DOCUMENT RESUME

ED 222 518 TM 820 664

AUTHOR Ellett, Frederick S., Jr.; Ericson, David P.
TITLE Towards an Algebra for Analyzing Causal Relations.
PUB DATE 8 Mar 82
NOTE 71p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Attribution Theory; *Correlation; Evaluation Methods; Models; *Path Analysis; *Probability; Regression (Statistics); *Statistical Analysis
IDENTIFIERS Causal Inferences; *Causal Models

ABSTRACT

Correlation-based approaches to causal analysis contain too much irrelevant information that masks and modulates the true nature of causal processes in the world. Both causal modeling and path analysis/structural equations give the wrong answers for certain conceptions of causation, given certain assumptions about the "error" variables. An alternative approach, the conditional probability approach (CP), uses conditional probability and not correlation as the key concept. The CP approach can avoid the shortcomings and problems of such methods as causal modeling and path analysis. It provides plausible composition and decomposition rules as well as a plausible measure of causal strength. Presented in the supplement are CP theorems which cover both dichotomous and continuous cases under two sets of assumptions about the "outside causes" of a system which involves probabilistic causation. (Author/PN)

Towards an Algebra for Analyzing Causal Relations

Frederick S. Ellett, Jr.
Graduate School of Education
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024

David P. Ericson
Graduate School of Education
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024

March 8, 1982

X9FSE/B

(E-Copy)

Paper to be presented at the Annual Meeting of the American Educational Research Association, March 1982.

Towards an Algebra for Analyzing Causal Relations

Frederick S. Ellett, Jr.

David P. Ericson

U.C.L.A.

ABSTRACT

A major part of the paper investigates the relationships among correlation, partial correlation, and various conceptions of causation. Two widely used rules of causal inference are explicated. It is argued that for each of the conceptions of causation the causal inference rules, which use partial correlations, are invalid.

Another part of the paper explicates the basic principles of path analysis and structural equation analysis. It is shown that these approaches are subject to three different but important problems.

The final, major part tentatively develops an alternative approach (the conditional probability approach) which uses conditional probability and not correlation as the key concept. It is shown that the C.P. approach can avoid the shortcomings and problems of such approaches as causal modeling and path analysis. It is also shown that it provides plausible composition and decomposition rules as well as a plausible measure of causal strength.

In a supplemental section, we present the theorems which hold for dichotomous systems under two sets of assumptions and the theorems which hold for continuous systems under two sets of assumptions, theorems where the probabilistic conception of causation is employed.

Towards an Algebra for Analyzing Causal Relations

OUTLINE

Towards an Algebra for Analyzing Causal Relations

I. Introduction

Educators often make assertions about causal relations among variables. Indeed, there are compelling arguments for holding that scientific analysis and policy formation must be concerned with causation.

This paper tentatively develops an algebra for analyzing causation. It develops a measure of causal strength and composition and decomposition rules for dichotomous systems which have important advantages over and are not subject to the various problems of such approaches as causal modeling and path analysis.

A major part of the paper investigates the relationships among correlation, partial correlation, and various conceptions of causation. Two widely used rules of causal inference are explicated. It is argued that for each of the conceptions of causation the causal inference rules, which use partial correlations, are invalid.

Another part of the paper explicates the basic principles of path analysis and structural equation analysis. It is shown that these approaches are subject to three different but important problems.

The final, major part tentatively develops an alternative approach (the conditional probability approach) which uses conditional probability and not correlation as the key concept. It is shown that the C.P. approach can avoid the shortcomings and problems of such approaches as causal modeling and path analysis. It is also shown that it provides plausible composition and decomposition rules as well as a plausible measure of causal strength.

In a supplemental section, we present the theorems which hold for dichotomus systems under two sets of assumptions and the theorems which hold for continuous systems under two sets of assumptions, theorems where the probabilistic conception of causation is employed.

II. Disagreements about the Nature of Modeling

Causal modeling is quickly growing in popularity among researchers in the social and behavioral sciences. Already a well-accepted heuristic and analytic instrument in economics, sociology, and biology, causal modeling is now being strongly promoted in psychology and education as well (see Bentler, 1980, for an overview of past and recent developments). Causal modeling techniques travel under a variety of names--multiple regression analysis, linear structural equation analysis, simultaneous equation methodology, path analysis, dependence analysis, covariance structural analysis, and simply, structural analysis. For many the intent is to model the causal relationships between variables so as to obtain the best "fit" with the data. Included in this are the tasks of identifying relevant variables and determining the direction of causality between those variables. Furthermore, especially in the case of path analysis, the attempt is made to estimate the numerical value of the coefficients once a model has been specified.

The techniques of causal modeling have now reached a high degree of statistical sophistication. But in this paper. we shall provide a treatment of the logic of causal inference. Thus, rather than dwelling upon purely technical problems, estimation or measurement issues, our focus is upon the substantive or theoretical interpretation of causal models.

One might respond that causal modeling has nothing to do with causation. Indeed, in an influential paper on fallacies in statistical inference, Guttman (1977), going beyond the well-known "correlation does not imply causation" (pp. 97-98), also states, "causal analysis does not analyze causes" (pp. 103-104). And in a review of Kenny (1979), Steiger (1980) concludes:

> "Causal analysis," then, seems to be, in an important sense, a misnomer. "Linear structural relations analysis" is a more modest, but more appropriate description...Since necessary and sufficient conditions for testing "causality" from correlations have not yet been proposed, we may run the risk of deluding the uninitiated (and perhaps ourselves as well) when we call such model-fitting "causal analysis." (p. 404)

Furthermore, in an important review essay, Bentler (1980), following Guttman, claims:

> Obviously, it is not necessary to take a stand on the meaning of "cause"...The word "cause" is meant to provide no philosophical meaning beyond a shorthand designation for a hypothesized unobserved process, so that phrases such as "process" or "system" modeling would be viable substitute labels for "causal" modeling. In such a definitional context, one need not worry about the criticism that "causal analysis does not analyze causes." (p. 420)

But it is important to note that Guttman (1977), in his paper at least, provides few grounds for his views about deriving causal inferences from correlational data or about the non-causal nature of causal modeling.

Not all proponents of causal modeling agree with Guttman. Rather, many are explicit about rendering a causal interpretation of causal modeling. Sewell Wright (1921), the developer of path analysis, was merely one of the first to argue for the causal interpretation. Simon, in his paper on spurious correlation concluded, "hence correlation is proof of causation if we are willing to make the assumptions of time precedence and non-correlation of the error terms" (1957, pp. 42-43). Blalock, following Simon, concurs, "...a method for inferring causal relationships from correlational data...involves sets of prediction from causal models where certain combinations of correlations can be expected to disappear" (1972, p. 51). Duncan (in Blalock 1971; originally published in 1966) also supports the causal interpretation. Goldberger is most explicit, "In a structural equation model each equation represents a causal link rather than a mere empirical association" (1973, p. 2). Asher (1976) and Kenny (1979) are but two further proponents of a causal interpretation of causal modeling.

There appears to be, then, a sizeable disagreement about the causal nature (or lack there of) of causal modeling.

Yet regardless of this disagreement, many researchers do agree on various rules of inference that provide the foundation for model specification, determiniation of direction among variables, and the correlation coefficient. We formulate two of these rules in Section V. For many researchers advancing a causal interpretation of causal modeling, the two rules are recognized as rules of causal inference. In Section VI we formulate another, broader set of rules of inference. Thus in the main body of this paper we shall explore the validity of these rules for drawing causal inferences. Our remarks on this issue of validity do not

necessarily generalize to the work of those who disavow the causal interpretation of causal modeling.

III. The Scientific and Practical Importance of Causation

It might be thought that it makes little difference whether we render a causal interpretation to our causal models. Perhaps very little is at stake. In that case, caution would favor a non-causal interpretation; there would be no need, then, to worry about the validity of causal inference rules. To the contrary, we think a good deal is at stake. First the logical status of the social sciences is at stake. Second, also at stake is whether social science findings, using causal modeling methods, have any clear implications for social policy.

In the first case, many believe that central to the enterprise of science is covering law explanation. And the scientifica search for such laws is the search for causal laws. Whatever else their differences, a majority of thinkers in the social sciences and philosophy of science agree upon this point (among social scientists see, for example, Campbell and Stanley (1966), Chomsky (1968), Cronbach and Meehl (1955), Cook and Campbell (1979), Hicks (1979), Simon (1957), Skinner (1953), and Suppes (1970); among philosophers of science see, for example, Giere (1979), Hempel (1967), Lakatos (1970), Nagel (1961), Popper (1972), Rudner (1967), Salmon (with contributions by Jeffrey and Greeno (1971) and Smart (1963)). This notion of causal explanation is quite different from that in merely predictive systems in which researchers are interested in the proportion of variance of variable Y "explained" by variable X. It is one thing to know that X predicts Y; yet it is quite another thing to know that X causally explains Y. If X merely predicts

Y, then we have no way of knowing whether X and Y are causally related or whether their co-variation is due to other variables (hence the concern of Simon and others with spurious correlation).

This is perhaps why Simon, Blalock, Goldberger, Asher, Duncan, et al. are prepared to give a causal interpretation of causal modeling. The causal interpretation results not from a lack of caution, but from an understanding of the central role of causal expalantion in science. Without the attempt to derive causal inferences from our causal models, we lose sight of the fact that a major aim of the social sciences, like the natural sciences, is to explain phenomena--that goes beyond description and prediction. And here we must state that we agree with Simon, Blalock, Goldberger, et al. on the importance of causal explanation in the social sciences. No matter how difficult the task, we do need a methodology for causal analysis in those areas of social research in which the strict experiment cannot be performed. For this reason, it is important to assess how well the proposed rules of causal inference enable us to legitimately draw such inferences from correlational data.

On the other hand, there is nothing wrong with predictive systems of causal modeling. They, too, are a wholly legitimate form of social science inquiry. However, if researchers are interested in developing a social science that conforms to the aims, rules, and canons of the physical and life sciences--physics, chemistry, biology, and medical science--then we must strive for more than prediction. We must strive, as well, for causal expalantion. Prediction is an important aim. But if it is taken to be the sole interest of social science inquiry, we should recognize that non-causal, predictive systems of causal modeling represent a form of social science inquiry quite different from its

cause-seeking counterpart. For this form has different aims, rules, canons, and results than that of the natural sciences. It represents a different vision of what the social sciences are or ought to become.

But there is a more practical second, point related to the first. Without causal explanation, the findings of the social and behavioral sciences have few clear implications for policy interventions. For policy simply is the attempt to intervene and change the world's established causal structure and its sequence of events. Predictive relationships, however, since they tell us nothing firm about how the causal structure of the world actually is, provide a limited basis for illuminating effective strategies. At best, they may help us to rule out certain strategies should they happen to actually coincide with the causal structure of the world.

Thus, the capacity of social science to play a major role in social change is predicated upon the ability to reveal the prevailing causal structure. This is a task that causal explanation can perform. For once, we understand how things do, in fact, work, we are also in a far stronger position to alter the structure in ways deemed more beneficial to social existence.

For these two reasons alone--for the sake of explanation and the practical relevance of the social sciences--the task of elucidating and testing the rules of causal inference in causal modeling is important. In so doing, the five most plausible understandings of causation will receive separate attention. As we shall see, each conception of cause yields different results as we probe the various rules of causal inference that relate correlation, partial correlation, and causation.

IV. Preliminary Remarks

In sections V, VI; and VII of this paper we shall be concerned with the widespread views on the underlying logic of causal inference for systems where the variables are dichotomus. In section IX we present our results for systems where the variables are continuous.

One view of the logic of causal inference is that the logic for systems with continuous variables is the same as the logic for systems with dichtomus variables. H. Blalock, for example, has put forward such a view (1971, 1972). In sections V, VI, VII we shall investigate the rules of inference which employ the conception of partial correlation used by such people as Blalock.

