accounted for more of the outcome variance in achievement than did the measured college environmental characteristics. In some cases, however, when input and treatment are more closely balanced or the input-treatment dependencies are sufficiently large, the reversal procedure would result in ambiguity rather than confirmation of the findings that emerge in doing the initial regression against input.

Some have defended the two-step procedure on grounds of temporal asymmetry; their argument is plausible and pragmatic, provided that one is reasonably sure that the input-treatment dependencies are caused by input variables or their antecedents and not by some feedback of information about college characteristics. This matter can be studied operationally by including among the input variables measures of why the student chose that college, and by including among the college characteristics variables measures of college admissions policies. The way in which these variables are distributed among the factors defined by the total system should throw more light on the question.

One also needs to be assured that the input-treatment dependencies are not a function of socioeconomic or of cultural factors which influence both input and treatment variables. A reanalysis by the orthogonal decomposition model (Creager and Boruch, 1969) of some of Werts' (1968) data about the effects of home and school variables on achievement yielded just such a result. It makes no sense to assign such sources of variance or hypothesized causes to either input or treatment sets of variables. This kind of experience has led me to prefer methods which isolate multicollinear dependencies and force examination of them, especially those associated with the confounding of input and treatment through nonrandom admission of students to colleges. Meanwhile, the two-step method continues to be used extensively (Astin, in press, b) with considerable insight and judgment.

Recently, Astin (in press, a) has been concerned with another problem that is critical regardless of the models used: the effect of errors of measurement, not only on our analyses, but also on our interpretation of results. Beyond the usual lip service, inadequate attention has been paid to this matter. In fact, the college effects literature and much of the sociological literature lack empirical estimates of reliability even for commonly used variables. We have, therefore, been obtaining some empirical estimates for many of the variables used in the Council's Cooperative Institutional Research Program.

The various procedures based on multiple part-correlation methods attempt to account for college impact through partitioning explained variance in the dependent variable; this accounting is done by comparing proportions of outcome variance explained by full and reduced regression models. The alternative approach involves the examination of variance in the full regression composite. This examination is sometimes done by looking at the algebraic terms of the standard formula for linear composite variance; it is annoying to see that this procedure continues to be used in spite of widespread criticism about its generally fallacious nature. Because it too stumbles over the multicollinearity problem, it is useful only for making rough qualitative judgments of the rank of importance in systems with low correlations and large differences in the weights.

There remains the procedure of complete orthogonal decomposition of the regression system in terms of <u>interpretable</u> common and unique factors (Creager and Boruch, 1969). Here again, the aim is to account for predictable variance in the dependent variables or outcomes, because such variance is directly and pragmatically a measure of differences in outcomes that one may ultimately be interested in modifying. This procedure is applicable to canonical and discriminant composites and to bipolar systems. Moreover, the results are invariant under reflection of either variables or factors. A simple computer program is now available (Creager, 1971) for inputting the weights on up to six linear composites, and inputting the results of the factor analysis of the system under study, to obtain the orthogonal decomposition of composite variance.

Two comments should be made about this method. The first has to do with attempting to analyze very small systems where clear factor definition may not be attainable. Such a situation may also occur in somewhat larger systems in which only those variables selected by stepwise regression are included in the analysis and thus adequate factor definition and interpretability may become a problem. Fortunately it is not necessary to confine the definition of factors to just those variables in some ad hoc system of immediate interest. One may include variables which were permitted to enter freely into regression but did not in fact do so, with no adverse effect on the analysis of one or more composites derived from the total system. Moreover, this kind of operation has advantages in a large-scale research program where a single factor analysis can serve as the basis for analyzing many derived composites. The second comment is that we do not yet know the effects of measurement errors on the outcome of such analysis. The problem can be virtually eliminated, however, by correcting correlations for attenuation prior to doing either the regression or the factor analyses.

With the orthogonal decomposition model, one can say: If I do something to change values on this variable, it will have certain effects on the values of the dependent variable and on any other variable in the system. This procedure does not preclude the use of previously specified hypotheses which may or may not be confirmed by the analysis. Thus we are interested in more than prediction and description of the status quo. Since the method is new, most of the effort so far has been one of understanding the limitations of the method, of perfecting it, and of providing computer software to handle it (Creager, 1971). Recently, the method has been put to its most severe test in a reanalysis of the data for Astin's study of undergraduate achievement and institutional excellence (1968). Three full model regression composites developed by free entry were analyzed by orthogonal decomposition and the results compared with those of Astin, who used two-step regression and the uniqueness-commonality account of variance. The conclusions were very similar, in that both strategies indicated that achievement is more strongly determined by input than by environment, and the agreement was partly a function of the fact that the multicollinearity within input and treatment sets was stronger than that between sets. However, the reanalysis also indicated that free-entry regression defines a more efficient and parsimonious predictor system and is less biased toward input. Orthogonal decomposition also provides more detailed information about variances, which may be pooled in various ways for heuristic interpretation. Further demonstration of the scientific value of this strategy remains on the agenda for the study of college impact.

About path analysis, I claim no expertise. Since we have others on the panel speaking on this topic, I shall limit myself to some comments and some questions regarding this procedure for studying college impact.

