74-17,220

YAP, Kim Onn, 1941-
A STUDY OF THE EFFICIENCY OF CAUSAL ANALYSIS CONDUCTED ON PANEL DATA.

University of Hawaii, Ph.D., 1973
Education, psychology

University Microfilms, A XEROX Company, Ann Arbor, Michigan

A STUDY OF THE EFFICIENCY OF CAUSAL ANALYSIS

CONDUCTED ON PANEL DATA

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN EDUCATIONAL PSYCHOLOGY

DECEMBER 1973

By

Kim Onn Yap

Dissertation Committee:

Peter Dunn-Rankin, Chairman
Dorothy Adkins
Frederick Bail
Harry Ball
Ian Reid

ACKNOWLEDGMENTS

The valuable assistance and encouragement that I received from all those who have taken an interest in the present study are gratefully acknowledged. In particular, I would like to thank Dr. Charles Portwood of the Curriculum Research and Development Group for his very useful comments and criticisms on earlier drafts of this dissertation.

I am also indebted to the Department of Educational Psychology for providing the unusually large amount of computer time needed to complete the study. A note of appreciation is extended to the Computing Center staff for giving me full access to the Center's facilities and professional consultation.

ABSTRACT

The primary purpose of the present study is to investigate the validity of models for causal analysis in the behavioral sciences. The first phase of the investigation provides a numerical definition of cause and on the basis of this definition formulates a computer simulation model to generate panel data with pre-specified causal characteristics. Three major causal parameters are built into the data: (1) the direction of causal influence, (2) the amount of causal influence, and (3) data reliability. These parameters are systematically manipulated to determine their effects on causal estimation. A total of 110 data sets are created. Each data set differs from the other with respect to one or more of the causal parameters.

Seven causal models are applied to the created data sets to determine their relative efficiency. The models are: (1) cross-lagged correlation model, (2) part correlation model, (3) econometric model, (4) variance components model, (5) frequencies-of-shift-across-median model, (6) frequencies-of-change-in-product-moment model, and (7) modified-frequencies-of-change-in-product-moment model.

The results suggest that the efficiency of the causal models is relative to the nature of the data. None of the models appear to be efficient for all data sets created under the simulation model. In general, when the time-one measures of the two variables involved are uncorrelated, causal estimation is likely to be valid. When the time-one measures are correlated, causal estimation appears to have low validity. Furthermore, it appears to be easier to detect congruent causal influence than to detect incongruent causal influence. When

both congruent and incongruent causal influences are present, causal estimation appears to be extremely difficult.

The above findings imply that causal analysis may best be conducted when the researcher has some prior knowledge about the data. They also suggest that the efficiency of causal analysis as reported in the literature is probably highly exaggerated. If the causal models are used as procedures for classifying individuals on the basis of causal influence, the results are likely to be erroneous. A summary is provided on the strengths and weaknesses of the seven causal models relative to the nature of the data to which the models are applied. This will hopefully serve as a guideline for future researchers who may need to use the models.

The second phase of the study consists of two causal analyses conducted on data obtained from real-life situations. In the first analysis, the causal relations between vocabulary and comprehension abilities are investigated. In the second analysis, the causal relations among height, weight, and strength are studied. The results of these analyses indicate that some knowledge about a given set of real-life data can be determined or assumed prior to causal analysis and that reasonable applications of causal models can be made to the data to infer causal relations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

PART ONE

A SIMULATIVE EVALUATION OF SEVEN CAUSAL MODELS

CHAPTER I

INTRODUCTION TO PART I

Scientific investigations have generally been conducted on the basis of (i) experimental manipulation and (ii) naturalistic observation. The physical scientist has been able to introduce experimental control in most of his domains of investigation. The social or behavioral scientist, for pragmatic reasons, has had to resort to naturalistic observation as a means of collecting data much more than his physical science counterpart. A major difference between experimental and quasi-experimental research appears to lie in the fact that in an experimental situation the researcher can manipulate the independent variable and its effects on the dependent variable can usually be determined on an analytical basis. In the jargon of the traditional analysis of variance model, various sources of variance can be identified and properly attributed. In quasi-experimental settings, however, variance attribution becomes much more difficult. In most cases, the distinction between the dependent and independent variables is arbitrary. Thus, given pupil attitudes and teacher attitudes (Yee, 1966) as two variables, either one may be considered to be the dependent or independent variable. Ordinarily, one of the two variables is arbitrarily assigned the role of the independent variable.

A confounding factor is the presence (typically in large amounts) of error variance in naturalistic observations. Few, if any, correlation coefficients, for instance, take the value of unity or zero. Error variance can be assumed to be always present. Yet, the specification of error variance in data gathered in quasi-experimental settings is usually not possible. It follows that the effects of error on inferences made on the basis of such data are also unknown.

In the absence of more appropriate procedures, the product-moment correlation coefficient has often been used as a means of analyzing data gathered in quasi-experimental settings. A correlation coefficient does not in any way dispel the ambiguity with respect to variance attribution and often adds to the uncertainty of the "explained" variance. It appears that it is partly because of this limitation of the correlation coefficient that many people have tended to confine their data interpretation to a level at which only association of the two variables involved is inferred. Students of inferential statistics are often warned in the first introductory course that a correlation coefficient implies only a "functional" relationship between two variables.

Yet, it is clear that a major portion of the data gathered by the behavioral scientist belongs to the quasi-experimental category. If behavioral sciences are to make further advances, it would appear evident that efforts should be made to push data interpretation beyond the purely associative level. Ambiguity with respect to the attribution of variance will have to be reduced or completely removed. Analytical techniques will have to be developed to identify independent variables which do not lend themselves to experimental manipulation.

A start has already been made in this direction (see Campbell, 1963; Campbell and Stanley, 1963; Yee, 1966; Blalock, 1969; Portwood, 1972). A number of techniques have been suggested as appropriate tools for making scientific inferences on the basis of quasi-experimental data beyond the associative level. In terms of variance attribution, the aim of these techniques is to identify the independent variable in quasi-experimental settings, and once this is done, to properly specify and assign variance. For lack of a better term, such procedures have been described as quasi-experimental designs or "causal" models.

A word should be said about the term "causal." Historically, the concept of cause often has been introduced as an "explanatory" construct. Hume (1880) admittedly denied the certainty of causal knowledge on the basis that (i) any conclusion derived from inductive argument is in fact not certain; and (ii) one can never be certain that the laws of the universe will not change in the future. Yet, what is often overlooked by Humean philosophers is that Hume never denied the usefulness of the concept of causality (see Portwood, 1972). Other writers (e.g., Wold, 1954, 1956; Blalock, 1969) have advocated the use of "causal models" in theory building. Within the context of the present study, a causal model is defined as a procedure designed to identify the independent variable in a quasi-experimental setting where both dependent and independent variables are not amenable to direct experimental manipulation.

Some of the causal models developed in the behavioral sciences are direct "borrowings" from traditional statistical procedures originally designed to determine the "functional" relations among variables, while others are developed for the specific purpose of disentangling "causal" relationships. When the former are used, causal interpretation becomes a matter of professional orientation or personal preferences. The empirical and theoretical difference between "functional" and "causal" dependence appears to be nothing more than the difference between scientific description and explanation. In other words, it is, more than anything else, a matter of semantics.

More recent statistical techniques developed to disentangle "causal" relationships are not spared the equivocality involved in data interpretation. The main problem, however, appears to lie with the degree of specificity with which the models are formulated. Most models are implicit rather than explicit with respect to assumptions which need to be made to uphold the validity of the models. Thus, different causal procedures may provide different results with questionable validity.

The purpose of the present study is to provide a numerical definition of cause andwithin the context of this definition to investigate the validity of the causal models which have been developed in the behavioral sciences. In the course of this investigation, the strengths and weaknesses of these models are made explicit so that future researchers may be aware of their limitations and efficiency.

CHAPTER II

REVIEW OF CAUSAL MODELS

Causal models developed in the social and behavioral sciences can be classified under two rubrics: models which deal with data on concurrent variables and models which deal with data of a time-series nature. Under the first classification are the Simon-Blalock approach and the econometric models, although the latter is often used to analyze time-series data. Under the second classification are the sixteenfold table model, the cross-lagged correlation model, the frequencies-of-shift-across-median model, and the frequencies-of-change-in-product-moment model. As the present study deals mainly with methodological issues, the review is confined to a discussion of the various models as techniques for causal analysis. Substantive issues are generally avoided. The main objective of the review is to present the technical aspects of the causal models in some detail. In the following sections, causal models which deal with data on concurrent variables will be reviewed first. This is followed by a presentation of models which deal with time-series data.

Concurrent variable models

One of the first explications on causal analysis was made by Simon (1954). Faced with the ambiguity of the correlation coefficient, he attempted to make a distinction between "true" correlation and "spurious" correlation. The author stated that while correlation in general might be no proof of causation, "true" correlation (as opposed to spurious correlation) did constitute such proof. To find out whether the correlation between X and Y is true or spurious, one would introduce a third variable Z that may account for the correlation. If the partial correlation between X and Y with Z held constant is close to zero while the zero-order correlation between X and Y is not, one would conclude that either (a) Z is an intervening variable or (b) the correlation between X and Y is spurious. Whether (a) or (b) is the case will depend on some "common sense" assumptions which can be made about the empirical data. Simon (1954) cited two examples to illustrate the application of this method.