Other rules of causal inference, however, have been developed primarily for systems with dichotomus variables. These rules have been provided by P. Kendall and P. Lazarfeld (1950), E. Nagel (1961), H. Reichenbach (1956) and P. Suppes (1970). In what follows in sections V, VI, and VII, we show, for example, that the partial correlation as defined by Blalock does not go to zero as the rules of causal influence claim it should. It can be shown that if the Blalock partial correlation does not go to zero, then the Kendall-Lazarsfeld and Nagel partial correlation and the Suppes' kind of partial correlation do not go to zero either. Hence, by showing that the rules which use Blalock's partial correlation are invalid, one also shows the rules advanced by Kendall and Lazarsfeld, Nagel, and Suppes are also invalid. We present these arguments in detail elsewhere.[1]

Our general inquiry, then, is not confined to such writers as H. Simon (1957) or Blalock. Rather, it extends to cover the entire gamut of causal modeling methods which employ various rules to make

causal inferences where the system has variables which take on a continuous range of values. In the literature much as been made about the various strengths and weaknesses of Simon's and Blalock's models which rely upon non-standardized coefficients and regression analysis, on the one hand, and the structural equation model approach which relies upon standardized coefficients and path analysis (see, e.g., Wright, 1960; Blalock, 1971; Goldberger, 1973; Duncan 1975; Asher, 1976). The alleged advantages of the structural equation model approach are that it makes possible the estimation of the numerical values of the path coefficients--and in so doing the estimation of the magnitude or strength of the causal links--and that, as Bentler (1980) notes, it can easily handle models with "latent" or unmeasured variables. Whatever differences there are, then, between Simon-Blalock and path analysis are primarily those of which measure to use rather than a difference in models and underlying rationale.[2] Indeed, Asher (1976, pp. 29, 34-35) claims that, in terms of model specification and determination of the direction of causality, both require the same investment in assumptions. And, since we are not concerned with measurement or estimation issues (e.g., with the measurement of unobservable variables), but only with whether causal modeling methods can, in theory, validly test causal hypothesis, our comments should prove to be general. In section IX we present our results for causal systems where the variables are continuous.

The last point concerns presentation. For simplicity's sake, we shall develop here our results for the dichotomus case in sections V, VI and VII. Hence, we shall discuss the fourfold point correlation, Ø (phi), where the two variables can take on only two values: 1 when present; 0 when absent. Given a 2x2 table the relationship between X and Y can

be represented as follows:

$X_i = 0$, if X is absent in occurrence i.

$X_i = 1$, if X is present in occurrence i.

$Y_i = 0$, if Y is absent in occurrence i.

$Y_i = 1$, if Y is present in occurrence i.

Also let:

a = the probability of an occurrence i where $X_i = 0$ and $Y_i = 1$

b = the probability of an occurrence i where $X_i = 1$ and $Y_i = 1$

c = the probability of an occurrence i where $X_i = 0$ and $Y_i = 0$

d = the probability of an occurrence i where $X_i = 1$ and $Y_i = 0$

The occurrences can be arranged in a 2x2 table of this form:

a	b	Y present
c	d	Y absent
X absent	X present	

The phi correlation coefficient between the scores in the population under study is:

$$\rho_{xy} = \phi = \frac{bc - ad}{[(a+b)(c+d)(a+c)(b+d)]^{\frac{1}{2}}}$$

The partial correlation between X and Z with Y "held constant" is:

$$\rho_{xz,y} = [\rho_{xz} - \rho_{xy} \cdot \rho_{yz}] / [(1 - \rho^2_{xy}) \cdot (1 - \rho^2_{yz})]^{\frac{1}{2}}$$

Notice that the partial correlation is zero if and only if the numerator term is zero. (This is the conception of partial correlation used by such writers as Blalock.)

Again, we have developed results and theorems for cases where the variables are continuous (where the regular Pearson product moment correlation is used). We present, without our proofs, these general results in the Supplement, section IX.

V. Purported Analyses of "Cause" and the Rules Using Partial Correlations

V.A. Two Rules of Causal Inference

In this section we will consider several purported analyses of causation and the relationships among each analysis, correlation, and partial correlation.[2] We employ Blalock's basic framework in which he discusses four conceptions of causation.

Let us begin by formulating the inference rule (R.S.B) which appears to be accepted by many social scientists. It is a straightforward interpretation of Simon's remarks (1957).[3]

> Rule S.B... Given standard assumptions,[4]
>
> if the correlation ρ_{xz} between X and Z is high positive (or negative) and the partial correlation $\rho_{xz,y}$ between X and Z with Y "held constant" is zero, then
>
> either a) Y is an intervening variable--the causal effect of X on Z (or vice versa) operates through Y;
>
> or b) Y is a common cause of X and Z--the correlation between X and Z is "spurious."

Notice that rule S.B. is a confirmational or an inductive rule. That is, rule S.B. purports to able to establish that certain causal relations exist from inferences built on correlations (or some other statistic) among the variables. There is another kind of rule of causal inference which we shall call falsificationist. A falsificationist rule postulates that certain causal relations exist and it then sets out to use the statistics to falsify the hypotheses. The rule C.S.B., which we formulate next, is a falsificationist rule. Path analysis and structural

equation approaches are also falsificationist. (See section VI below). (We believe that certain kinds of falsificationist's rules are valid in certain situations. See section VII.)

Let us formulate an inference rule (R.C.S.B.) which appears to be more widely accepted than Rule S.B. It fits Blalock's remarks that causal inferences don't prove causal relations, rather they rule out or reject causal relations.[5]

Rule C.S.B... Given standard assumptions,

if either (a) Y is an intervening cause variable (the causal effect of X on Z, or vice versa, operates through Y) or (b) Y is a common cause of X and Z; then the correlation ρ_{xz} between X and Z is high positive (or negative) and the partial correlation $\rho_{xz,y}$ between X and Z with Y "held constant" is zero.

Notice that these rules assume that Y is either an intervening variable between X and Z or Y is a common cause of X and Z, but not both. (In Section VI, VII, and the Supplement the assumption is relaxed.) Our point is that given this assumption (and assumptions about the "outside causes"), the two rules of causal inference are invalid for certain conceptions of causation.

V.B. Causes as Necessary Conditions: Type 1

Consider some familiar examples of causation. The striking of a match causes it to ignite or a person drinking a poison causes him or her to die. Of course, were oxygen not present or were the match wet, it would not ignite. And if an antidote were administered, the person would not die. These examples suggest that to talk about the causal

conditions for any change is to talk about those changes that are necessary for its occurrence. Were those conditions not to occur, the change in question would not occur.

Given that a causal condition X of any change Y is a necessary condition of the change Y, what follows about the correlation between X and Y? First, it follows that a = 0. For when X = 0, then Y = 0 because X is a necessary condition for Y; there will be no cases where X = 0 and Y = 1. ρ_{xy} simplifies to the following:

$$\rho_{xy} = [b/(b+d)]^{\frac{1}{2}}[c/(c+d)]^{\frac{1}{2}}.$$

But, it also follows that whenever d is non-zero and whenever X is necessary for Y, no determinate (positive) value of the correlation can be deduced. Any value of ρ, high or low, is logically compatible with the fact that X is a necessary condition for Y! (Note, however, that ρ must be non-negative.)

On the other hand, given only that a = 0, it does not follow that X is a necessary condition for Y. It could turn our that X is merely correlated with the actual causal condition, but played no role in the causation. It is quite possible that no match has ever ignited except in the presence of some gravitational force, yet the presence of such a force is not causally necessary for the ignition of the match.

The important point to note in such cases as the match and the quaffed poison is that <u>any</u> positive value, high or low, of the ρ correlation is compatible with the fact that x is causally necessary for Y. The actual value will depend upon the number of occurrences of such events as the match's being wet, the oxygen's being absent, and the antidote's being administered, and so on.

Given the analysis that causes are necessary conditions, what can be shown about the causal inference rules which relate correlation, partial correlation, and a conception of causation?

It is easy to find cases which invalidate rule S.B., cases where the correlation ρ_{xz} is high (positive or negative) and the partial correlation $\rho_{xz,y}$ is (near) zero, but where it does not follow that Y is an intervening cause or a common cause of X and Z. Suppose Y is a necessary condition for Z, but X is not necessary (nor sufficient) for Y. Here some subset of not-X causes Y. Since there is no causal path from X to Y, Y is not an intervening variable (and it isn't a common cause either). Here the antecedents of rule S.B. are satisfied but the consequences are not.

Given that rule S.B. is invalid, it might be hoped that rule C.S.B. is valid. Unfortunately, rule C.S.B. is also invalid. Suppose Y is a common, necessary cause of X and Z.[6] Consider the common cause Y where we have the following correlation tables:

0	25	X
50	25	not X
not Y	Y	

0	25	Z
50	25	not Z
not Y	Y	

0	25	Z
75	0	not Z
not X	X	

Here the antecedents of rule C.S.B. are satisfied, but the partial correlation between X and Z with Y "held constant" does not go to zero. (For, $\rho_{xz} \neq \rho_{xy} \cdot \rho_{yz}$. Remember, the numerator of $\rho_{xz,y}$ is $\rho_{xz} - \rho_{xy} \cdot \rho_{yz}$.)

V.C. Causes as Sufficient Conditions: Type 2

Let us consider another set of familiar examples of causation. For example, disturbances on the sun cause radio wave interference on earth or the explosion of an atom bomb causes a building to collapse. Here we

have the feeling that no matter what else happens, the sun's disturbances would cause radio wave interference and the bomb's blast would collapse the building. These examples suggest the view that a cause is sufficient for its effect. In more general form, to say a set of conditions is a cause is to say that the set of conditions is sufficient for the effect.

Blalock, for one, apparently argues that a cause is a sufficient condition for the event:

> According to Bunge, one of the essential ingredients in the scientist's conception of a cause is the idea of "producing", a notion that seems basically similar to that of forcing. If X is a cause of Y, we have in mind that a change in X is followed by or associated with a change in Y. Thus, although constant conjunction may be a part of one's definition of causality, conjunction is not sufficient to distinguish a causal relationship from other types of associations. (1972, p. 9; emphasis in the original)

Two important points need to be made here. First, it is unfortunate that Blalock never attempted to give any independent explications of "producing" or "forcing". For example, he might have attempted to link producing to human manipulability or to examples such as the bomb blast where the forcing is very evident. On the other hand, it is easy to see how difficult it is to model the forcing notion with statistical concepts. Second, constant conjunction for Blalock comes to the view that all X's are followed by Y's, yet he clearly distinguishes constant conjunction from causation. Thus, Blalock appears to be committed to the view that to say X is the cause of Y is to say X is sufficient for Y. This view is compatible with the way Blalock goes about "finding" causes (1972, Chs. II and III).

Further support of this interpretation is provided by Simon's (1957) more formal analysis of causation. In that paper Simon explicates a version of what has been termed the "covering law" or hypothetical-deductive model of scientific explanation. The model can be briefly specified as having the form:

$$\begin{array}{l} X \supset Y \\ \underline{X} \\ \therefore \quad Y \end{array}$$

Where $X \supset Y$ is a general law specifying that if X occurs, then Y occurs and X specifies instantiations of the initial conditions and Y the effect. In other words, Simon provides a sufficient condition analysis of causation.[7] It is not unreasonable to believe that Blalock simply took it over from Simon.

At any rate, if causes are a sufficient condition for their effects, we cannot infer the existence of X from the presence of Y. There can be a plurality of causes of Y, each sufficient, but none necessary for Y. But, suppose now that X does cause Y. What follows about the correlation between X and Y? First, it follows that d=0. For when X = 1, Y = 1 because X is sufficient for Y. Hence, there will be no cases where X = 1 and Y = 0. Given that d = 0, then ρ becomes:

$$\rho = [b/(a+b)]^{\frac{1}{2}} \bullet [c/(a+c)]^{\frac{1}{2}}$$

But it can then be determined that when a is non-zero and X is sufficient for Y, no determinate value, high or low, of the ρ correlation can be deduced. The value of ρ depends on the actual number of occurrences in the 2x2 table cells--on the numbers a, b, and c. Any (positive) value of ρ is logically compatible with the fact that X is sufficient for Y.