I do not think path analysis and the orthogonal analysis of prediction systems are interchangeable, although at several points, they are similar and possibly compatible. It is more likely that the two methodologies are mutually complementary in the sense that both can contribute substantive information for a synthesis of the larger picture of higher education. The two methods differ in how they define direct and indirect effects, how they express these effects in empirically derived numbers, and how they add up to some kind of total effect. In the orthogonal decomposition method, orthogonal portions of variance represent sources of variance that can be explicitly subtotaled across inputs, treatments, and other sets of variables in the system, and totaled to the predictable variance of the dependent variable. In path analysis, path coefficients -- which are usually regression weights -- and their products add up, under certain conditions, to the zero-order correlations with the dependent variable. Thus, it partitions and accounts for the slopes of the individual regressions in terms of the slopes of partial regressions. Much of my objection to interpreting regression weights as anything but a particular set of weights that maximizes prediction has resulted from attempts to interpret them, their squares, and their products as independent portions of variance, which they are not. Regression weights treated as slopes of particular partial regressions, and correlations treated as slopes of total regressions, indicate change in the dependent variable produced by change in the independent variables. This alternative to describing these changes in terms of variance is a quite legitimate and useful one. However, regression weights -- and therefore path coefficients -- are not orthogonal to each other so that the meaning of what adds up by the internal algebraic consistency of path models is not entirely clear. But the path analyst cannot be accused of completely ignoring multicollinearity. External to the algebra is the causal model embodied in the path diagrams which imply the observed relationships. In some path models, however, correlations among exogenous variables are left unexplained, although they are used in computing the coefficients for other paths. Doesn't this imply only a partial attention to multicollinearity, and does it not invite surprising side effects when results are applied in practice?

Path analysis is a class of methodologies consisting of models implied by external hypotheses about causal relations. Any realistic study of college effects would seem to require recursive systems from multistage, multivariate models. When one has many input, cultural, and treatment variables, the resulting path analysis would be extremely complex, to say the least. In orthogonal decomposition, this complexity is advantageous in the sense that one can obtain a finer factor resolution of the system. If the path analyst solves the problem of complexity by doing many analyses of small pieces of the system, can the results of the many analyses be synthesized in a logically and statistically consistent way? Perhaps so, although I have only seen very simple systems studied by path analysis in college effects studies. Another new direction in college impact studies with path analysis might well demonstrate the consistency of causal relations confirmed in two or more fractionated studies with those confirmed by a single and more complex analysis.

Concerning stochastic models, there is little to add to Feldman's remark that they are promising but have to be employed in actual analysis of data before they can be properly evaluated. I simply note here that classical Markov models are not generally realistic for college impact studies. The Cornell models developed by McGinnis and his associates (1968) try to achieve more realism by changing axioms: e.g., to permit more than one-step dependency. This approach is clearly a relevant line of future development. A different line of development assumes that rectangular, rather than square transition matrices will be required for realistic applications to college impact studies. The states meaningfully defined at college entry and at later stages are not necessarily the same.

Both the stochastic and path methods have the problem of keeping the number of transitions from getting out of hand, and if this problem is handled by fractionation, then synthesis of results is required. Prior regression analyses and application of the hierarchical grouping algorithm may help to solve this problem.

So far, stochastic approaches to college impact studies are little more than extensive cross-tabulations with chi-square tests to detect nonrandom input, which we already know exists, and to detect the existence of college effects. There seems to be a kind of built-in separation of input and treatment, and obviously the description of flow of students is useful for purposes of manpower and resource allocation. Beyond this, it is not yet clear whether this approach can be developed to produce any real understanding of the system of higher education.

The extension of college impact methodology to the study of multiple outcomes and their interrelations is available by canonical regression with orthogonal decomposition of the canonical variates and as multiple outcome states in the stochastic methods. I see no reason why path analysis could not be similarly extended to study multiple dependent variables. However, the additional complexity involved may present problems.

At the present state of the art, we need to press forward with all three methodologies, not only to detect, but also to estimate, the extent of college impacts. We must also press forward to polish our methods and ensure their integrity, which involves not only their internal logical and statistical consistency, but the soundness of any interpretations of substantive results. We hope, and expect, results to be taken seriously in practical decision making. We had, therefore, better be right. On the other hand, if we wait too long to be sure, the world will have passed us by and our results will no longer be relevant.

REFERENCES

[1] Astin, A. W., "Undergraduate achievement and institutional 'excellence,'" Science, 161 (1968), 661-668.

[2] Astin, A. W., "The methodology of research on college impact," Part I, Sociology of Education, 43 (1970), 223-254; Part II, in press. (a)

[3] Astin, A. W. Predicting Academic Achievement in College. New York: The Free Press, in press. (b)

[4] Bottenberg, R. A., and Ward, J. H., Jr. Applied Multiple Linear Regression. Technical Documentary Report PRL-TDR-63-6. Lackland Air Force Base, Texas: 6570th Personnel Research Laboratory, Aerospace Medical Division, (1963).

[5] Creager, J. A., "A fortran program for the analysis of linear composite variables," in press.

[6] Creager, J. A., and Boruch, R. F., "Orthogonal analysis of linear composite variance," Proceedings, 77th Annual Convention, American Psychological Association, (1969), 113-114.

[7] Cronbach, L. J., and Furby, L., "How we should measure 'change'--or should we," Psychological Bulletin, 74 (1970), 68-80.

[8] Feldman, K. A., "Research strategies in studying college impact," ACT Research Report, 34 (1970).

[9] McGinnis, R. A., "A stochastic model of social mobility," American Sociological Review, 33 (1968), 712-722.

[10] Werts, C. E., "The partitioning of variance in school effects studies," American Educational Research Journal, 5 (1968), 311-318.