Example one dealt with three variables: the average amount of candy consumed per month by each subject, the average age of subjects, and the percentage of subjects who were married. A high negative correlation was found between marital status and the amount of candy consumed. There was also a high negative correlation between marital status and age. When age was held constant, the partial correlation between marital status and candy consumption was nearly zero. This finding suggested that either age was an intervening variable between marital status and candy consumption or the correlation between marital status and candy consumption was spurious, being a joint effect of age. The author argued that common sense tells us that the latter explanation was correct.

Example two dealt with the following three variables: the percentage of female employees who were married, the average number of absences per week per employee, and the average number of hours of housework performed per week per employee. A high positive correlation was found between marriage and absenteeism. However, when the amount of housework was held constant the partial correlation between marriage and absenteeism was virtually zero. The author argued that common sense tells us that the average number of hours of housework was an intervening variable (i.e., marriage results in a higher average amount of housework, which in turn results in more absenteeism.)

It is clear that, using Simon's method, the researcher is able to make a causal interpretation only if he is prepared to assume, a priori, that certain other causal relations do not hold among the variables. In example one, he assumes that the age of the person does not depend on either his candy consumption or his marital status. In example two, he assumes that marital status is not causally dependent on either the amount of housework or absenteeism. In cases where such assumptions cannot be made, utility of the technique becomes limited.

The work of Simon (1954) led to more studies on spurious correlation, most of which tended to cast more shadow than light on the subject matter. Linn and Werts (1969) for instance, studied four different methods of removing spurious association between variables: (1) the partial correlation between X and Y with the causally prior variables controlled; (2) the part correlation of X and Y with the causally prior variables partialled out of X; (3) the part correlation of X and Y with the causally prior variables partialled out of Y; (4) the standardized partial regression weight of Y on X computed from the regression equation in which the causally prior variables and X are the independent variables. The authors showed that the four procedures may yield widely divergent estimates of the magnitude of the influence of X on Y.

It will be noted that the four procedures can be described in terms of the sequence of regression equations used to obtain the respective coefficients. The regression steps for partial correlation are: (1) regress X on the common antecedent (or causally prior) variables to obtain X' or the X residuals; (2) regress Y on the common antecedent variables to obtain Y', or the Y residuals; (3) regress Y' on X'. The standardized regression weight for X' thus obtained is the part correlation measuring the influence of X' on Y. This model implies that all other causes of X are uncorrelated with the common antecedent variables and it is the hypothetical variable X' which influences Y. Similar regression steps are involved in computing the part correlation of X and Y with the influence of the common antecedent variables removed from Y. In addition, it will be noted that if X is influenced by the common antecedent variables and X influences Y', then the common antecedent variables will be correlated with Y'. This will create a logical contradiction in the model since Y' is assumed to be uncorrelated with the common antecedent variables. To avoid this contradiction we must assume that the antecedent variables do not influence X. The standardized partial regression model involves only the regression of Y on X and the common antecedent variables. In this model, X and the antecedent variables may be correlated, but it is assumed that all other causes of Y are uncorrelated with both X and the common antecedent variables.

It can be shown that all four methods are in fact simply variations of the unstandardized partial regression coefficient. The relationship between the standardized and unstandardized regression coefficients is given by the following equation:

$$B^*_{yx.w} = B_{yx.w} \frac{\sigma_x}{\sigma_y}$$

Where B* is the standardized coefficient, W is a common antecedent variable, and B is the unstandardized coefficient. The above equation suggests that the standardized measure is highly subject to fluctuation due to changes in the variances of X and Y from one sample to another.

For this reason, the generalizability of the four procedures discussed earlier should be limited to very specifically defined populations. Under circumstances where the range of the sample is restricted on one or more of the variables, these methods can lead to seriously biased estimates. Unstandardized regression coefficients, on the other hand, have the advantage of not being affected by changes in sample heterogeneity. It follows that they are also more stable than the correlation coefficient since the latter is equivalent to the standardized regression weight.

In addition to the above limitations, Richards (1966) showed that part or partial correlation might severely underestimate themagnitude of the true relationship between two variables, Tukey (1954) suggested the use of partial regression as a more satisfactory procedure. Werts and Watley (1968) provided hypothetical models which showed that a control that reduced the magnitude of the partial correlation did not affect the expected value of the corresponding regression coefficient. It was stated that the regression coefficient would generally provide a more interpretable estimate of the relationship between two variables than would part or partial correlations.

Another prominent proponent of causal analysis is Blalock (1960, 1961_a, 1961_b, 1962_a, 1962_b, 1965, 1967, 1969) who argued that symmetrical correlation coefficients can be used to make asymmetrical assertions about causal relationships because correlation coefficients possess magnitude, and the value of each correlation is assumed to be determined in part by certain specific variables rather than others (see Blalock 1962_b; Polk, 1962). Some variables might be only indirectly related to each other through the operation of other variables. In such cases, the correlation between any two indirectly related variables should be weaker than the correlations between either of these variables and some intermediate variable. The paradox that symmetrical coefficients are used to make inferences as to the direction of causality (which implies asymmetry) is real only if one looks at single pairs of variables.

Blalock (1962a) extended the Simon model to cover the multiple-variable case. Using a recursive set of equations, the author showed that if some of the regression coefficients could be assumed to be zero, certain restrictions would have to be imposed on the data if the equations were to remain mutually consistent. Specifically, for each of the regression coefficients set to zero (i.e., it is assumed, a priori, that there is no direct causal link between the two variables concerned) one would impose the condition that the corresponding partial correlation should also be zero. The predictions thus obtained could then be tested on the empirical data.

In one of his earlier applications of causal analysis in sociology, Blalock (1960) used a set of data on North American Indian tribes. The variables included division of labor, residence, land tenure, and descent. The hypothesis being tested was that a matridominant division of labor (W) should be followed by matrilocal residence (X), then by matricentered land tenure (Y), and finally by a matrilineal system of descent (Z).

Blalock (1960) noted that among the four variables the largest correlations should occur between adjacent variables (i.e., the correlation between (W) and (Y) should be smaller than those between W and X, and X and Y.) and the smallest correlations should occur between variables furthest removed from each other in the causal chain (i.e., W and Z). The author showed that one can predict exactly how large each correlation should be in relation to others. Specifically, r_{xz} should equal $r_{xy}r_{yz}$, r_{wy} should equal $r_{wx}r_{xy}$, and r_{wz} should equal $r_{wx}r_{xy}r_{yz}$. Had the researcher constructed an alternative model by adding a direct link between division of labor (W) and land tenure (Y), he would have expected r_{xz} to equal $r_{xy}r_{yz}$ and r_{wz} to equal $r_{wy}r_{yz}$. The results of this particular study on North American Indian Tribes showed that the second model had a higher degree of fit for the data.

The above approach is of course not without its limitations. First, it is important to realize that this approach does not enable us to actually prove the correctness of any causal inferences. One set of inferences may be more adequate than another in terms of goodness of fit, but their validity is never empirically established. Second, there will usually be several sets of inferences which predict exactly the same empirical results. For instance, the following models postulate reverse causal chains: $W \rightarrow X \rightarrow Y \rightarrow Z$ and $Z \rightarrow Y \rightarrow X \rightarrow W$. Yet they predict identical intercorrelations among the variables. The researcher will have to depend on theoretical reasoning, knowledge of time sequences, or even common sense to choose between these two causal models. Third, it should be cautioned that random measurement errors can lead to faulty causal inferences. Given a causal model where Z is the intervening variable between X and Y, a control for Z (if measured perfectly) will reduce the association between X and Y to zero, barring sampling errors. However, if Z is measured with error, one will be controlling only for Z', the measured value of Z, and not Z itself. Thus, in the case of measurement error, a control for Z' will reduce but not eliminate the correlation between X and Y. The degree of reduction in the correlation will depend on the actual amount of measurement error involved. Under these circumstances, a number of alternative hypotheses about the original causal model become plausible: (i) Z has, in fact, been perfectly measured but there is also a direct causal relationship between X and Y; (ii) the relationship between X and Y is indeed spurious but an additional variable W also needs to be controlled; (iii) Z is part of a developmental sequence in which Z causes X, which in turn causes Y. Thus, whenever there is random measurement error in one or more of the variables in a causal system, the causal relationships become ambiguous. This gives rise to a larger number of rival hypotheses which cannot be ruled out on the basis of empirical data.

Data analysis conducted in the realm of econometrics appears to be essentially causal in nature. The econometric models (Blalock, 1969; Wonnacott and Wonnacott, 1970) were developed from regression analysis and are sometimes simply called the regression models (Darlington, 1968). While they do not explicitly use the causal terminology, these models do infer causal relationships between variables, a notable example being the causal influence between supply and demand. The regression coefficients are taken as indices of the strength of causal links between exogenous (independent) and endogenous (dependent) variables.

The central issues in econometrics have to do with the stability of the causal system and the identification of the regression equations. The notion of stability is directly related to the magnitude of change which occurs in the endogenous variable over time. If the magnitude of change increases at successive time periods, the causal system will not be able to stabilize. It will, as a matter of fact, explode. If, on the other hand, the magnitude of change decreases over time, then the system will sooner or later reach a point of stable equilibrium.