On the other hand, given that d = 0, it does not follows that X is a sufficient condition for Y. It could turn out that X is merely correlated with Y.

It is easy to find cases which invalidate rule S.B., cases where the correlation ρ_{xz} is high positive (or negative) and the partial correlation $\rho_{xz,y}$ is (near) zero, but where it does not follow that Y is an intervening sufficient cause or a common sufficient cause of X and Z. Suppose that X is a sufficient cause for Y, but Y is neither necessary nor sufficient for Z. Here some subset of notY causes Z. Since there is no causal path from Y to Z, Y is not an intervening variable (and it isn't a common cause either). Here the antecedents of rule S.B. are satisfied, but the consequences are not.

Rule C.S.B. is also invalid. Suppose Y is a common sufficient-cause of X and Z. Consider the common cause Y where we have the following correlation tables:

	0	50	Y
	200	50	not Y
	not X	X	

	100	50	Z
	150	0	not Z
	not Y	Y	

	100	50	Z
	100	50	not Z
	not X	X	

Here the antecedents of rule C.S.B. are satisfied, but the partial correlation between X and Z with Y "held constant" does not go to zero. (For $0 = \rho_{xz} \neq \rho_{xy} \cdot \rho_{yz} > 0$.)

Of course, there are several problems with attempting to analyze causation as a sufficient condition. One counter example to the analysis is the fact that the presence of fire is sufficient for the presence of oxygen, yet surely the fire did not cause the presence of oxygen. Furthermore, if X is a sufficient condition for Y, then any other state S of the world which occurs can be conjoined with X so that X and S are

sufficient for Y. For example, if frustration were sufficient for aggression, then frustration and Antarctica's being cold would also be sufficient for aggression. Clearly, however, Antarctica plays no causal role in aggression. Hence, there are difficulties with the view that causes are sufficient conditions.

Even given Blalock's preferred interpretation of causation as a sufficient condition, we have shown that any positive value (high or low) of the correlation ρ is logically compatible with the fact that X caused Y. We have also shown that the two widely held causal inference rules are invalid. It remains to be seen whether the method will prove more adequate with our three remaining understandings of cause.

V.D. Causes as Necessary and Sufficient Conditions: Type 3

The Type 1 analysis of cause involves a restrictive notion of what a cause is. For example, it would rule out the possibility that there might be a plurality of causes for a certain effect in which each cause is sufficient for (or increases the probability of) an effect, but none of which is necessary. (For instance, there are many ways of causing a person's death or of bringing about academic achievement, but not a one and only way.) And we have seen there are problems with the Type 2 analysis. But perhaps it is still possible to hold that if X causes Y, then X (or the totality of necessary conditions for Y) is necessary and sufficient for Y. Such an analysis was widely held in the 1950s and early 1960s.

Suppose, then, that "X is the cause of Y" means "X is a necessary and sufficient condition for Y." First, notice that a = 0, for when X =

0, Y = 0 because X is necessary (there can be no cases in which X = 0 and Y = 1). Second, notice that d = 0, for when X = 1, Y = 1 because X is sufficient for Y (X cannot = 1 and Y = 0). Given that a = 0 and d = 0, the correlation ρ equals 1.

It should be pointed out, of course, that given a correlation of 1, we cannot directly infer that X caused Y. X and Y can be perfectly correlated, but not causally related, if X and Y are the unalterable consequences of some third event Z; X could be the side effect of a virus (Z) and Y the disease.

At this point, using correlations as a method in causal inquiry appears to be justified. If the view of causation is sound, then it is a necessary condition that the correlation be 1. Although it does not necessarily prove causal relationships, one can use this as a test to reject causal claims.

However, given our previous arguments and examples, it follows that rule S.B. is invalid here, too. But rule C.S.B is no longer susceptible to the counterexamples of the previous sections. When a cause is a necessary and a sufficient condition, it follows that the correlation ρ_{xz} is high (1) and that the numerator of the partial correlation $\rho_{xz,y}$ is zero. Except for the practical difficulties encountered in finding such causes, it appears the position is theoretically sound. But as we shall see, other difficulties quickly arise.

Blalock (1972, p. 31) has anticipated this analysis of causation which states that strictly speaking cases of type 1, type 2, and type 4 (causes as neither necessary nor sufficient conditions; to be discussed below) can all be reduced to necessary and sufficient conditions. This can be done, he maintains, by the device of re-defining what is included

under X and Y. As we have already noted, however, Blalock has mostly espoused and applied the view that causes are merely sufficient conditions. Furthermore, he strongly relies (1972, pp. 40, 68-69, 73-76) on what may be termed "the principle of decreasing correlation" which states that when X causes Y and Y causes Z, one can expect the correlation between X and Z to be smaller than that between X and Y or Y and Z. But the problem is that this principle is applicable only if there be other causes for Y and Z. Thus, in his many exemplifications of Simon's model, Blalock has implicitly assumed that a cause X of Y need not be the only cause of Y, that X is not necessary for Y. (If a cause were necessary and sufficient for its effect, the ρ correlation would be 1.) Blalock has apparently presented inconsistent views.

Even were Blalock to present specific arguments and procedures for reducing all cases to those of type 3, a major theoretical difficulty attends the claim that causes are necessary and sufficient conditions. For an adequate analysis of causation should be able to distinguish cause from effect. "X is the cause of Y" should not entail "Y is the cause of X"; the causal relationship is an asymmetric one. Yet if X is sufficient for Y, then Y is necessary for X; and if X is also necessary for Y, then Y is sufficient for X. If "X is the cause of Y" means "X is necessary and sufficient for Y", then it also follows that Y is the cause of X. Thus the asymmetry of the causal relationship has been violated. Therefore, this analysis of causation is inadequate.

Blalock is aware of this difficulty. And like most other writers on causal modeling (e.g., Asher, 1976 and Kenny, 1979), he has noted that because temporal sequences are also asymmetrical, time precedence of one variable over another might help to resolve the direction of

causation. But, unlike most other writers, Blalock is less comfortable with the time sequence-gives-you-causal-asymmetry-view. He is explicitly aware that the forcing or producing idea is not contained in the notion of temporal sequences. Thus, he argues that our conception of causality should not depend on temporal sequences, except for the impossibility of an effect preceding its cause (1977, p. 10). Blalock, we think, is mainly correct on this observation. However cause is understood, it cannot be part of the definition of cause that it temporally precede its effect. This has implications for ordering variables in a model. But even Blalock appears to violate this caveat when he later seems to employ temporality to distinguish cause and effect (1972, p. 43).

Counterexamples to the temporal-sequence condition fall into two sets. By all reason, the first set--contemporaneous causes and effects--is the larger. For example, when a locomotive is pulling a caboose and the two are tightly coupled, the motion of the locomotive is sufficient for the motion of the caboose. Once primary inertia has been overcome, the motion of the caboose is also sufficient for the motion of the locomotive. But here they move at the same time. Furthermore, plentiful examples of this sort might be drawn from psychology and education.

The second set of counterexamples are more exotic since they involve effects which, partly at least, precede their causes in time. Evidently, many physicists believe that this is what occurs with certain highly charged sub-atomic particles in a cloud-chamber. Many philosophers of science have also avoided ruling out, a priori, the possibility of "backward causation " (See, e.g., Sayre, 1977 and Dummett and Flew, 1954).

One last example is instructive. The change in atmospheric conditions is regularly associated with the falling barometer, although it is the change in atmosphere that causes the storm. Here, both the atmospheric and barometer changes temporally precede the storm. Hence, the difference between regular association and causal relationship is not merely a matter of temporal sequence.

Yet more troubling for the Simon-Blalock procedures, even in a case such as this, one cannot use the partial correlations to distinguish cause from spurious cause. For though all the pairwise correlations are 1, the partial correlation is undefined. The difficulty is that the denominator of the partial correlation is zero. Hence, the rule C.S.B. is actually invalid afterall. Blalock (1972, pp. 87-89), recognizes this problem of multicollinearity and admits that there are no defenses available (1972, pp. 87-89).

Thus, such a view of causal inference does not allow us to distinguish spurious from causal relations in general, given a type 3 view of cause. This is unfortunate since actual cases in the social and behavioral sciences are likely to be many times more recalcitrant than the barometer example. But then again, the conception of cause as necessary and sufficient condition is probably the least plausible one that we have examined. We turn to type 4 cases which may prove to be a most plausible understanding of causation. It is also the conception held by many social scientists.

V.E. Causes as INUS Conditions: Type 4

For one reason or another most social scientists have rejected the conceptions of causation which treat causes as a sufficient condition or

which treat causes as a necessary condition for some event E. Social scientists are inclined to hold that such invariable and close relationships between a cause and its effect are not likely to be found in social inquiry. (For example, see D. Heise, 1975)

The use of such strategies as causal modeling or path analysis can be reconstructed by using either the conception of causation as an INUS-condition or the conception of causation as probabilistic causation. In this section we consider the validity of the two rules when the former conception is used; in the next section we consider the validity of the rules when the latter is used.

The INUS conception of causation holds the following. To say that X is the cause of Z is to say that X is an insufficient but nonredundant part of a set of conditions which is unnecessary but sufficient for Z. Thus, wherever there is a plurality of causes of Z, and wherever a conjunction of conditions which includes X is sufficient for Z, X will turn out to be an INUS condition of Z, (Heise, 1975; Mackie, 1974). In more abstract terms, it may turn out that all cases of (AX or WB) are also cases of Z. It can be seen that an INUS conception of causation requires that strict determinism be true locally, that there be sufficient conditions of Z.

Given that X causes Z means that X is an INUS causes of Z, what follows about the correlation between X and Z? For simplicity, let's assume that Z can be sufficiently-caused by [(X and C) or W]. Consider the terms in the 2 x 2 correlation table. First, it follows that d is not zero, for X is not sufficient for Z. Together X and C bring about Z, but when notC occurs with X, then notZ occurs. Thus, it follows that

b is not zero and that the ratio b/(b+d) equals the number of occurrences of (X and C) divided by the number of occurrences of X. Second, it follows that a is not zero, for X is not necessary for Z; W can also bring about Z. It is also possible that notW occur with notX, and hence that not Z occur. At any rate, the ratio a/(a+c) equals the ratio of the number of occurrences of (notX and W) divided by the number of occurrences of notX. Therefore, when X is an INUS cause of Z, the terms a, b, and d are non zero and it is possible for c to be non-zero also. Furthermore, the ratios d/b and c/a can take on values such that ρ_{xz} ranges from high positive to high negative. Also, ρ_{xz} can equal zero when d/b equals c/a.

Given the analysis that X causes Z means that X is an INUS cause of Z, what can be shown about the causal inference rules which relate correlation, partial correlation and the INUS conception of causation?