For simplicity, consider the following equation as a causal system:

$$X_2 = A + \hat{B}X_1 + e,$$

where X_1 is a measure of variable X at time-one, X_2 is a measure of the same variable at time-two, and A is a constant. In this equation, X at time-two is expressed as a function of itself at time-one. It is clear that the stability of the causal system now hinges on the magnitude of $\hat{B}$--provided, of course, that one could ignore the error term. If $\hat{B}$ is positive, then an increase in X will bring about further increases in successive time intervals. If the numerical value of $\hat{B}$ is greater than unity, then an initial change of one unit will bring about a greater than unit change in the next time interval which will, in turn, be followed by successively greater changes. The successive increases in the magnitude of change will eventually lead to the explosion of the system. On the other hand, if the numerical value of $\hat{B}$ is less than unity, then the successive changes in X will become smaller and smaller, bringing the system to a position of equilibrium.

In real life situations such explosive causal systems probably do not exist. At any rate, inasmuch as the present study deals only with the identification of the source and direction of causal influence, the stability of the causal system has little relevance here. For convenience, one may simply assume that the system is approaching the point of stable equilibrium. In Yee's (1966) study, for instance, one may assume that the attitude change--caused by either teachers or pupils--becomes smaller and smaller at successive time intervals within the teacher-pupil interaction period.

The identification problem has a direct bearing on how the estimating equations are formulated. Specifically, if the equations are underidentified, one will not be able to obtain unique estimates of parameters. While the fine points of the identification problem will not be discussed here, an intuitive understanding of the problem can be gained by means of an analogy. Given a set of equations where the number of unknowns exceeds the number of equations, one ordinarily cannot obtain unique solutions for the variables involved. Such a set of equations may be as follows:

$$X = Y + 2Z$$

$$Y = 3X$$

No unique values can be solved for X, Y, and Z in the above equations.

As a rule of thumb, an equation is exactly identified if the number of exogenous variables excluded from the equation is equal to the number of endogenous variables included in the equation minus one. An equation is underidentified if the former is less than the latter and overidentified if the former is greater than the latter.

For example, in the following causal systemwhere both Y_1 and Y_2 are endogenous variables (since they are caused by each other), both equations (1) and (2) are underidentified:

$$Y_1 = \hat{B}_1 Y_2 + e_1 \qquad (1)$$

$$Y_2 = \hat{B}_2 Y_1 + e_2 \qquad (2)$$

One could of course reformulate the above equations so that they are exactly identified. This is accomplished by adding appropriate exogenous variables into the causal system:

$$Y_1 = B_1 Y_2 + B_2 X_2 + e_1 \qquad (3)$$

$$Y_2 = B_3 Y_1 + B_4 X_1 + e_2 \qquad (4)$$

Here, two exogenous variables X_1 and X_2 have been added to the causal system. Applying the rule of thumb, one finds that equation (3) excludes one exogenous variable (X_1) and includes two endogenous variables (Y_1, Y_2). That is, the number of exogenous variables excluded from the equation is equal to the number of endogenous variables included in the equation minus one. The equation is therefore exactly identified. The same argument applies to equation (4).

One may further reformulate the causal system so that the equations are overidentified, as follows:

$$Y_1 = B_1Y_2 + B_2X_1 + B_3X_3 + e_1 \qquad (5)$$

$$Y_2 = B_4Y_1 + B_5X_2 + B_6X_4 + e_2 \qquad (6)$$

Note that two more exogenous variables (X_3, X_4) have been added to the causal system. The number of exogenous variables excluded from equation (5) is now two (i.e., X_2, X_4) while the number of endogenous variables (i.e., Y_1, Y_2) included in the equation remains the same. Equation (5) is thus overidentified. For the same reason, equation (6) is also overidentified.

As mentioned earlier, when a causal equation is underidentified, no unique estimation can be made of the parameters. An exactly identified equation yields a single unique estimate for each parameter and an overidentified equation gives more than one estimate for each parameter. When two or three unique estimates are possible, the researcher will have the advantage of being able to cross-validate his theoretical formulations of causal influence. That is to say, if his causal theory is valid the multiple estimates for the same parameter would be highly similar to each other.

It will be noted that, in general, the econometric models require the researcher to make simplifying assumptions about the causal system. Among other things, the exogenous variables have to be identified for inclusion into the causal model. Another drawback of the econometric models (when used in the behavioral sciences) is that they assume a one-directional causal influence. Where congruent and incongruent influences co-exist, the models may yield misleading results.

Time-series models:

Time series data dealt with in the present study is confined to the two-variable-measured-twice category. Causal analysis conducted in the two-variable-measured-twice setting dates back to the first half of the century when Lazarsfeld (1947) developed the sixteenfold table to study the causal influence of voting behavior. The model examines shifts of dichotomous variables at two time points and measures the strength of causal influence on the basis of frequencies of shifts. The method has generally been used to infer the relative strengths or depths of various attitudes and may best be illustrated by such an example (see Campbell and Stanley, 1963). Suppose measures of warmth and coldness are taken from 100 teachers and measures of responsiveness and unresponsiveness are taken from their students. The researcher conducts a correlational analysis on the data and finds a positive correlation between warm teachers and responsive students. The equivocality of the correlation coefficient leaves the researcher with the question: does teacher warmth cause student responsiveness, or do responsive students bring out warmth in teachers. To resolve the ambiguity, measures of warmth and responsiveness may be taken from the same teachers and students at a second time point generating sixteen response patterns:

Student change from time 1 to time 2

Teacher change from time 1 to time 2	R → R	R → U	U → U	U → R
W → W				
W → C				
C → C				
C → W				

C: Cold R: Responsive
W: Warm U: Unresponsive

Suppose the following results are obtained from this particular study on teacher and students attitudes:

Time 1

Student	Teacher: Cold	Warm
Responsive	20	30
Unresponsive	30	20

Time 2

Student	Teacher: Cold	Warm
Responsive	10	40
Unresponsive	40	10

It will be noted that two opposite conditions may bring about the shifts that have occurred between time 1 and time 2:

Condition 1: Teacher warmth causing student responsiveness.

Student	Teacher: Cold	Warm
Responsive	10 10	30
Unresponsive	30	10 10

Condition 2: Student responsiveness causing teacher warmth.

Student	Teacher: Cold	Teacher: Warm
Responsive	10 10 →	30
Unresponsive	30	← 10 10

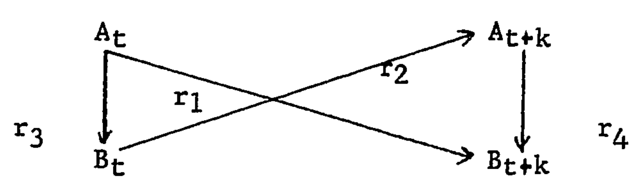
The above diagrams involve four types of shifters which tend to increase the correlation. All four types of response patterns could, of course, occur simultaneously. In such cases, any inference with respect to the direction of causation would be based on the predominance of one over the other. If only one (or two compatible shifters) occurred, it would be plausible to make a one-directional causal inference. For instance, if only condition 1 held, it would be highly plausible that teacher attitudes are changing student attitudes.

One limitation of the sixteenfold table is obvious: the model is applicable to only dichotomous data. Another weakness of the design, as suggested by Campbell and Stanley (1963), appears to be that regression becomes a problem if the marginals of either variable are badly skewed (e.g., 10-90 splits instead of the 50-50 splits used in the example).

Lipset et al. (1954) introduced the notion of "qualifier" to the use of the sixteenfold table. A "qualifier" is simply a variable that can be used to create subsets of data which can then be analyzed separately. The procedure is designed to enhance the power or precision of the sixteenfold table. Using "political interest" as a "qualifier," Lipset et al. (1954) found that subjects with originally high interest in political affairs were positively affected by party contact. Subjects with originally low interest were, however, negatively affected by party contact. Had the two interest groups been considered together (i.e., without using the "interest" qualifier) the differential effects of party contact would have been obscured.

The use of "qualifier" may, however, give rise to statistical regression effects. Campbell and Clayton (1961), for instance, pointed out that for situations in which three variables were positively but imperfectly correlated, of two groups of subjects equally high on one variable, those also high on a second would have regressed less on the third than those low on the second. This was consistent with the general rule that any variable is better predicted from the sum of two predictors than from either alone. Thus, if one regards the "qualifier" matchings as the summing of predictors, it would become obvious that many "qualifier" analyses were probably overinterpreted.

Campbell and Stanley (1963) generalized the sixteenfold table technique to non-dichotomous data to develop what they described as the cross-lagged correlation model. The method was also proposed by Pelz and Andrews (1964) apparently independent of Campbell and Stanley. The cross-lagged correlation method is based on the rule of time precedence. When a given event consistently precedes the occurrence of another and the reverse does not hold, two hypotheses become plausible: (a) Event 1 is a cause of Event 2; or (b) both Events are the effects of some more general cause. (See Crano et al. 1972.) The model makes causal inferences on the basis of the relative magnitudes of the cross-lagged correlations. This may be illustrated by the following diagram:



If A and B are the cause and effect variables, respectively, r_1 will be greater than r_2, k being a remeasurement interval which is close to the causal interval needed for a change in A to effect a change in B. In addition, cross-lagged correlation r_1 should also be greater than either simultaneous correlation r_3 or r_4, if k is close to the causal interval.

Several assumptions are involved in the cross-lagged correlation model. First, the model assumes that changes constantly occurred in the state of variable A for each subject. If all subjects' A value remained the same for both time points, all cross-lagged correlations among any A or B would approach whatever value existed for r_1. Second, it is assumed that the causal effect of A on B is not immediate but occurs over a time interval. If A and B were to change simultaneously, r_3 would equal unity and the cross-lagged correlations would have the same value. Third, it is assumed that A is not completely consistent or markedly inconsistent over time. If there were marked fluctuations in A, and if the interval of remeasurement differed from the interval of causation, the cross-lagged correlations might become negative, and the direction of the difference opposite to what would be predicted. Fourth, it is assumed that the consistency of A or B does not change over time. If A or B were progressively more or less stable, the method might run into difficulty. Finally, the relationship between A and B is assumed to be linear. Causal analysis conducted with the model is valid only to the extent to which these assumptions are satisfied.

It should also be noted that the cross-lagged correlation model does not consider the direction or nature of the causal influence. Specifically, it does not take into account cases where the causal influence is toward incongruity (i.e., cases where the cause variable causes the effect variable to change in the opposite direction). The co-existence of congruent and incongruent influences tends to lower the cross-lagged correlation, leading to erroneous interpretation.

This limitation of the cross-lagged correlation method lies in the fact that the results the model yields are necessarily composite in nature. Yet, a set of panel data involving two variables produces four possible causal relations as follows:

$$X \xrightarrow{+} Y$$

$$X \xrightarrow{-} Y$$

$$X \xleftarrow{+} Y$$

$$X \xleftarrow{-} Y$$

The plus and minus signs indicate congruent and incongruent influence, respectively. In most social situations it is not always possible to rule out some of these causal relations a priori. Moreover, there is always the possibility that unidentified subgroups exist in the total sample for which the causal relations of the variables are different from that for the other groups. Thus, Crano et al. (1972) report that the predominant causal sequence over all subjects in their study was in the direction of intelligence causing later achievement. Dividing the total sample into core and suburban group, however, revealed that this sequence held only for the suburban group. The causal relationship of intelligence and achievement for the core subgroups was opposite to that of the suburban group. The authors conclude that a combination of the data from the core and suburban subjects can be extremely misleading. Yee (1966) has also found different causal influences between teacher and pupil attitudes for the lower and middle social class samples. In the Eron study (Eron et al., 1972) viewing violent television programs was interpreted as having causal influence on later aggressive behavior. It is however, also plausible that scenes of violence might actually reduce later aggression, as the results of the Feshback and Singer study (1971) seemed to suggest. Similarly, the relationship between absenteeism and school grade may be congruent or incongruent. One could hypothesize that good grades, as a positive reinforcer, enhance school attendance. On the other hand, it also appears plausible that good grades may induce some students to cut classes because of their sense of security.

Another limitation of the cross-lagged correlation model was pointed out by Crano et al. (1972). The authors argued that the cross-lagged correlation model assumes that the common factor structure of the measures employed at both time points remains constant. A necessary consequence of this "stationarity" assumption is that the synchronous correlations are equal at both points in time. Two different sources of change may account for the synchronous correlation differences. First, the loading of a test on one common factor changes while the loading on another common factor remains the same or changes in the opposite direction. This is described as a change in kind. A good example of a change in kind is seen in infant "intelligence" tests. These tests tend to measure motor skills more than mental ability. In the case of intelligence tests for older children the reverse is true. If intelligence (I) and some motor skill (M) are measured at ages 1 and 5 for the same sample and it is found that $r\ I_1M_5$ is greater than $r\ I_5M_1$ we cannot conclude that intelligence causes motor skill. It is more sensible to conclude that two measures of motor skills correlate higher than a measure of motor skill with mental ability. Second, all the loadings of a test may change by a multiplicative constant. That is to say, the common factor structure of the tests does not change, but there are changes in the amount of communality and, therefore, uniqueness. This is described as a change in amount. Changes in amount can occur when (1) the real reliability of Test A decreases over time while that of Test B increases over time; or (2) the specific variance of Test A increases while that of Test B decreases over time.

This assumption of "stationarity" makes it necessary for the researcher to devise some means of correcting for differential reliability and specificity deviations that might occur between measurement periods if the full value of the cross-lagged panel technique is to be realized.

Rozelle and Campbell (1969) proposed the use of a no-cause base line to enhance the power of the cross-lagged correlation model. The authors suggested that in cases where the unlagged or synchronous cross correlation remains constant, this value might be used as the no-cause comparison base. A difficulty of this approach is the attenuation of correlation over time. Inasmuch as one would expect some variables to be weakened with time much more than others, inferring the amount of attenuation would be difficult. The problem can, however, be resolved through the use of some reliability measure. Specifically, the internal consistency reliability coefficients may be computed for each of the two measures at each time point. The attenuation coefficients which would reduce these momentary reliability values to the observed $r_{a_1a_2}$ and $r_{b_1b_2}$ could then be used to attenuate the observed $r_{a_1b_1}$ and $r_{a_2b_2}$ values into cause-free expectations for $r_{a_1b_2}$ and $r_{b_1a_2}$. Rozelle and Campbell (1969) pointed out that this procedure for estimating an acausal attenuated comparison value made it possible to draw causal inference even where no inequality between $r_{a_1b_2}$ and $r_{b_1a_2}$ existed.

From the sixteenfold table and the cross-lagged correlation model Yee (1966) derived two methods of causal analysis and described them as the frequencies-of-shift-across-median model and the frequencies-of-change-in product-moment model. The two models are described in the following sections.

The frequencies-of-shift-across-median may properly be called the eighty-one-fold table technique. The model trichotomizes continuous variables measured at two time-points by their respective medians. Measures of the variables are said to be (1) above the median or high, (2) on the median, or (3) below the median or low. On the basis of shifts of time-one and time-two measures across the medians, a total of 81 response patterns can be identified. Using Yee's study (1966) of teacher and pupil attitudes as an example, a 9 x 9 table may be constructed as follows:

Teacher change from time 1 to time 2	1 H→H	2 H→L	3 L→H	4 L→L	5 H→M	6 L→M	7 M→H	8 M→L	9 M→M
1 H→H									
2 H→L									
3 L→H									
4 L→L									
5 H→M									
6 L→M									
7 M→H									
8 M→L									
9 M→M									

Pupil change from time 1 to time 2

H: High M: Median L: Low

The response patterns suggest that a total of 81 resolutions for the relationship of teachers' and pupils' attitudes are possible. The attitude measures can remain without change relative to pretest medians and posttest medians [High to High (H→H), Low to Low (L→L), and Median to Median (M→M)]. They can shift from pretest medians across posttest median [Median to High (M→H), and Median to Low (M→L)], or shift to and across posttest medians from positions above or below pretest medians [High to Low (H→L), Low to High (L→H), High to Median (H→M), and Low to Median (L→M)].

The influence operating in each of the 81 resolutions is judged to be teacher caused or pupil caused on the basis on who shifted most and who shifted least relative to the medians of their pretest and posttest measures. For instance, if the teacher stayed high and the pupil shifted from low to high, then the teacher will be considered to be the source of influence. In cases where teachers and pupils remained without change relative to the medians of pretest and posttest the influence is considered to be uncertain.

The direction of the influence is determined on the basis of the complementarity of teachers' and pupils' attitudes (i.e., whether the relationship is positively or negatively correlated.) If teachers and pupils shifted to or remained in resolutions where their attitudes are more similar than previously, the influence is considered to be toward congruity. If teachers' and pupils' attitudes shifted to or remained in resolutions where their attitudes are more dissimilar, the influence is considered to be toward incongruity.

In Yee's study (1966) the 81 resolutions are classified as follows:

Classification of influence*	Cell**
TC	1.3,1.6,1.7,4.2,4.5,4.8,5.2,6.3,7.3,8.2
TI	1.2,1.5,1.8,4.3,4.6,4.7,5.3,6.2,7.2,8.3
TU	9.2,9.3,9.5,9.6,9.7,9.8
PC	2.4,2.5,2.8,3.1,3.6,3.7,5.4,6.1,7.1,8.4
PI	2.1,2.6,2.7,3.4,3.5,3.8,5.1,6.4,7.4,8.1
PU	2.9,3.9,5.9,6.9,7.9,8.9
UC	1.1,2.2,3.3,4.4,5.5,5.8,6.6,6.7,7.6,7.7, 8.5,8.8
UI	1.4,2.3,3.2,4.1,5.6,5.7,6.5,6.8,7.5,7.8, 8.6,8.7
UU	1.9,4.9,9.1,9.4,9.9

*The upper case letters have the following meanings:

TC: Teacher influence toward congruity.
TI: Teacher influence toward incongruity.
TU: Teacher influence in uncertain direction.
PC: Pupil influence toward congruity.
PI: Pupil influence toward incongruity.
PU: Pupil influence in uncertain direction.
UC: Uncertain influence toward congruity.
UI: Uncertain influence toward incongruity.
UU: Uncertain influence in uncertain direction.

**Cell designations from 81-cell table; first numbers refer to row and second numbers refer to column.

Hypotheses can then be formed on the basis of the classification of causal influence. Various combinations can be made by adding categories together that are similar in value with respect to the source or direction of influence. They are then tested by the chi-square test. The following are some examples:

Hypothesis 1: Teacher-pupil pairs indicating teacher influence toward either congruity or incongruity are more frequent than those indicating pupil influence toward either congruity or incongruity (TC + TI + TU > PC + PI + PU).