It is easy to find cases which invalidate rule S.B., cases where the correlation ρ_{xz} is high (positive or negative) and the partial correlation $\rho_{xz,y}$ is (near) zero, but where it does not follow that Y is an intervening cause between X and Z nor that Y is a common cause of X and Z. (Remember that the numerator of $\rho_{xz,y}$ equals $\rho_{xz} - \rho_{xy} \cdot \rho_{yz}$, so the partial is zero if and only if the latter quantity is zero.) Suppose Y is an INUS cause of Z but that X is not on INUS cause of Y. Consider the INUS cause Y where we have the following correlation tables:

96	0	Y
24	120	NotY
NotX	X	

108	12	Z
36	84	NotZ
NotY	Y	

30	90	Z
90	30	Not Z
NotX	X	

Here the antecedents of the rule are satisfied ρ_{xz} is high positive (.5) and the partial correlation $\rho_{xz,y}$ is zero), but Y is not an intervening cause between X and Z (nor is Y a common cause of X and Z). (For .5 $= \rho_{xz}$ which equals $\rho_{xy} \cdot \rho_{yz} = [(-96)/(96 \cdot 144)^{\frac{1}{2}}].[(-72)/(96 \cdot 144)^{\frac{1}{2}}].$)

Given that rule S.B. is invalid, it might be hoped that rule C.S.B. is valid. Unfortunately, rule C.S.B. is also invalid. There are cases where $\rho_{yz} = 0$ and where the partial correlation is zero, but there are also cases where the partial correlation is not zero. Suppose X is an INUS cause of Y and Y is an INUS cause of Z. Consider the following correlation tables:

25	75	Y	50	50	Z	62.5	37.5	Z
75	25	NotY	50	50	NotZ	37.5	62.5	NotZ
NotX	X		NotY	Y		NotX	X	

Hence ρ_{xz}, which is large negative, does not equal $\rho_{xy} \cdot \rho_{yz}$, for ρ_{yz} equals zero. Thus, even though Y is an intervening INUS-cause between X and Z, the partial correlation $\rho_{xz,y}$ is not zero.

There are cases where Y is an intervening INUS-cause between X and Z and where ρ_{yz} is nonzero but the partial correlation $\rho_{xz,y}$ is zero, but there are also cases where the partial correlation is not zero. Suppose X is an INUS-cause of Y and Y is an INUS-cause of Z. Consider the following correlation tables:

25	75	Y	16 2/3	33 1/3	Z	25	25	Z
75	25	NotY	83 1/3	66 2/3	NotZ	75	75	NotZ
NotX	X		NotY	Y		NotX	X	

Here ρ_{xz}, which is zero, does not equal $\rho_{xy} \cdot \rho_{yz}$, both of the latter terms are nonzero. Therefore, Y is an intervening INUS cause but the partial correlation $\rho_{xz,y}$ is not zero.

Given these two kinds of counter examples, then, it follows that when Y is an intervening INUS cause of X and Z the rule C.S.B. is invalid.

Difficulties also arise for the rule C.S.B. when Y is taken to be a common INUS cause of X and Z. Consider the following correlation tables:

0	75	Z
125	0	NotZ
NotX	X	

25	50	Z
75	50	NotZ
NotY	Y	

25	50	X
75	50	NotX
NotY	Y	

Here ρ_{xz}, which equals 1, does not equal $\rho_{xy} \cdot \rho_{yz}$, the latter correlations are both less than 1. In order for Y to be a common INUS cause of X and Z there must be some other factor(s) which is (are) a cause of X otherwise Y couldn't be an INUS cause of X. Similarly, there must be some other factors which is (are) a cause of Z; otherwise, Y couldn't an INUS-case of Z. The counterexample above assumed that the other factor is also an INUS common cause of X and of Z, but other kinds of counterexamples are easily constructed without this assumption. At any rate, given these counterexamples, it follows that when Y is a common INUS cause of X and of Z, it need not be the case that ρ_{xz} is high (positive or negative) and that the partial correlation $\rho_{xz,y}$ is zero. Therefore, the rule C.S.B. is invalid.

There are, of course, several problems with the analysis of causation as an INUS condition. First, suppose Y is an INUS condition of Z. Then Y is part of factor which has the following form: C•Y or W. If Y is an INUS conditon of Z, then so is C. But usually we select only one

of these INUS conditions and call it the cause. At best, then the INUS-analysis of causation is incomplete.

Second, suppose that Y is an INUS cause of Z and that X is an INUS cause of Y. In such a situation, z is an insufficient but nonredundant part of a set of conditions which are unnecessary but sufficient for Y. That is, Z is an INUS condition of Y. But if a cause is merely an INUS condition, then in this situation it follows that Y is the cause of Z and that Z is the cause of Y. Surely this is an unacceptable result, for the causal relationship is an asymmetric one. An adequate analysis of causation should be able to distinquish cause from effect. Since the analyses of causation as an INUS condition does not, it is not an adequate analysis.

Again, one might be tempted at this point to add temporality to the INUS condition analysis. After all, temporal sequences are also asymmetrical; time precedence of one variable over another might help to resolve the direction of causation. But, we have previously seen in our discussion of cause as a necessary and sufficient condition (Section V.D.) that the temporal sequence condition is subject to serious counterexamples. Thus, it cannot be used to save the INUS analysis of cause from this defect.

Still, there is one final problem for the analysis of causation as an INUS condition. An INUS condition Y is part of a set of conditions which is unnecessary but sufficient for Z. Thus, in order for Z to have an INUS condition, some set of conditions must be sufficient for Z. One might say that locally, at least, determinism is true. But many writers have argued that modern physics and biology use a conception of causation--probabilistic causation--which rules out such local determinism. In

such situations, W is a probabilistic cause of Z even though W does not belong to a set of conditions which is sufficient. In particular, many writers have held that $P(Z/W) > P(Z)$. We consider probabilistic causation in detail in section V.F.

Given that there are intelligible applications of the concept of probabilistic causation, then the analysis of causation as an INUS condition is unacceptable as a general analysis. It may turn out, however, that there are several related but distinct conceptions of causation. Perhaps the INUS condition analysis is adequate for one of these conceptions of causation. We leave this issue to another time.

We must notice in passing from causation as an INUS condition to probabilistic causation, that there seems to be room for an INUP analysis of causation. Y is an INUP cause of Z if Y is neither a sufficient nor a probabilistic cause but is a nonredundant part of a set S of conditions which is an unnecessary but a probabilistic cause of Z. Although Y does not "influence" Z by itself, the set S, which has Y as a nonredundant part, does "influence" Z. We leave consideration of INUP conditions to another time.

V.F. Cause as Probabilistic-cause: Type 5

A probabilistic cause can be partly understood by contrasting it to the familiar notions of necessity and sufficiency. Roughly, to say that 'X is a probabilistic cause of Y' is to say 'X causes Y even though X is insufficient and often unnecessary for Y.' In more formal terms, many writers have held that $P(Y/X) > P(Y)$. In other words, it has been held that the presence of X increases the probability that Y will occur.

Now many may find that the notion of probabilistic causation, rather than teeming with plausibility, is highly problematic. Somehow,

it might be maintained, the notion 'X causes Y' must be reducible to necessary or sufficient conditions or both. Thus, Mackie's (1974, p. 62) notion of an INUS condition (as discussed in V.D.) appears to be a special instance of a probabilistic cause, but it is really merely part of a sufficient condition. To say that 'X causes Y' is to say that 'X is an insufficient but necessary factor in a set of conditions that are unnecessary, but sufficient for Y.' In other words, X belongs to a set of conjuncts in a series of disjuncts, none of which are necessary, but all of which are sufficient for Y. For example, the effect is brought about if XC or AB or LM... Thus, with greater specification, we obtain invariable causal relations (if we are able to specify all of the disjuncts, the entire set would be necessary as well as sufficient for Y). An INUS condition requires that strict determinism be true. A probabilistic cause does not require this.

But such a hoped-for reduction to invariability in all cases appears to be more a matter of faith than being faithful to the facts. From medical research, we know that a variety of substances cause cancer even though only a small percentage of those exposed contract the disease. In quantum mechanics, indeterminancy and, hence, probabilistic causation plays an important role. And given the kind of creatures that human beings are, it is very likely that the notion of probabilistic cause has a large role to play in the social and behavioral sciences. To hanker after a reduction to invariable causal relations is, as Suppes argues, as pointless as it is unjustified (1970, pp. 7-8). (For other major treatments of probabilistic causation, see Good, 1961-62; Reichenbach, 1956; and Salmon, 1980).

But now, assuming the conception of causation as probabilistic cause, what follows for our standard methods? Following arguments similar to those involving INUS cause, it can be seen that X causes Y is compatible with any value of the ρ correlation. In such cases, the value, high or low, of the ρ correlation cannot be used to defend causal claims, nor can it be used to reject causal claims. Not only may the value of the correlation be high (or low) and positive (or negative) it may also be zero!

There are several kinds of counterexamples to show that when probabilistic causation is involved the rule S.B. is invalid. There are cases where ρ_{XZ} is high negative and the partial correlation $\rho_{XZ,Y}$ is zero, but where Y is neither an intervening cause between X and Z nor is it a common probabilistic cause of X and of Z. Also, there are cases where ρ_{XZ} is high positive and the partial correlation $\rho_{XZ,Y}$ is zero, but where Y is neither an intervening cause nor a common cause. Suppose Y is a probabilistic cause (.125) of Z but that X is neither a sufficient cause nor a probabilistic cause of Y. Consider the following correlation tables:

.125	.375 Z	.40	.00	.45	.05
.375	.125 NotZ	.10	.50 NotY	.15	.35 NotZ
NotX	X	NotX	X	NotY	Y

Here ρ_{XZ} is high positive and $\frac{1}{2} = \rho_{XZ} = \rho_{XY} \cdot \rho_{YZ} = (.816)\cdot(612)$, but Y is not an intervening cause between X and Z, for X is not a cause of Y! (And Y is not a common cause either.) Given these counterexamples, then, rule S.B. is invalid.

Counterexamples can be found which also invalidate rule C.S.B. when causation is analyzed as probabilistic causation. There are cases where $\rho_{yz} = 0$ and the partial correlation $\rho_{xz,y}$ is zero, but there are cases where $\rho_{yz} = 0$ and the partial correlation is not zero. Suppose X is a probabilistic cause (.5) of Y, Y is a probabilistic cause (.5) of Z; suppose that notX is a probabilistic cause (.5) of Y and notX is a probabilistic cause (.75) of Z. (For there is a hidden fourth variable which never occurs when either X or Y occur and which is a probabilistic cause of Z.) Consider the following correlation tables:

.25	.25	Y		.25	.25	Z		.375	.125	Z
.25	.25	NotY		.25	.25	NotZ		.125	.375	NotZ
NotX	X			NotY	Y			NotX	X	

Here $\rho_{xz} \neq \rho_{xy} \cdot \rho_{yz}$, even though Y is an intervening probabilistic cause.

There are also cases where Y is an intervening probabilistic cause (between X and Z) where ρ_{yz} is nonzero and the partial correlation $\rho_{xz,y}$ is zero, but there are also cases where the partial correlation is nonzero. Suppose X is a probabilistic cause (.5) of Y, Y is a probabilistic cause (.5) of Z, and notX is a probabilisitic cause (.5) of Z. (For there is a hidden fourth variable which never occurs when X or Y occur and which is the probabilistic cause of Z.) Consider the following correlation tables:

0	.25	Y		.25	.125	Z		.25	.125	Z
.50	.25	NotY		.50	.125	NotZ		.25	.375	NotZ
NotX	X			NotY	Y			NotX	X	

Here $\rho_{xz} \neq \rho_{xy} \cdot \rho_{yz}$ even though Y is an intervening probabilistic cause. This set of counterexamples assumed only that Z had more than one probabilistic cause!

Difficulties also arise for rule C.S.B. when Y is taken to be a probabilistic common-cause of X and Z. Although there are several kinds of counterexamples, the most striking is the following. Suppose Y causes X with probability .5; Y causes Z with probability .5; but Y causes (X or Z) with probability 1. (Consider the case of a coin toss or the case of radioactive decay in nonoverlapping time intervals.) Consider the following correlation tables:

.25	0	Z
.50	.25	NotZ
NotX	X	

0	.25	Z
.50	.25	NotZ
NotY	Y	

0	.25	X
.50	.25	NotX
NotY	Y	

In such cases, the correlation ρ_{xz} will always be negative. And since Y is a necessary probabilistic cause for X and a necessary probabilistic cause for Z, ρ_{yx} and ρ_{yz} will always be positive. Thus, $\rho_{xz} \neq \rho_{xy} \cdot \rho_{yz}$. In this kind of case, only Y is available to serve as the common cause of X and Z, but Y doesn't cause X and Y in an "independent" manner.

Thus, when causation is conceptualized as probabilistic causation, a conception which is used by many social scientists, both the rule S.B. and the rule C.S.B., which involve partial correlations are shown to be invalid rules of causal inference. In particular, X can be a probabilistic cause of Z and the correlation be high (or low) positive (or negative) or zero. Furthermore, the partial correlations being zero is neither a sufficient nor a necessary condition for Y's being an intervening cause between X and Z or for Y's being a common-cause of X and Z.

VI. Problems with Path Analysis/Structural Equation Approaches

In the previous sections it has been shown that the causal inference rules S.B. and C.S.B. which link correlation, partial correlation

and a conception of causation are invalid for certain conceptions of causation (given certain assumptions about the error variables). Of particular importance is the finding that if Y is an intervening causal factor between X and Z, or if Y is a common cause of X and of Z, but not both, then it can be shown that the partial correlation $\rho_{xz,y}$ need not go to zero when the probabilistic or INUS conceptions of causation are employed.

In our discussion up to this point, however, it has been implicitly assumed that Y is an intervening cause or a common cause, but not both. Many have claimed that approaches such as Simon-Blalock causal modeling also makes such an assumption (Asher, pp. 19-20). At any rate, it is now time to relax the assumption and consider cases where a cause X has both direct and indirect effects on Z.

It has been argued that one virtue of approaches such as path analysis and of structural equation models is that they allow one to consider causes which have both direct and indirect effects. (Asher, pp. 32-35; see also Duncan, 1975.)

Before we relax our assumption, however, notice that when there are no causes which have both direct and indirect effects, the path analysis/ structural equation approaches lead to a similar analysis as the Simon-Blalock causal modeling approaches. Thus, as we have shown in the previous sections, both Simon-Blalock causal modeling and path analysis/ structural equations give the wrong answers for certain conceptions of causation given certain assumptions about the "error" variables.

The results of Sections IV and V show that where the probabilistic or the INUS conception of causation is used and where one assumes the pairwise mathematical expectations of the "outside" causes (or error

variables) are zero, the two rules are invalid because the partial correlation need not go zero. We have also derived the result that where the probabilistic or the INUS conception of causation is used and where one assumes the pairwise correlations of the outside causes (or error variables) are zero, it can be shown that the partial correlation is zero where X causes Y and Y causes Z, but the partial correlation need not equal zero where X is a common cause of Y and Z.

Thus, suppose we are considering a single chain where X causes Y and Y causes Z. Contrary to the views of P. Lazarsfield and P. Kendall, E. Nagel, Patrick Suppes, and Hans Reichenbach, we have found that the partial correlation need not go to zero if the pairwise expectations are zero. We have found that the partial correlation must go to zero if the pairwise correlations are zero.

In the remainder of this section we will use the probabilistic conception of causation and we shall assume that the pairwise correlations of the outside causes are zero. Given these assumptions we shall show that such approaches as path analysis and/or structural equation models give the wrong answers. (In the Supplement we present the general results for dichotomus and continuous cases for both the assumption that the pairwise expectations of the outside causes are zero and the assumption that the pairwise correlations are zero.)

Approaches similar to path analysis and structural equation analysis, then, claim that they can handle causes which have both direct and indirect effects. Furthermore, such approaches claim that they can then compare the magnitude of the direct and indirect effects which would identify the operative (or underlying) causal mechanisms. (For example, see Asher, esp. p. 32f; Duncan, 1975, ch. 3, 4.)

Also, there are several equivalent path analysis techniques which enable one to decompose the correlation between any two variables into a sum of simple and compound paths. Some of these compound paths will be causally (or substantively) meaningful indirect effects while others may not be (Asher, p. 32). Whether a compound path is causally meaningful depends on both the equations and one's theory or model of the causal processes. Path-analysis and structural equation approaches are basically falsificationist; they hypothesize a causal structure and see if the observational consequences conform to it. (See, for example, Duncan, 1975, p. 47).

Suppose X is a probabilistic cause of Y (with probability α) and that Y is a probabilistic cause of Z (with probability β). Suppose also that X is a probabilistic cause of Z (with probability ξ). Thus, X is both a direct and indirect cause of Z. We want to show that the basic equations developed by path analysis and structural equation approaches do not apply to this causal system. The basic set of equations is the following (see Asher, pp. 32-33; Duncan, pp. 51-53):

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = \alpha \qquad \ldots \text{(a)}$$

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = \rho_{xy}\frac{\sigma_y}{\sigma_x}\cdot\beta + \rho_{xx}\cdot\xi \qquad \ldots \text{(b)}$$

which can also be written

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = \alpha\cdot\beta + \xi \qquad \ldots (b^1)$$

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = \rho_{yy}\cdot\beta + \rho_{xy}\cdot\frac{\sigma_x}{\sigma_y}\cdot\xi \qquad \ldots \text{(c)}$$

which can also be written

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = \beta + \alpha\cdot\xi\left(\frac{\sigma_x}{\sigma_y}\right)^2 \qquad \ldots (c^1)$$

In these equations, α, β, ξ are supposed to represent the magnitidues of causal influence of one variable on another. For example, α is supposed to represent the magnitude of causal influence of X on Y.

For the moment we'll only be concerned with the magnitude of the direct causal influence of X on Y--i.e., α. Given the suppositions made above, path analysis/structural equation approaches claim that this magnitude α should equal $\rho_{xy} \frac{\sigma_y}{\sigma_x}$. We believe this is a serious mistake.

Remember, we are supposing that the correlation between the "outside cause" of X and the "outside cause" of Y is zero, and that one is using the probabilistic conception of causation where X is the probabilistic cause (α) of Y. Given these assumptions, we have found that

$$\rho_{xy} \cdot \frac{\sigma_y}{\sigma_x} = (\alpha) \cdot (1-P_2),$$

where P_2 is the probability of the "outside cause" of Y. (Note also that the expectation of the outside cause is P_2.) Unless $P_2 = 0$, it is <u>not</u> the case that $\rho_{xy} \cdot \frac{\sigma_y}{\sigma_x}$ equals α, where α is the probabilistic cause of Y by X. Hence, the path analysis/structural equation approaches do not apply, unless $P_2 = 0$!

Furthermore, suppose for the moment that X is a sufficient condition of Y. Whenever X occurs, so does Y. Given that there are other (outside) causes (R_u) of Y, however, the expression $\rho_{xy} \cdot \frac{\sigma_y}{\sigma_x}$ will change if we vary the probability of R_u, even though X remains a sufficient condition for Y. Surely this is a doubly absurd result. First, if X is sufficient for Y, then it ought to take on the maximum magnitude for causal influence (i.e., 1). Second, X's causal influence on Y ought to be independent of the other causal influences on Y. Given that path analysis/structural equation approaches lead to such problematic results, they must be rejected as unjustifiable.

There are, however, further problems for the path analysis/structural equation approaches.

Problem 2. Suppose X is a probabilistic cause (.5) of Y, yet if X doesn't lead to Y, then it leads to Z. But X never leads to both Y and Z. Suppose also that Y is a probabilistic cause (1.0) of Z. Again, then X is both a direct and indirect cause Z. For simplicity, suppose also that no other factors cause Y and that no other factors cause Z. Consider the following correlation tables:

0	.67	Z		.33	.33	Z	0	.33	Y
.33	0	Not Z		.33	0	Not Z	.33	.33	Not Y
Not X	X			Not Y	Y		Not X	X	

If one followed the path analysis/structural equation approaches in this situation, one might expect

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = \alpha = 1/2 \qquad \ldots(i)$$

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = \alpha\cdot\beta + \gamma = (1/2)\cdot 1 + 1/2 = 1 \qquad \ldots(ii)$$

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = \beta + \alpha\cdot\gamma = 1 + (1/2)(1/2) = 5/4 \qquad \ldots(iii)$$

But if we calculate the appropriate correlations and variances for the left hand side we get

$$1/2 = 1/2 \qquad \ldots(i^1)$$

$$1 = \alpha\cdot\beta + \gamma = 1 \qquad \ldots(ii^1)$$

$$1/2 = \beta + \alpha\cdot\gamma = 5/4 \qquad \ldots(iii^1)$$

Thus, the path analysis/structural equation approaches lead to absurd results. Clearly, the equations one derives from such approaches do not hold for such a causal system, a system which involves probabilistic causation.

Problem 3. Suppose that X is a probabilistic cause (.5) of Y and that X is also an (independent) probabilistic cause (.5) of Z. Suppose also that Y is a probabilistic cause of (1.0) of Z. Suppose there are no other causes of Y, and no other causes of Z. Consider the following correlation tables:

0	.50	Z
.33	.17	Not Z
Not X	X	

0	.33	Y
.33	.33	Not Y
Not X	X	

.17	.33	Z'
.50	.0	Not Z
Not Y	Y	

The path analysis/structural equation approaches would lead one to expect that:

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = \alpha = \tfrac{1}{2} \quad \ldots(iv)$$

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = \alpha\cdot\beta + \xi = \tfrac{1}{2}\cdot 1 + \tfrac{1}{2} = 1 \quad \ldots(v)$$

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = \beta + \alpha\cdot\xi = 1 + \tfrac{1}{2}\cdot\tfrac{1}{2} = 5/4 \quad \ldots(vi)$$

But, if we calculate the appropriate correlations and variances for the left hand sides, we get:

$$\tfrac{1}{2} = \alpha = \tfrac{1}{2} \quad \ldots(iv^1)$$

$$3/4 = \alpha\cdot\beta + \xi = 1 \quad \ldots(v^1)$$

$$3/4 = \beta + \alpha\cdot\xi = 5/4 \quad \ldots(vi^1)$$

Thus, the path analysis/structural equation approaches again lead to absurd results. Clearly, the equations do not hold for such causal systems that involve probabilistic causation.

Given these difficulties and problems with the basic logical structure of these approaches, we conclude that such approaches to causal analysis (or causal inference) are unjustifiable for probabilistic causation (and for INUS causation). Accounting for the variance between variables is one task; determining the degree of causal influence of one variable on another variable is a quite different task. (In the Supplement we present our general theorems for probabilistic causation which set out the proper relations).

VII. An Alternative Approach (The Conditional Probability Approach)

Here we would like to outline an approach to causal analysis (or causal inference) which we shall call the conditional probability approach. This approach gets its name from the key role played by conditional probabilities--for example, P(Y/X). It can be shown that the C.P. approach affords us decompositon rules; it can also be shown that it affords us composition rules for a wide range of cases. In particular, it can handle causes which have both direct and indirect effects. Finally, it gives a quite plausible measure of one variable's causal influence on other.

Although the C.P. approach offers a rich and rather straight forward approach to many different kinds of causal systems, here we shall show how C.P. can resolve the three problems which lead to the downfall of path analysis. In doing so, we shall lay out some of the main principles and formulas of the C.P. approach. (Note: Here we shall assume the pairwise expectations of the outside causal variables are zero.)

Let's consider problem one. Suppose X is only a direct cause of Y, but that there are other causes of Y. (Suppose also, as above, that X

is a direct cause of Z and that Y is a direct cause of Z). In such a situation, the C.P. approach maintains that the magnitude of X's causal influence on Y is given by the conditional probability P(Y/X). In such a situation, the conditional probabilty is unaffected by the number or the strength of other causes of Y. Unlike the term $\rho_{xy}\frac{\sigma_y}{\sigma_x}$ which equals P(Y/X) - P(Y/notX), the conditional probability P(Y/X) remains unaffected by changes in the other causes of Y. Furthermore, the conditional probability can take on values from 0 to 1. It can take on the value 1 when X is a sufficient condition for Y or when X is a probabilistic cause (1.0) of Y. Thus, the conditional probability gives us the correct maximum value when X is sufficient condition for Y or when X is probabilistic cause (1.0) of Y. Notice also that when one is using the probabilistic or the INUS-conception of causation, the correlation ρ_{xy} can be zero even where X is a cause of Y. Conditional probability is not susceptible to this problem either. Principle: The causal link from X to Y should be assigned the magnitude P(Y/X). Thus, the C.P. approach easily resolves problem one.