Hypothesis 2: Teacher-pupil pairs indicating teacher influence toward congruity are more frequent than those indicating pupil influence toward congruity (TC > PC).

Hypothesis 3: Teacher-pupil pairs indicating teacher influence toward incongruity are more frequent than those indicating pupil influence toward incongruity (TI > PI).

Summary statements of causal influence are made on the basis of the chi-square test results. If, for instance, the result for hypothesis 1 turns out to be significant, then teacher attitudes will be inferred as the source of the causal influence.

A major limitation of the frequencies-of-shift-across-median model stems from the implicit assumption that the variable showing less change is the source of influence. Thus, if observations of the first variable shift from high to median and the corresponding observations of the second variable shift from high to low, the first variable is inferred to be the source of influence. Yet, not all observations of the variables can be expected to shift across the median. In such cases (or when both variables shift from low to high or high to low) causal influence is said to be uncertain. The uncertain cases (i.e., UC, UI, UU) are not included in the chi-square analysis. Tests of hypotheses are therefore based on a restricted number of cases.

Another implicit assumption of the model is that the two sets of measures (e.g., teacher attitudes and pupil attitudes) are equally reliable. If measures of teacher attitudes are more reliable than measures of pupil attitudes, the effects of error variance in the latter might be such that pupil attitudes are shown to have changed more than teacher attitudes. Under these circumstances, teacher attitudes could then be erroneously inferred as the cause of pupil attitudes.

The frequencies-of-change-in-product-moment model is developed to overcome a serious shortcoming of the cross-lagged correlation technique. This limitation, pointed out in an earlier section, pertains to the determination of the direction of influence between correlated variables. In Yee's (1966) study, for instance, the researcher may find $r_{P_1T_2} > r_{T_1P_2}$. If this is the case, the cross-lagged correlation technique would have him infer that pupils' attitudes influence teachers' attitudes. The inference, however, is not the only one possible. The difference between the cross-lagged correlations could have resulted not from pupil influence toward congruity, but rather from teacher influence toward incongruity. That is, teacher influence may be greater than pupil influence. The ambiguity cannot be resolved by the cross-lagged correlation technique because the method confounds the degree and the direction of influence between the two variables.

The frequencies-of-change-in-product-moment model involves the following steps:

(1) Raw scores (e.g., of teachers' and pupils' attitudes) are converted to z scores for each variable and for each time point.

(2) The direction of causal influence (i.e., congruent or incongruent) is determined by the relative magnitudes of the cross-products of pretest z scores and posttest z scores. If the cross-product of the posttest z scores ($z_{T_2}\ z_{P_2}$) is greater than the cross-product of the pretest z scores ($z_{T_1}\ z_{P_1}$), the direction of influence is congruent. If the cross-product of posttest z scores is less than the cross-product of the pretest

z scores, the direction of influence is incongruent. It will be noted that this procedure for determining the direction of influence is logically related to the basic formula for product-moment correlation coefficient:

$$r = \frac{\sum z_x z_y}{N - 1}$$

(3) The source of influence is determined on the basis of the relative magnitudes of cross-lagged z products (i.e., $z_{T_1} z_{P_2}$ and $z_{P_1} z_{T_2}$). When the direction of influence is congruent, the variable whose pretest measure is contained in the more positive product is the source of influence. When the direction of influence is incongruent, the variable whose pretest measure is contained in the more negative product is the source of the influence.

In Yee's (1966) study, each teacher-pupil pair is thus classified under one of four influence categories: TC, TI, PC, PI. The symbols have the same meanings as those presented in the discussion of the frequencies-of-shift-across-median model. Hypotheses with respect to the frequencies of TC, TI, PC, and PI can be formed and tested by the chi-square test. Similar to procedures used in the frequencies-of-shift-across-median model, summary statements of causal influence are made on the basis of the significance of chi-square test results.

To recapitulate, the frequencies-of-change-in-product-moment model focuses on the change of the z score product-moments of variables measured at two time-points. Causal inferences are made on the basis of the relative magnitudes of the product-moments of z values. Specifically, the model assumes that if the product of the time-two z scores (i.e., posttest z scores) is more positive (greater) than that of the time-one z scores (i.e., pretest z scores) for the two variables, the causal influence is toward congruity. Conversely, if the product of the time-two z scores is more negative (less) than that of the time-one z scores, the causal influence is toward incongruity. The source of influence is determined on the basis of the magnitudes of the cross-lagged z products. When causal influence is congruent the source of influence is the time-one variable in the more positive product. When causal influence is incongruent, the source of influence is the time-one variable in the more negative product.

It will be noted, however, that for any individual case the product of the posttest z scores is more positive (greater) than that of the pretest z scores under the following conditions: (i) z_{x_1} is less than z_{x_2} and z_{y_1} is less than z_{y_2}; (ii) z_{x_1} is less than z_{x_2} and z_{y_1} is greater than z_{y_2} (if z_{x_2} is equal to z_{y_1} and z_{x_1} is less than z_{y_2} or z_{x_1} is equal to z_{y_2} and z_{x_2} is greater than z_{y_1}). Condition (ii) is clearly a case of incongruent influence. Yet in the frequencies-of-change-in-product-moment model, this will be classified as an example of congruent influence. Conversely, the product of the posttest z scores is more negative (less) than that of the pretest z scores under two conditions: (i) z_{x_1} is greater than z_{x_2} and z_{y_1} is greater than z_{y_2}; (ii) z_{x_1} is greater than z_{x_2} and z_{y_1} is less than z_{y_2} (if z_{x_2} is equal to z_{y_1} and z_{x_1} is greater than z_{y_2} or z_{x_1} is equal to z_{y_2} and z_{y_1} is greater than z_{x_2}). Here, again, condition (i) is clearly a case of congruent influence which will be misclassified by the frequencies-of-change-in-product-moment model as an example of incongruent influence. It seems evident, then, that apart from the "higher order" difference of z products, the "first order" difference between the four z scores should also be taken into account.

These limitations of the frequencies-of-change-in-product-moment model appear to point out the need for developing a new causal model on the basis of the "first order" difference among the z scores. Such an attempt will be made in a later section of this dissertation.

CHAPTER III

THE PROBLEM

The importance of causal inference in non-experimental settings has been generally recognized by social and behavioral scientists. This has given rise to a number of causal models for investigating social and behavioral phenomena. These models have generated controversies which have contributed little toward the verification of their validity. Several factors appear to have perpetuated the attendant haziness and ambiguity of the various causal methods.

First, most social or behavioral phenomena are multi-faceted and involve an extremely complex system of relationships among the variables. The determination of causal influence is made difficult by this complex network of relationships and further complicated by the nature or direction (i.e., whether the influence is congruent or incongruent) of the relationships.

Second, causal models proposed in the literature are seldom explicit with regard to the assumptions inherent in the models. While it is a truism in research that the validity of a model, causal or otherwise, is relative to the assumptions involved, relatively little attention has been paid to the explicit specification of assumptions in the field of causal statistics.

Third, empirical proof of causal relationships can be obtained only by experimental manipulation. Without being able to manipulate the "cause" variable, one can never be sure that the causal theory involved is valid. Yet, in many, if not most social or behavioral settings, manipulation of the "cause" variable, or for that matter any variable, appears to be extremely difficult. As a matter of fact, it is precisely when such manipulation is not possible that the researcher resorts to the use of causal models. It seems clear therefore that in most cases the validity of a causal model will have to be established (not proved in the empirical sense of the word) through means other than experimental manipulation.

Fourth, and most important, much of the haziness and controversy that pervade the field of causal statistics has been a direct result of the lack of a precise definition of "cause." The discussion of necessary and sufficient conditions of causality has tended to be too philosophical to be of direct value for empirical research aimed at establishing the validity of causal models.

In general, the results obtained in the application of the causal models to empirical data are at best tentative and at worst ambiguous or erroneous. When different models are applied to the same set of data, divergent results are often obtained.

In Yee's (1966) study, hypothesis testing consisted of a series of chi square tests performed on the results of the frequencies-of-change-in-product-moment model (FCP) and the frequencies-of-shift-across-median model (FSM). For instance, if either technique identified 60 PC cases (i.e., Pupils influencing teacher in the Congruent direction) and 40 TC cases (i.e., Teacher influencing pupils in the Congruent direction) a chi square test with 1 df would be performed on the results to test the hypothesis that PC $>$ TC (Pupil influence was greater than Teacher influence, i.e., $60 > 40$).

Yee's study employed five teacher variables and 12 pupil variables. Three hypotheses of particular interest were:

(1) Teacher influence was greater than pupil influence

or TC + TI > PC + PI.

(2) Teacher congruent influence was greater than pupil congruent influence or TC > PC.

(3) Teacher incongruent influence was greater than pupil incongruent influence or TI > PI.

After the hypotheses were tested with regard to each teacher-pupil variable pair, 180 chi square tests were performed. This process was repeated on several subsamples of teachers and pupils.

In a pilot study, the results obtained by Yee (1966) were further analyzed to show that the FCP and FSM might yield different results when applied to the same set of data.