Let's consider problems 2 and 3. This group of problems divides into two classes of problems. Problem 2 belongs to a class of problems involving pure-"or"-systems. Problem 3 belongs to a class of problems involving pure-"and"-systems. We shall clarify these terms as we proceed.

Up to this date, researchers have been most unclear about the distinction between pure-or-systems and pure-and-systems. Indeed, we can show that there is an entire continuum of systems which lie between the pure-or-system and the pure-and-system extremes. We'll discuss the continuum in another paper.

The pure-or-system is a causal system where X is a probabilistic cause (P) of Y and X is a probabilistic cause (1-P) of Z, but where X never leads to both Y and Z on a single occasion. Roughly, in this system X always leads to Y or Z, but never both. Modern physics can supply us with many examples of this type of setup. For example, when an excited electron decays from the higher energy levels to the lower levels, there will be nonzero probabilities that the electron will go to certain lower levels. Yet the electron can decay to only one of those lower levels; it can't go to two (or more) levels at once. Perhaps the radioactive decay of an atom into disjoint time intervals is another example. (Physicists claim the atomic structure of the substance causally explains such outcomes.)

The pure-and-system is a causal system where X is a probabilistic cause (P) of Y and where X is also an (independent) probabilistic cause (Q) of Z. Here X can lead to the occurrence of both Y and Z. This situation is analogus to a situation where two coins are tossed at the same time but where they don't interfere with each other's outcomes. The first coin, a blue-red coin, has the probability (P) of coming up blue while the second coin, a black-white coin, has the probability (Q) of coming up black. Here a tossing of the coins can lead to a black-blue outcome.

Problem 2 above involves a pure-or-system where X is a probabilistic cause (P = .5) of Y, and if X doesn't cause Y, then it causes Z (and vice versa). But X never leads to both Y and Z. Furthermore, Y is a probabilistic cause (R = 1.0) of Z. Hence, X is both a direct and indirect cause of Z. As we have argued above, the strength of each causal link in the system should receive the magnitude equal to the

conditional probability of the link's variables. Thus, the link from X to Y receives the value equal to P(Y/X) = P = .5; the link from Y to Z receives the value equal to P(Z/Y) = R = 1.0; and the link from X to Z receives the value P(Z/X) = Q = 1 - P = .5. In such a causal system, it can be shown that the overall causal influence of X on Z is given by the following formula:

$$[P(Z/X)]_{or} = P(Y/X) \bullet P(Z/Y) + P(Z/X) = P\ R + (1-P).$$

Given the numbers above, $[P(Z/X)]_{or} = (.5)(1) + (.5) = 1$. Notice that this formula for the overall causal influence of X on Z (in a pure-or-system) has the appropriate limiting values as the strengths of the causal links vary and that the link from Y to Z receives the maximum value 1 which is appropriate for a probabilistic cause (1.0). For example, if the strength of this link from Y to Z were reduced to 0, then the strength of the overall system reduces to the strength of the remaining link, 1-P. And if P were then to go to 1, both the link strength of X on Z and the overall strength would go to zero. As another example, if R were to remain fixed at 1 but P were to go to 1, then the strength of the link from X to Z would go to zero yet the overall strength of X on Z would remain at the maximum, 1.

The formula given above suggests the principle that the strength of a causal link be given by the conditional probability of the link's two variables and that the strength of a path with two links be given by the product of the strengths of the two links. Thus, the link from X to Z should receive the value P(Z/X) = (1-P) while the path from X to Y to Z should receive the value $P(Y/X) \circ P(Z/Y) = P \circ R$. Indeed, this result can be shown to generalize. Thus, the C.P. approach in this kind of situation gives the appropriate magnitudes of both the direct causal influence of X on Z and the indirect causal influence of X on Z.

For illustration, consider this system. Here, X is a probabilistic cause (P = .33) of Y; if X does not lead to Y, it leads to W (and vice versa). But X never leads to both Y and W. Furthermore, Y is a probabilistic cause (R = 1.0) of Z and W is a probabilistic cause (S = 1.0) of Z. Given the C.P. approach, the causal link from X to Y should equal P(Y/X) = P = .33; the link from X to W should equal P(W/X) =Q = 1 - P = .67; the link from Y to Z should equal P(Z/Y) = R = 1.0; and the link from W to Z should equal P(Z/W) = S = 1.0. In such a causal system, it can be shown that the overall causal influence of X on Z is given by the following formula:

$$[P(Z/X)]_{or} = P(Y/X)\cdot P(Z/Y) + P(W/X)\cdot P(Z/W) = P\cdot R + (1-P)\cdot S.$$

Given the numbers above, $[P(Z/X)]_{or} = (.33)\cdot 1 + (.67)\cdot 1 = 1$. Again, notice that this formula has the appropriate limiting values as the strengths of the causal links vary. For example, if R goes to zero and S goes to zero, then the overall causal strength goes to zero while the causal link from X to Y remains at P.

Problem 3 above involves a pure-and-system. X is a probabilistic cause (P = .5) of Y and X is also an (independent) probabilistic cause (Q = .5) of Z. Notice that Q need not equal (1-P) here. Furthermore, (Y and not X) is a probabilistic cause (R = 1.0) of Z. Given the C.P. approach, the causal link from X to Y should equal P(Y/X) = P = .5; the link from X to Z should equal P(Z/X) =Q = .5; and the link from Y to Z should equal P(Z/Y) = R = 1.0. In such a causal system, it can be shown that the overall causal influence of X on Z is given by the following formula:

$$[P(Z/X)]_a = P(Y/X) \bullet P(Z/Y) + P(Z/X) - P(Y/X)P(Z/Y)P(Z/X).$$

$$= P \bullet R + Q - P \bullet R \bullet Q.$$

Given the numbers above, $[P(Z/X)]_a = (.5)\bullet(1) + (.5) - (.5)(1)(5) = .75.$ Notice that this formula has the appropriate limiting values as the strengths of the causal links vary. For example, if P, R, and Q all become 1, then the overall strength is 1. The C.P. approach gives the correct answer even in cases where there are two (independent) paths to Z which are sufficient!

It is important that one compare the formula for overall causal strength in a pure-or-system with the formula for the overall strength in a pure-and-system. In the pure-or-system, the strengths of the two paths from X to Z simply add up to give the overall causal strength. In the pure-and-system, the strengths of the various paths from X to Z do not add up to give the overall strength. Although the formula for overall strength in a pure-and-system contains the sum of the strengths of the two paths, an quantity equal to the product of the strengths of the two paths must be substracted from the sum.

Thus, we have shown that the C.P. approach can avoid the problems and difficulties which beset the path analysis/structural equation approaches. And, it can give a plausible rendering of the strength of a causal link and a plausible account of the strengths of both direct causes and indirect causes (causal paths). And most importantly, the C.P. approach can produce simple formulas which express the magnitude of the overall causal influence of X on Z as a function of the magnitudes the causal links and paths which the theory or model asserts to exist from X to Z.

The principles, formulas, and other findings which we have presented here are but part of those we have developed for the conditional probability approach. Not only can the C.P. approach handle causal systems with more than two paths, the approach can also handle cases which fall into the continuum of cases which range from the pure-or-system to the pure-and-system. It can even handle systems which have both or-aspects and and-aspects. Furthermore, because of the limiting properties of the principles and the formulas, it can handle mixed causal systems in which there are several different kinds of causation (say, probabilistic causation and sufficient causation).

We trust we have been able in this short space to present a plausible and viable alternative approach to such problematic approaches as path analysis or structural equation methods. We believe that there are good reasons for holding that the conditional probability approach can provide an algebra for analyzing causal systems, an algebra with both decomposition and composition rules. (In the Supplement we present our general conditional probability theorems which hold for dichotomus and continuous cases under two sets of assumptions about the "outside causes" of the system.)

VIII. Conclusion

As as have shown, then, the two rules of causal inference underlying causal modeling are invalid for the five conceptions of causation. The widely held rule S.B. and rule C.S.B., which relate correlation, partial correlation and a conception of causation, are subject to counter-examples. Since structural equation models and/or path models share the most important assumptions (especially concerning the partials) with

Simon and Blalock, our results appear to generalize to those who try to make causal inferences. And difficulties for such approaches as path analysis and/or structural equation methods extend beyond those cases of causal systems in which a cause is a direct cause or indirect cause, but not both at once. The inability of the Simon-Blalock method to handle causal systems that hypothesize causes as having direct and indirect effects on the same variable strongly limits the usefulness of that method. All things being equal, theoretical and practical considerations favor any alternative that can adequately handle both kinds of systems. The alleged superiority of the other approaches rests, in part, on their claim to be able to handle such situations. We have shown, however, that even in situations where a cause has direct and indirect effects, the path analysis/structural equation methods yield incorrect solutions.

In sum, then, the major weakness in the expanding literature of causal inference is sufficiently clear. Though measurement and identification problems remain, it is the foundations for deriving causal inferences from "causal models" that require critical scrutiny. For without validity here, measurement issues surely become secondary. Of course, one alternative is simply to fall back on non-causal, predictive systems of causal modeling. Then, we need not worry about the validity of any newly proposed rules of causal inference. (Though the examination of predictive systems of causal modeling has been beyond our focus, we should remember that many of these approaches share the two inference rules with the causal interpretation version. Given the general kinds of problems encountered by the two rules relating correlation and partial correlation in the causal version, we should not be surprised to

find that merely predictive systems have problems here, too.) But, such a step would be to relinquish the goals of explanation and strong policy relevance for the social sciences. At present, this seems too large a step to take. For, despite our disagreements with Simon, Blalock, Asher, Duncan, Goldberger, and others, we share a similar vision of what the social sciences should be. There simply is no recourse from re-thinking the rules of causal inference from the foundations up.

As we have mentioned (pp. 33-34), we believe that a major problem with any correlation - based approach to causal analysis suffers from the same major difficulty. They all contain too much irrelevant information that masks and modulates the true nature of causal processes in the world. The overall result of using correlations, as we have seen throughout, is wrong answers. For this reason, we are convinced that Blalock is wrong when he states, "It is the regression coefficients which give us the laws of science" (1972, p. 51; emphasis in the original). Similarly, we do not think that path analysis or structural equations methods will yield the "laws of nature." No amount of tinkering can overcome the difficulties we have uncovered if the basic problem is inherent in the notion of correlation.

Furthermore, we have shown that in attempting causal analysis, it is vitally important for researchers to concern themselves with the substantive interpretation of cause. It cannot be brushed aside as an unimportant issue. The various results, as we have seen, depend upon our understanding of what a cause is. On the other hand, we must be cautious about any attempt to reduce the notion of cause to any single understanding. The likelihood is that causes in the world are of various sorts; some may be necessary for their effects, some sufficient,

some necessary and sufficient, some INUS conditions, and others--perhaps a majority--probabilistic. Attempts to reduce 'cause' to any one of these understandings have ended in failure--for in every case, counter-examples can be marshalled. Given this state of affairs, it is also likely that causal chains are composed of different types of causes; they are mixed chains.[8] This points to the requirement that our causal analysis methodology--if it is to be successful-must be able to handle such complexity. It is clear that our present methodology cannot.

Thus, we require an approach to causal analysis that is far more delicate and precise in detecting and dissecting causal networks. It must be an approach that (1) is non-correlational and (2) can handle causal systems containing mixed chains. We tentatively believe that the conditional probability approach that we have outlined can do precisely that. Our success in dealing with three complex cases that path analysis/structural equation approaches could not handle is an auspicious beginning. It remains for us to show that the conditional probability approach can handle a wider range of cases and different types of systems as well. Towards this end, we have already made progress. If laws be possible at all in the social sciences, the conditional probabilities may give us the laws of nature. If so, we shall have a methodology for causal analysis that is at once more simple, more subtle, and far more powerful than those we now possess. We shall have an algebra for analyzing causal relations.