The procedures used consisted of (1) a tabulation of the results of the chi square tests on three subsamples; and (2) A comparison between FCP and FSM using an application of the chi square test on the results tabulated. The following example illustrates the tabulation:

As a result of a series of chi square tests performed on Yee's data we obtain the following:

	Hypothesis 1					Hypothesis 2					Hypothesis 3				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
P1					*					*					
P2	*										*				
P3	*									*	*				
P4	*	*				*	*				*				
P5															
P6															
P7															
P8						*									
P9															
P10	*	*	*	*	*	*			*	*	*	*			
P11					*										
P12															

Note: The above table is actually a simplified version of Table 66 in Yee's monograph (p. 107).

where the T's and P's are teacher and pupil variables, respectively, and an asterisk indicates that a significant ($p < .05$) chi square value was computed when FCP was applied to the data.

Thus, when FCP was used, the chi square test yielded four significant values in column one under hypothesis one; three significant values in column one under hypothesis two; and four significant values in column one under hypothesis three, etc.

How closely do the results of FSM correspond to this pattern of chi square values? In answer to this question, a contingency table is set up as follows:

		FSM +	FSM -
FCP	+	I	II
	-	III	IV

where + indicates that the chi square value is significant and - suggests the reverse. Ideally, (i.e., if FSM and FCP are the same or interchangeable) the frequencies of cells II and III would be zero. To the extent that they are not, we will conclude that the two techniques are in actuality different from each other.

Five such contingency tables were constructed for each of three subsamples included in the pilot study--making a total of 15 contingency tables. Then a chi square test was performed on these contingency tables to determine the degree of association (dependency) between FCP and FSM. Except for the total, which consisted of 540 observations, each of the contingency tables comprised 36 observations. As all contingency tables had 1 df, Yates' correction was applied in the chi squares. The chi square values obtained and their corresponding significance levels are tabulated as follows:

Contingency table	chi square	p
1	2.822	.089
2	1.575	.207
3	2.338	.122
4	.219	.645
5	.823	.632
6	.298	.592
7	12.388	.001
8	.032	.852
9	3.568	.056
10	.106	.744
11	.201	.658
12	.253	.621
13	.770	.616
14	.001	.975
15	3.568	.056
Total	39.877	.000

The above table shows that except for the total, only one chi square value was found to be significant beyond the .05 level. The evidence clearly does not support claims that the FCP and FSM are comparable. While the chi square value for the total contingency table is highly significant, the actual relationship between the two variables (i.e., FCP and FSM) as measured by phi correlation was only .27.

In conclusion, it may be said that the FCP and FSM have been shown to have a low degree of similarity and would not yield comparable results when applied to the same set of data. This finding, in turn, implies that at best, only one of the methods could be depended upon for causal analysis of a particular set of data. At worst, both might be erroneous. Alternatively, the usefulness of the two models may be relative to the nature of the data to which they are applied. In other words, the limitations of the models may be clearly defined if the implicit assumptions about the data are made explicit.

It thus appears that what is needed in the field of causal statistics is a way to verify the validity of such models. This entails a precise operational definition of "cause" and a set of causal data that is consistent with the definition. Furthermore, the causal nature of the data will have to be known, a priori. The obstacle that one would encounter in such an endeavor is clear: while an operational definition of "cause" may not be too difficult to derive, it appears impossible to know, a priori, the causal nature of a set of social or behavioral data. Thus one could argue that high achievement is the cause of high IQ just as convincingly as one could argue that high IQ is the cause of high achievement. As a matter of fact, even such biological variables as height, weight, and strength may present a rather ambiguous network of interrelationships in causal terms.

Yet, causal data do not have to come from the empirical world. As long as they are consistent with a definition of "cause" that is mathematically and conceptually satisfying, such data can be generated. In fact, the use of simulated data appears to have at least two advantages: (1) the causal nature of the data can be predetermined (and will therefore be known, a priori); and (2) their various parameters can be manipulated to test the efficiency of the different causal models. This latter characteristic of simulated data is particularly desirable in view of the fact that many of the causal methods involve implicit assumptions about the data.

CHAPTER IV

PROCEDURES

Definitions and Abbreviations

A number of terms used in this dissertation take on meanings which are different from those customarily assigned to them in the current measurement literature. Abbreviations are used extensively in describing the various causal models in order to make the dissertation more concise. To avoid confusion, key terms and abbreviations used in the dissertation are defined or explained as follows:

I. Definitions

Independent variable: A variable which can be used to explain or account for variance in another variable. It is equivalent to "variance source" as the latter is used in the traditional analysis of variance jargon. As defined and used in this dissertation, an independent variable may or may not be amenable to direct experimental manipulation. In cases where the independent variable is amenable to experimental manipulation, the term has the same meaning as it is used in traditional measurement literature.

Dependent variable: A variable whose variance may be attributed to or explained by another variable--the independent variable. The term, as defined and used in this dissertation, has exactly the same meaning as it is used in current measurement literature.

Quasi-experiment: An experiment or research effort in which the independent as well as the dependent variables are not amenable to direct experimental manipulation.

Causal model: A procedure designed to identify the independent variable in a quasi-experimental situation.

Causality: A relationship between two variables observed in a quasi-experimental setting one of which is identifiable as the independent variable and the other as the dependent variable by means of a causal model. While not directly manipulated by the researcher, the independent variable thus identified is interpreted as though it could be manipulated to have certain effects on the dependent variable.

Cause: The independent variable in a quasi-experimental situation.

Cause variable: A variable identified as the independent variable in a quasi-experiment situation.

Effect variable: A variable identified as the dependent variable in a quasi-experimental situation.

Causal influence: The effect that the independent variable (identified in a quasi-experimental situation by means of a causal model) has on the dependent variable.

Source of influence: A variable identified as the independent variable by means of a causal model. It is equivalentto a cause variable.

Direction of influence: The positive or negative relationship or correlation between the independent variable and the dependent variable in a quasi-experimental situation. If the relationship is positive, the direction of influence is said to be congruent. If the relationship is negative, the direction of influence is said to be incongruent.

Magnitude of influence: The amount of variance in the dependent variable which can be explained by or attributed to the independent variable in a quasi-experimental setting.

Causal parameters: The source, direction, or magnitude of influence which exists between the independent and dependent variables in a quasi-experimental situation.

Influence coefficient: A parameter through which the amount of causal influence as well as the direction of causal influence is manipulated in the creation of the various causal data sets. The parameter takes the following values: 1.0, .50, .25, -.25, -.50, -1.0. For each data set created, the amount of causal influence is directly dependent on the absolute value of the parameter. The sign of the parameter determines the direction of the causal influence (plus for congruent influence and minus for incongruent influence). In describing the causal data sets in later chapters, the influence coefficient is used as an index of the amount and direction of causal influence in a particular data set.

Error ratio: The ratio between the true component and the error component of a measure. An error ratio of 9/1, for instance, simply means that on the average the true score is nine times as great as the error score. Error ratios should be considered as lower bound reliability coefficients. In other words, an error ratio of 9/1 represents a reliability coefficient which is in fact higher than .90.

Causal estimates: The approximations of causal parameters yielded by the various causal models.

Congruity: A state in which the causal influence between the independent and dependent variables in a quasi-experimental situation brings about positive correlation between the two variables.

Incongruity: A state in which the causal influence between the independent and dependent variables brings about a negative correlation between the two variables.

<u>Validity</u>: The degree of accuracy with which a causal model estimates or makes approximations of the causal parameters.

<u>Efficiency</u>: The relative validity of the causal models as when they are compared with one another.

<u>Note</u>: The terms "validity" and "efficiency" are sometimes used interchangeably in this dissertation to refer to the accuracy of causal estimates yielded by the various causal models.

II. <u>Abbreviations</u>

a. <u>Causal models</u>:

CLC: Cross-lagged correlation model.

EC: Econometric model.

FSM: Frequencies-of-shift-across-median model.

FCP: Frequencies-of-change-in-product-moment model.

MFCP: Modified-frequencies-of-change-in-product-moment model.

PC: Part correlation model.

VC: Variance components model.

b. <u>Classification of causal influence</u>:

TC: Teacher influence toward congruity.

TI: Teacher influence toward incongruity.

TU: Teacher influence in uncertain direction.

PC: Pupil influence toward congruity.

PI: Pupil influence toward incongruity.

PU: Pupil influence in uncertain direction.

UC: Uncertain influence toward congruity.

UI: Uncertain influence toward incongruity.

UU: Uncertain influence in uncertain direction.

YC: Variable Y is the source of influence; the direction of influence is toward congruity.

YI: Variable Y is the source of influence; the direction of influence is toward incongruity.

YU: Variable Y is the source of influence; the direction of influence is uncertain.

XC: Variable X is the source of influence; the direction of influence is toward congruity.

XI: Variable X is the source of influence; the direction of influence is toward incongruity.

XU: Variable X is the source of influence; the direction of influence is uncertain.

Defining Cause

Cause is defined on the basis of change. In other words, if a change in a variable can be attributed to some influence, it is conceptually satisfying to say that the influence is the cause of the change, i.e., the cause of the variable. Defining cause in this fashion sidesteps the requirement of asymmetry as it occurs in a correlation coefficient and makes the concept of necessary and sufficient conditicns irrelevant. This study considers cause in terms of change over time as opposed to the if-then logic. It will be noted parenthetically that this definition of cause in no way contradicts the definition listed in a previous section (where cause is defined simply as the independent variable in a quasi-experimental situation). The definition presented in this section is an expansion of the earlier definition.

Specifically, the present study defines cause as follows:

$$\begin{aligned} \text{Let } X_1 &= x + e_1, \\ Z &= z + e, \\ X_2 &= x + e_2 + f(z), \end{aligned}$$

where X_1 and X_2 are measures of the same variable obtained at two time points, the e's are error terms, Z is the cause variable, and f(z) is a function of z. From the above, we have:

$$X_2 - X_1 = x - x + e_2 - e_1 + f(z).$$

Since e_1 and e_2 are error terms, we can assume that their means are zero. This gives $X_2 - X_1 = f(z)$, which says in essence that the change in X (between X_2 and X_1) is a function of z. Since z is contained in Z, Z is defined as the cause of X. Note that this definition of cause represents a basic causal model which can be extended to include more than one "cause" variable and adapted to fit a variety of empirical situations. As an example, the pretest-posttest situation may be represented as follows:

$$\begin{aligned} X_1 &= a + e_1, \\ X_2 &= a + e_2 + f(b), \\ Y_1 &= b + e_3, \\ Y_2 &= b + e_4 + f(b), \end{aligned}$$

where X_1 and Y_1 are the pretest values, X_2 and Y_2 are the posttest values, the e's are error terms, a,b, and f(b) are elements that make up the values of X_1, X_2, Y_1, and Y_2. It will be realized that the above model implies that Y is the cause of X.

Generating Causal Data

Based on the above definition of cause, a computer simulation program was prepared to generate a variety of causal data sets. These data sets are built from orthogonal vectors of random numbers provided by a random number generator. Various causal characteristics are then built into the data. Specifically, 30 causal data sets were generated with the following general data model called the uncorrelated model A:

$$X_1 = a + e_1,$$

$$X_2 = a + e_2 + f(b),$$

$$Y_1 = b + e_3,$$

$$Y_2 = b + e_4 + f(b),$$

where, as indicated earlier, X_1 and X_2 are the pretest and posttest measures of variable X and Y_1 and Y_2 are the pretest and posttest measures of variable Y. As an example, the values of a, b, and e's for the first data set may be obtained as follows:

$$a = .9\ n_1,$$

$$e_1 = .1\ n_2,$$

$$e_2 = .1\ n_3,$$

$$b = .9\ n_4,$$

$$e_3 = .1\ n_5,$$

$$e_4 = .1\ n_6,$$

$$f(b) = 1.0\ b,$$

where the n's are sets of two-digit random numbers. These random numbers are generated by a computer subroutine--GAUSS (IBM, 1968). This subroutine computes normally distributed random numbers with a given mean and standard deviation, which are chosen to be 50 and 25, respectively, for the present study. GAUSS employs the power residue method, which has been extensively studied and thoroughly tested. The method has passed the various tests of randomness, including the chi square test, the runs test, and the autocorrelation test--all with satisfactory results (IBM, 1959; IBM, 1968).

To further demonstrate the statistical properties of the random numbers actually used in the present study, coefficients of skewness and kurtosis were computed for each of the seven sets of random numbers. The coefficients are tabulated as follows:

Random number set	Skewness	Kurtosis
1	-.11	-.57
2	.14	-.24
3	-.04	-.32
4	-.02	.62
5	.34	.17
6	-.17	.31
7	-.06	.21

These indices suggest that the seven sets of random numbers can be considered as normally distributed. The slight departure from normality should have negligible effects on the results obtained in the present investigation. Subroutine GAUSS, incidentally, has the capability of generating 2^{29} random numbers before repeating the cycle. The subroutine is therefore sufficient to produce the desired data sets.

From the general data model 30 causal data sets were created. This was done by varying the coefficients for the n's and f(b) for the equation $X_2 = a + e_2 + f(b)$. The purpose of varying the coefficients of the n's is to manipulate the reliability of the data. For instance, in the first data set, we have set $a = .9\ n_1$ and $e_1 = .1\ n_2$. This is equivalent to setting the error to be one-tenth of the observed value of X_1. If we had set $a = .8\ n_1$ and $e_1 = .2\ n_2$, we would have had a larger error and less reliable value for X_1.

In the present study, the error ratios (i.e., the ratios between the coefficients for a and e) are varied from 9/1 to 5/5 with intermediate steps of 8/2, 7/3, and 6/4. Similarly, by varying f(b), we are in effect changing the amount of influence that the cause variable Y has on X. If we set f(b) = b, we would expect about 50 percent of the variance of X_2 to be attributable to the cause variable Y_1. If we set f(b) = ½b, the amount of variance in X_2 attributable to the cause variable will be reduced to about 25 percent. In the present study the amount of influence that the cause variable has on the effect variable is varied in three steps: from b to ½b and finally to ¼b.

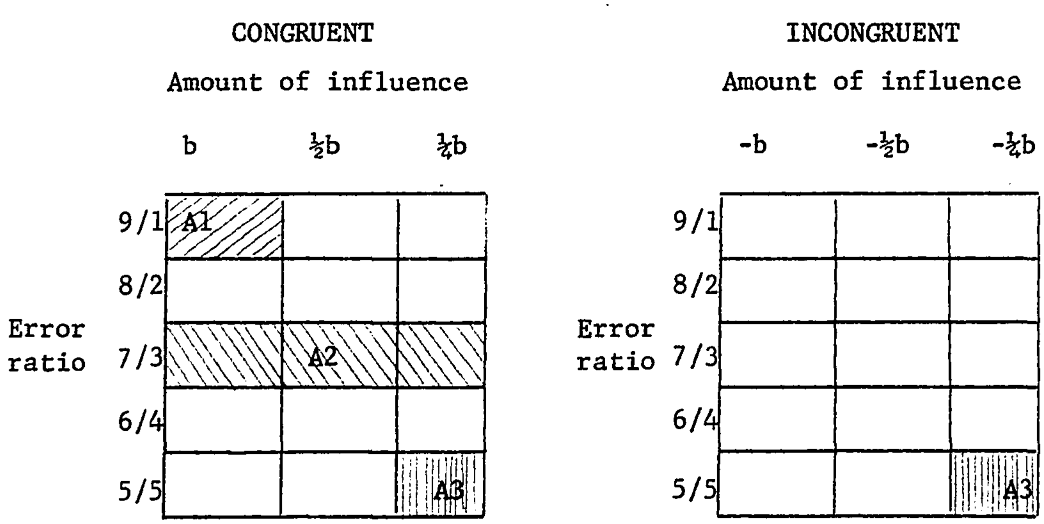
The data sets discussed thus far are either 100 percent congruent or 100 percent incongruent, depending on whether f(b) is positive or negative. Moreover, the amount of causal influence is uniform for all cases in each data set. Pure and simple cases like these are of course not likely to occur in real life situations, especially in social settings. To obtain more "realistic" data sets, two variations of the general data model are created.

The first variation consists of data sets with varying amounts of causal influence (i.e., b, ½b, and ¼b). Specifically, one third of the cases are created with an influence coefficient of 1.0 and the other two thirds are created with influence coefficients of .50 and .25, respectively.

The second variation consists of data sets where both congruent and incongruent causal influences are present. Specifically, one-half of the cases are created with congruent influence and the other half of the cases are created with incongruent influence. In the present study, these data sets are formed simply by combining the corresponding congruent and incongruent data sets created under the original general data model.

The first type of data sets (where causal influence is uniform within each data set) are designated as model A1 data sets. The second type of data sets (where the causal influence is varied) are designated as model A2 data sets. The third type of data sets (where both congruent and incongruent causal influences are present) are designated as model A3 data sets. For each of the model A1 data sets 200 cases were generated. For each of the model A2 data sets 300 cases were generated. For each of the model A3 data sets, 400 cases were obtained. These data sets are diagrammatically presented as follows:

<u>Uncorrelated Data Sets</u>



Note that each cell in the above matrices stands for a model A1 data set. Each three-cell row stands for a model A2 data set. Combining a congruent data set with the corresponding incongruent data set yields a model A3 data set. Thus, 30 model A1 data sets, 10 model A2 data sets and 15 model A3 data sets were created.

In the general data model, X_1 and Y_1 are uncorrelated. While this is perfectly appropriate for model testing purposes, in real life, the two variables could be correlated to some degree, as in the data on teacher and pupil attitudes used in the Yee (1966) study. A correlation between X_1 and Y_1 will also be expected if the data are viewed from a time-series perspective. That is to say, X_2 and Y_2 which are correlated will become X_1 and Y_1 if another set of data is collected at time three.

To generate data where the pretest measures of the two variables are correlated, a second general data model is created as follows called the correlated model B:

$$X_1 = a + b + e_1,$$

$$X_2 = a + b + e_2 + f(c),$$

$$Y_1 = b + c + e_3,$$

$$Y_2 = b + c + e_4.$$

Note that X_1 and Y_1 will be correlated since they have the common element b. Furthermore, the magnitude of the correlation will depend on the magnitude of b relative to a, c, and the error terms. The cause element here is c. As in the first general data model, the magnitude of the cause element in the second general data model is manipulated by assigning to it a coefficient that varies from 1.0 to .25. The magnitude of the error term is also varied according to the following error ratios: 9/1, 8/2, 7/3, 6/4, 5/5.