(In the Supplement we present our general conditional probability theorems which cover both dichotomus and continuous cases under two sets of assumptions about the "outside causes" of a system which involves probabilistic causation.)

IX. Supplement (General Conditional Probability Theorems)

Is it possible to extend the conditional probability approach from dichotomus variables to continuous? We think yes. In this section, we will merely present (and not prove) the general theorems of the conditional probability approach.

IXA. Dichotomus Variables

First, let us briefly introduce notation used in the two dichotomus situations. We'll assume, 1) that X is probabilistic cause of Y (with probability (b + d)), 2) that Y is a probabilistic cause of Z (with probability β), and 3) that X is a probabilistic cause of Z (with probability (a + b)). Roughly, the system can be specified in the following way:

the probability of Y and Z, given only X, equals b;

the probability of Y and not Z, given only X, equals d;

the probability of not Y and Z, given only X, equals a;

the probability of not Y and not Z, given only X, equals c;

and the probability of Z, given only Y, equals β . We also assume that there are outside causes, V_2 of Y and V_3 of Z (where X equals V_1). We also have the following:

The probability of Y, given X and not V_2, equals b + d;

The probability of Y, given X and V_2, equals 1.

Similar remarks hold for the "additivity" of the other outside causes.

IXA. Dichotomus Variables; Pairwise Expectations of (V_1, V_2, V_3) zero

Given that the pairwise expectations of the outside causes are zero, the following can be shown.

$$\frac{EXY}{EX^2} = b + d \qquad \ldots(s1)$$

$$\frac{EXZ}{EX^2} = a + b + d\beta \qquad \ldots(s2)$$

$$\frac{EYZ}{EY^2} = \frac{P_2\beta + P_1 \cdot (b + d\beta)}{P_2 + (b+d) P_1} \qquad \ldots(s3)$$

(Here, P_1 is the probability of $V_1 = W$, P_2 is the probability of V_2, and P_3 is the probability of V_3.)

Special Case 1 ($a + b = 0$; simple chain)

$$\frac{EXY}{EX^2} = d \quad ; \quad \frac{EXZ}{EX^2} = d\beta \quad \text{and} \quad \frac{EYZ}{EY^2} = \beta .$$

Thus, $\frac{EXZ}{EX^2} = \left(\frac{EXY}{EX^2}\right)\left(\frac{EYZ}{EY^2}\right)$.

Special Case 2 ($b = 0$, $a = 1-d$; a pure-or-system)

$$\frac{EXY}{EX^2} = d; \quad \frac{EXZ}{EX^2} = a + \beta d; \quad \frac{EYZ}{EY^2} = \beta .$$

Special Case 3 ($(a + b)(b + d) = b$; a pure-and-system)

$$\frac{EXY}{EX^2} = b + d; \quad \frac{EXZ}{EX^2} = (a + b) + (b + d)\beta - (b+d)(a + b)\beta ;$$

$$\text{and } \frac{EYZ}{EY^2} = \beta + \frac{b(1-\beta)P_1}{P_2 + (b + d) P_1} .$$

IXA2. Dichotomus Variables; Pairwise Correlations of (V_1, V_2, V_3) zero

Given that the pairwise correlations of the outside causes are zero, the following can be shown:

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = (b + d) \cdot (1-P_2) \qquad \ldots(s4)$$

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = (1-P_3)[a + b + d\beta - (a + b + d)\beta \cdot P_2] \qquad \ldots(s5)$$

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = \{(1-P_3)(1-P_2)((1 - (a + b + d)P_1)[P_2\beta - \beta((a + b + d)P_1P_2 - dP_1)) + a \cdot P_1P_2 + bP_1] - aP_1[P_2 + dP_1 - (a + b + d)P_1P_2]\}) \text{ divided by } [P_2 + (b + d)P_1(1 - P_2)] \cdot [1 - P_2 - (b + d)P_1(1 - P_2)] \qquad \ldots(s6)$$

Special Case 1 (a + b = 0; simple chain)

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = d(1-P_2); \quad \rho_{xz}\frac{\sigma_z}{\sigma_x} = (1-P_3)(1-P_2)d\beta;$$

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = (1-P_3)\beta.$$

Thus, $\rho_{xz} = \rho_{xy} \cdot \rho_{yz}$

IXB. Continuous Variables

Let us briefly introduce notation used in the two continuous variable situations. Here the regular, continuous case; Pearson Product moment correlation is used. We'll assume that X is the probabilistic cause of Y's being AX with probability (b + d) and of Y's being zero with probability 1 - (b + d), that Y is the probabilistic cause of Z's being BY with probability β and of Z's being zero with probability 1 - β; and that X is a probabilistic cause of Z's being CX with probability (a + b) and of Z's being zero with probability 1 - (a + b). Roughly, the system can be specified in the following way:

The probability of Y's being $AX + V_2$, given X and V_2, equals b + d;
the probability of Y's being V_2, given X and V_2, equals 1 -(b + d);
the probability of Z's being $CX + V_3$, given X, V_2, and V_3, equals $(a + b)(1-\beta)$;
the probability of Z's being $CX + ABX + BV_2 + V_3$, given X, V_2, V_3, equals βa; and so forth.

As before, we assume that there are outside causes of V_2 of Y and V_3 of Z (where X equals V_1).

IXB. Continuous Variables; Pairwise Expectations of (V_1, V_2, V_3) zero

Given that the pairwise expectations of the outside causes, V_1,V_2 and V_3 are zero, the following can be shown:

$$\frac{EXY}{EX^2} = (b + d)A \qquad \ldots(s7)$$

$$\frac{EXZ}{EX^2} = (b + d)\beta AB + (a + b)C \qquad \ldots(s8)$$

$$\frac{EYZ}{EY^2} = \beta B + AC \cdot b\left(\frac{EX^2}{EY^2}\right) \qquad \ldots(s9)$$

Special Case 3 ((a + b)(b + d) = b; a pure-and-system)

$$\frac{EXY}{EX^2} = (b + d)A \; ; \quad \frac{EXZ}{EX^2} = (b + d)\cdot\beta\cdot AB + (a + b)C$$

and $\frac{EYZ}{EY^2} = \beta B + (a + b)C(b + d)A\left(\frac{EX^2}{EY^2}\right)$

IX. B.2 Continuous Variables; Pairwise Correlations of (V_1, V_2, V_3) zero

Given that the pairwise correlations of the outside causes, V_1, V_2 and V_3 are zero, the following can be shown:

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = (b + d)A \quad \ldots(s10)$$

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = (b + d)A \cdot \beta \cdot B + (a + b)C \quad \ldots(s11)$$

$$\rho_{yz}\frac{\sigma_z}{\sigma_y} = \beta B + \frac{(EX^2)ACb - (EX)^2AC(a + b)(b + d)}{\sigma_y^2} \quad \ldots(s12)$$

Special Case 1 ($a + b = 0$; simple chain)

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = dA; \quad \rho_{xz}\frac{\sigma_z}{\sigma_x} = dA\beta B;$$

and $\rho_{yz}\frac{\sigma_z}{\sigma_y} = \beta B.$

Thus, $\rho_{xz} = \rho_{xy}\rho_{yz}$.

Special Case 4 ($b = 1, \beta = 1$; a deterministic linear system)

$$\rho_{xy}\frac{\sigma_y}{\sigma_x} = A;$$

$$\rho_{xz}\frac{\sigma_z}{\sigma_x} = AB + C;$$

and $\rho_{yz}\frac{\sigma_z}{\sigma_y} = B + AC\left(\frac{\sigma_x}{\sigma_y}\right)^2.$

Thus, the conditional probability approach can handle causal systems which involve probabilistic causation. The approach can handle both dichotomus and continuous variables under the assumption that the pairwise expectations of the outside causes are zero and under the assumption that the pairwise correlations are zero. Notice also that the conditional probability approach gives the appropriate answers when the probabilities limit to deterministic sufficiency. (Special case 4.)

Notes

[1]The two critical steps in establishing these results are the following. (1) Kendall and Lazersfeld, Nagel and Suppes take rule C.S.B. and replace the requirement that $\rho_{xz,y}$ be zero with the requirement that $S_{xz,y}$ be zero, where $S_{xz,y}$ equals zero if and only if $P(XZ/Y) = P(X/Y) \cdot P(Z/Y)$ and $P(XZ/\text{not } Y) = P(X/\text{not } Y) \cdot P(Z/\text{not } Y)$. (2) It can then be shown that if $\rho_{xz,y}$ is not equal to zero, then $S_{xz,y}$ is not equal to zero. For a detailed discussion see Ellett and Ericson (1982).

[2]Elsewhere we compare and contrast in detail the Blalock kind of partial correlation approach with the approaches of H. Reichenbach (1956) and P. Suppes (1970). There, we establish that if Blalock's partial correlation is not zero then neither are the "partial correlations" used by Kendall and Lazarsfeld (1950), Nagel (1961), Reichenbach and Suppes. (See footnote 1.)

[3]Actually, it is far from difficult to estimate the regression coefficients of each equation, as Blalock argues (1967). So, this is hardly a problem for Blalock. (It is also claimed that, for example, path analysis can handle causes which have both direct and indirect effects, whereas the Simon-Blalock approach assume the cause is either a direct or an indirect cause but not both. When a cause is either a common cause or an intervening cause but not both, the approaches give the same result. In the first part of this essay we argue that in such situations they give the wrong result. In the latter part of this essay we argue that, for example, path analysis gives the wrong results in the other situations as well. See Section VI.)

[4]One of the important assumptions is that any "outside" factors also causing Y, or causing Z, are such that the pairwise mathematical expectations between these "outside" factors are zero. In sections IV and V, our examples satisfy this assumption because we have constructed cases where such (dichotomus) variables or factors never co-occur. In the supplement IX, we present our results for systems with dichotomus variables and for systems with continuous variables under the assumption that the expectations are zero and under the assumption that the correlations are zero.

[5]In a more technical paper (in preparation) we discuss the problems concerning defensibility of the key assumptions. For example, the assumption that the error terms be uncorrelated with the predictors or with each other.

[6]To be more precise, rule C.S.B. is invalid because there are no cases where Y is a common necessary cause of X and Z, given the assumption that the pairwise expectations of the outside causes are zero. The tables we presented here have assumed that the pairwise correlations of the outside causes are zero. Thus, they show the rule C.S.B. is invalid under both sets of assumptions.

[7]There is a major problem with Simon's analysis. Though he recognizes that causes are asymmetric with their effects (if X causes Y, it cannot be that not--Y causes not-X), Simon's use of the predicate calculus containing "material implication" ($\supset$) commits him to saying that not-Y causes not-X. For $X \supset Y$ is equivalent to $notY \supset notX$ via the rule of logical inference called modus tollens. See also Simon's use of the "truth tables" which illustrates the commitment (1957, pp. 58-59).

8 Consider the following mixed case involving probabilistic and sufficient causes. It might be imagined as something like a pinball machine. Here X and Y are mutually exclusive, exhaustive, and equiprobable results from W (the probabilistic cause), whereas X or Y is sufficient for Z (Z = X or Y), and where W occurs with twice the frequency as not-W. For then, $\rho_{WZ} = 1$ but $\rho_{XZ} = 1/2$.

References

Asher, H. B. Causal Modeling. Beverly Hills, CA: Sage Publications, 1976.

Bentler, P. M. "Multivariate Analysis with Latent Variables: Causal Modeling." Annual Review of Psychology, 31, 1980.

Blalock, H. M. "Causal Inferences, Closed Populations, and Measures of Association." American Political Science Review, 61, 1967. Also in Blalock (Ed.) (1971) below.

Blalock, H. M., Causal Inferences in Nonexperimental Research. New York: W.W. Norton, 1972.

Blalock, H. M., (Ed.). Causal Models in the Social Sciences. Chicago: Aldine Publ. Col, 1971.

Campbell, D. T. & Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNalley, 1966.

Chomsky, N. Language and Mind. New York: Harcourt Brace, 1968.

Chronbach, L. J. & Meehl, P. E. "Construct Validity in Psychological Tests." Psychological Bulletin, 52, 4, July 1955.

Cook, T. D. & Campbell, D. T. Quasi-Experimentation. Chicago: Rand-McNalley, 1979.

Dummett, M. & Flew, A. "Can an Effect Precede its Cause?" Proceedings of the Aristotelian Society, Supp. Volume XXXVIII, 1954.

Duncan, O. D. "Path Analysis: Sociological Examples." American Journal of Sociology, 72, 1966. Also in Blalock (Ed.) (1971).

Duncan, O. D. Introduction to Structural Equation Models, New York: Academic, 1975.

Ellett, F. S., Jr. & Ericson, D. P. "Correlation, Partial Correlation and Causation." Unpublished manuscript 1982.

Giere, R. N. Understanding Scientific Reasoning. New York: Holt, Rinehart, and Winston, 1979.

Goldberger, A. S. "Structural Equation Models: An Overview." In Goldberger, A. S. and Duncan, O. D. (EDS.). Structural Equation Models in the Social Sciences. New York: Seminar Press, 1973.

Good, I. J. "A Causal Calculus I-II." British Journal for the Philosophy of Science, XI, 44, 1961 and XII, 45, 1962.

Guttman, L. "What Is Not What in Statistics." The Statistician, 26, 1977.

Heise, D. Causal Analysis New York: Wiley, 1975.

Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart, and Winston, 1963.

Hempel, C. The Philosophy of Natural Science. Englewood Cliffs, N.J.: Prentice-Hall, 1967.

Hicks, S. J. Causality in Economics. New York: Basic Books, 1979.

Kendall, P. and Lazersfeld, P. F. "Problems of Survey Analysis," in R. K. Merton, and P. F. Lazersfeld (eds). Continuities in Social Research: Studies in the Scope and Methods of "The American Soldier." Glencoe, Ill., 1950, 133-196.

Kenny, D. A. Correlation and Causation. New York: John Wiley and Sons, 1979.

Lakatos, I. "Falsification and the Methodology of Scientific Research Programmes." In Lakatos, I. and A. Musgrave (Eds.), Criticism and the Growth of Knowledge. Cambridge: Cambridge University Press, 1970.

Mackie, J. L. The Cement of the Universe. Oxford: Oxford University Press, 1974.

Nagel, E. The Structure of Science. New York: Harcourt Brace, 1961.

Popper, S. K. Objective Knowledge. Oxford: Oxford University Press, 1972.

Reichenbach, H. The Direction of Time. Berkeley: U. of California Press, 1956.

Rudner, R. The Philosophy of Social Science. Englewood, N.J.: Prentice-Hall, 1967.

Salmon, W. C. "Probabilistic Causality." Pacific Philosophical Quarterly, 61, 112, Jan.-April 1980.

Salmon, W. C. with contributinons by R.C. Jeffrey and J. G. Greeno, Statistical Explanation and Statistical Relevance. Pittsburgh: University of Pittsburgh Press, 1971.

Sayre, K. M. "Statistical Models of Causal Relations." Philosophy of Science, 44, 1977.

Simon, H. A. Models of Man. New York: John Wiley and Sons, 1957.

Skinner, B. F. Science and Human Behavior. New York: Free Press, 1953.

Steiger, J. H. Review Section of Psychometrika, 45, 3, September, 1980.

Suppes, P. A Probabilistic Theory of Causation. Amsterdam: North-Holland, 1970.

Wright, S. "Correlation and Causation." Journal of Agricultural Research, 20, 1921.

Wright, S. "Path Coefficients and Path Regressions: Alternative or Complementary Concepts?" In Blalock (Ed.), 1971 above.

Appendix to

"Towards an Algebra for Analyzing Causal Relations"

F. S. Ellett, Jr. & D. P. Ericson

Graduate School of Education

University of Calfornia

405 Hilgard Avenue

Los Angeles, California

90024

March 1982

The address given at AERA's Annual Meeting in New York City at Session 58.03, Causal Modeling.

An Overview of "Towards an Algebra for Analyzing Causal Relations"

F.S. Ellett, Jr. & D.P. Ericson

UCLA

Educators of all types make the following kind of assertions: "A particular teaching method causes the students to learn." But what can it mean to say that one kind of event (or thing) causes another kind of event? This is a philosophical question, for it is concerned with the conceptual analysis of 'causation.' Suppose now that we have an (adequate) analysis of 'causation.' What rules of inference can be validly used to test hypotheses about the causal relations between events (or variables)? This is also a philosophical question, for it is concerned with the logic of causal inference--with establishing that certain rules are sound and valid.

In this essay we address ourselves to these two general but important philosophical questions : (1) What is a plausible analysis of 'causation'? and (2) what rules of causal inference are valid? As we show in the paper, an answer to the question about valid rules of inference depends on one's analysis of causation. Furthermore, we consider several of the widely used rules of causal inference. We conclude that in general these rules are invalid. To be more specific, we consider the partial-correlations rule used by Simon and Blalock, the "partial-correlation" rule used by Kendall and Lazarsfeld, E. Nagel, H. Reichenbach, and P. Suppes, and the rules associated with path analysis and structural equation methods. Again, we conclude that in general these rules are invalid.

In the time that remains we'll present more of the details about the methods and conclusions of our paper. In doing so we'll present an overview and outline of the paper.

As we've already noted, educators often make assertions about the causal relations among events or variables. Indeed, we believe there are compelling arguments for holding that scientific analysis and policy formation must be concerned with causation. In an early section of the paper we present some of the major reasons for holding these views about the importance of causation.

The next part of the paper investigates the relationships among correlation, partial correlation, and various analytical conceptions of causation for dichotomous variables. Here we consider the Blalock-Simon partial-correlation rule and the partial-correlation rule advanced by Kendall and Lazarsfeld, Nagel, and Suppes. Roughly, the partial correlation rules assert that for intervening variables or common causes certain partial correlations must go to zero. In other words, the partial correlation's being zero (or not) provides a test of the causal hypothesis.

In this part of the paper we also present five philosophical "theories" of the analysis of causation. We refer each of these purported analysis as a conception of causation. The five conceptions of causation are:

1) a cause C is a necessary condition for E.
2) a cause is a sufficient condition
3) a cause is a necessary and sufficient condition
4) a cause is an INUS condition

and 5) a cause is a probabilistic cause.

The first three of these should be familiar to you. Perhaps the last two are not so familiar. So let us briefly explain them. An event C is an INUS cause of E if the event C is an insufficient but non-redundant part of a set of conditions which is not necessary, but which is sufficient for E. Hence, the label I.N.U.S. Here's an example. Striking a match is an INUS cause of the match's flame because the striking also requires oxygen's being present and the match's being dry to have its effect. Yet, striking the match is not the only way to get it to light. Thus, striking the match is an INUS-cause.

Notice that it is quite plausible to hold that teaching is an INUS-cause of student's learning. Clearly teaching isn't sufficient and in most cases the child must have the requisite background knowledge and be trying to learn.

Thus, if C is an INUS-cause of E, than C is part of a set of conditions which is sufficient for E. In the paper we show that an event C can be an INUS-cause of E and the correlation between C and E can be negative or zero!

The fifth conception of causation is called probabilistic causation. One of the earliest writers to advance such a conception was P. Suppes. In many ways, the "mathematics" of a probabilistic causation is similar to those of an INUS-causation. If C is a probabilistic cause of E, however, it does not follow that C is a part of a set of conditions which is sufficient for E. Probabilistic causation applies to the case where there is genuine indeterminism. Here, an example will help clarify the term. Suppose an atom has an electron at one of its higher energy levels. Then there are finite probabilities that the electron will "decay" or fall to various other levels, but, as some physicists

maintain, there is no set of conditions which determines the level to which the electron will decay. We are not saying, of course, that there are such cases of probabilistic causation in physics or education. We are only saying it is a perfectly intelligible and legitimate conception of causation.

At any rate, in the paper we show that C can be a probabilistic cause of E and the correlation between C and E can be negative or even zero. You will recall that when one of the first three conceptions of causation is used, the correlation must be positive. So INUS-causation and probabilistic causation behave differently from the first three conceptions of casuation.

Our investigations of the relationships among correlation, partial correlations, and conceptions of causation conclude that for each of the conceptions of causation the partial-correlation rules are invalid.

We have spent time here today elucidating the conceptions of INUS-causation and probabilistic causation for two reasons. First, we believe that educators are not adequately aware of them, but they should be. Second, the next parts of the paper primarily employ the concept of probabilistic cause, although our results also hold for INUS-causation.

The next major section of the paper explicates the basic principles of path analysis and structual equation methods. Indeed, as many people have said, path analysis and structural equation methods have, in many ways, replaced the earlier approaches associated with Simon and Blalock. We show however, that when the variables are dichotomus and where probabilisitc causation is involved, path analysis and structural equation methods give the wrong answers. In other words, path analysis and structural equation methods present rules of causal inference which are invalid.

At this point it may be helpful to distinguish among three different kinds of probability-statements which can be about a causal system which involves probabilistic causation. First, a causal system (or model) is partly constituted by certain variables. Let us call them X, Y, and Z. Now the variables X, Y, and Z will be distributed in various ways in the population. Such distributions give us the first kind of probability statements. Of course, we have only samples of the general population. This gives rise to those probability statements involving measurement error and estimation, and so on. This is the second kind of probability statement. We trust these two kinds of probability statements are quite familiar. But, notice that so far we have said nothing about the nature of the causal relationships among variables X, Y and Z. Perhaps X causes Y, and Y causes Z and also X directly causes Z. Here X is a direct and indirect cause of Z. This is, of course, a further specification of the causal relations among the variables X, Y, and Z. Indeed, this is the kind of basic system which path analysis and structural equation methods claim they give a better analysis than the Simon-Blalock approaches.

Yet, a further kind of specification is still required. For as we have shown there are five different conceptions of causation. It is our belief that path analysis and structural equation methods must conceive of the causal relations between, say, X and Y to be causation as sufficiency. But, in the paper, we show that one can conceive of the causal relations among X, Y and Z as probabilistic-causation. This is what gives rise to the third kind of probability statement concerning the system under investigation.

Furthermore, and this is the important point, when one conceives of the causal relations among the variables as probabilistic-causation,

then the rules and principles associated with path analysis and structural equation methods are invalid.

At this point the attentive listener is likely to feel uneasy. Afterall, the listener might think; didn't Ellett and Ericson say they were dealing only with dichotomus variables? What about systems with continuous variables? And what is the general importance of these findings? Finally, perhaps Ellett and Ericson have done a typical philosophical job in exposing the problems and errors in current methods, but can they offer any constructive views about what should be done?

We have answers to all three of the attentive listener's questions. First, there are indeed pecularities with causal systems with dichotomus variables. But, we have extended our results to causal systems with continuous variables. For causal systems with continuous variables we have found that path analysis and structural equation methods, when probabilistic causation is involved, still give the wrong answers. In such cases, their rules of causal inference are invalid.

Second, the general importance of our findings is that rules of causal inference need not be restricted to cause as sufficiency. They can be extended to systems where there is INUS-causation or probabilistic-causation. And, it is likely that such conceptions of causation "will fit" educational phenomena. Recall our earlier remarks that teaching may well be an INUS-cause.

Thirdly, and finally, it is true that we have been critical of many of the current writers who are concerned with methods of causal inference. We do believe that current methods are severely limited. But, we do share their commitment to trying to find those methods and rules of causal hypothesis testing which can yield valid inferences.

And, toward the constructive end of finding rules and methods for analyzing the causal relations, in the paper we have presented methods and equations for causal systems which involve probabilistic causation. We believe these methods and equations articulate valid rules of inference, valid rules of inference for both systems with dichotomous variables and for systems with continuous variables.

We believe a whole new dimension of causal analysis has been opened. We hope you'll explore this dimension with us.