The seven sets of normal random numbers generated by the computer subroutine GAUSS (i.e., n_1, n_2, n_3, n_4, n_5, n_6, n_7) are used to make up the values of X and Y. The element a is made up of n_1 multiplied by a coefficient. Similarly, b, e_1, e_2, c, e_3, and e_4 are made up of n_2, n_3, n_4, n_5, n_6, and n_7, respectively, each multiplied by a coefficient. For instance, the various elements of a data set may be obtained by the following scheme:

$$a = .45\ n_1,$$

$$b = .45\ n_2,$$

$$e_1 = .1\ n_3,$$

$$e_2 = .1\ n_4,$$

$$c = .45\ n_5,$$

$$e_3 = .1\ n_6,$$

$$e_4 = .1\ n_7,$$

$$f = 1.0$$

The data set thus created will have an error ratio of 9/1 and the amount of causal influence is 1.0. The influence is made congruent or incongruent by switching the sign for f (the coefficient for c) from positive to negative, and vice versa.

Using the procedure described above, three types of correlated data sets (B1, B2, and B3) were created. These are diagrammatically presented as follows:

<u>Correlated Data Sets</u>

CONGRUENT

Amount of influence

Error ratio	c	$\frac{1}{2}$c	$\frac{1}{4}$c
9/1	B1		
8/2			
7/3		B2	
6/4			
5/5			B3

INCONGRUENT

Amount of influence

Error ratio	-c	-$\frac{1}{2}$c	-$\frac{1}{4}$c
9/1			
8/2			
7/3			
6/4			
5/5			B3

Each of the cells in the above matrices stands for a data set where the causal influence is uniform for all cases. These are designated as model B1 data sets. Each row with three cells stands for a data set where the causal influence is varied. These are designated as model B2 data sets. Combining a congruent data set with the corresponding incongruent data set, we obtain a data set where both congruent and incongruent causal influences are present. These are designated as model B3 data sets. For each B1 data set, 200 causal cases were created. For each B2 data set, 300 causal cases were created. Each B3 data set consisted of 400 causal cases.

A word should be said about the relationship between what is described as the error ratio in this dissertation and reliability coefficient as traditionally known in the measurement literature. An error ratio of 9/1 simply means that on the average the true score is nine times as great as the error score. That is, in

$$X = a + e,$$

if $a = .9\ n_1$, then $e = .1\ n_2$, where X is the observed score, a is the true score, e is the error score, n_1 and n_2 are random elements.

As defined by Gulliksen (1950), the reliability coefficient is the ratio of the true variance to the observed variance. Since multiplying an element by a constant (i.e., .9 and .1 in the present case) will multiply the variance by the square of the constant, the ratio between true score variance and error variance will, in the present case, be greater than 9/1. Thus, error ratios should be considered as lower bound reliability coefficients. That is to say, an error ratio of 9/1 represents a reliability coefficient which is in fact higher than .90.

Proposing Three New Causal Models

Three new models are presented in the following sections. These may be described as the modified-frequencies-of-change-in-product-moment (MFCP), the part correlation model (PC), and the variance components model (VC).

Modified-frequencies-of-change-in-product-moment (MFCP)

MECP examines the change in z value of variables measured at two time points. In a two-variable situation, all possible differences of the four z scores (i.e., pretest-posttest z scores for both variables) are considered. These include six initial differences: D_1 is the difference between z_{x1} and z_{y1}; D_2 is the difference between z_{x2} and z_{y2}; D_3 is the difference between z_{x1} and z_{x2}; D_4 is the difference between z_{y1} and z_{y2}; D_5 is the difference between z_{x1} and z_{y2}; D_6 is the difference between z_{x2} and z_{y1}.

From these differences, three "second order" differences may be derived: C_1 is the difference between D_1 and D_2; C_2 is the difference between D_3 and D_4; C_3 is the difference between D_5 and D_6. The following causal interpretations are made on the basis of the magnitude of the "second order" differences:

(1) If C_1 is greater than or equal to zero, which implies that D_1 is greater than or equal to D_2, then it can be inferred that the influence between the two variables is congruent (i.e., the posttest z scores are more similar to each other than the pretest z scores). When the influence is congruent, the following causal links become plausible:

(a) If C_2 is less than zero, which implies that D_3 is less than D_4; then it can be inferred that X, being the variable that changes less (i.e., the difference between z_{x1} and z_{x2} is less than the difference between z_{y1} and z_{y2}), is the source of influence.

(b) Conversely, if C_2 is greater than or equal to zero, we would infer that Y is the source of influence.

(c) If C_3 is less than zero, which implies that D_5 is less than D_6, then it can be inferred that X, being the variable that forces Y to be more similar to itself (i.e., the difference between z_{x1} and z_{y2} is less than the difference between z_{y1} and z_{x2}), is the source of influence.

(d) Conversely, if C_3 is greater than or equal to zero, we would infer that Y is the source of influence.

(2) If C_1 is less than zero, which implies that D_1 is less than D_2, then it can be inferred that the influence between the two variables is incongruent (i.e., the posttest z scores are more dissimilar to each other than the pretest z scores). When the influence is incongruent, the following causal links become plausible:

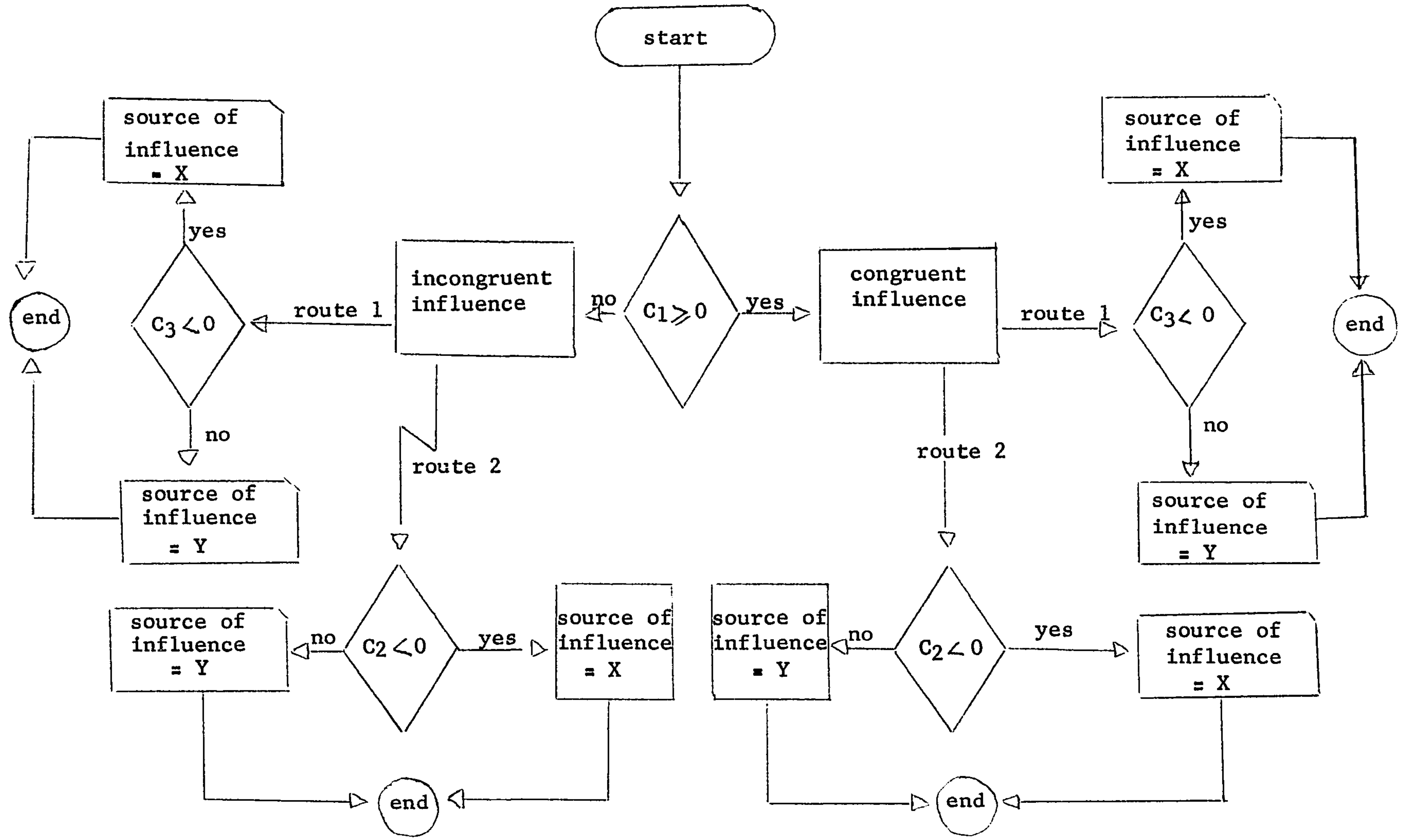
(a) If C_2 is less than zero, which implies that D_3 is less than D_4, then it can be inferred that X, being the variable that changes less (i.e., the difference between z_{x1} and z_{x2} is less than the difference between z_{y1} and z_{y2}), is the source of influence.

(b) Conversely, if C_2 is greater than or equal to zero, we would infer that Y is the source of influence.

(c) If C_3 is less than zero, which implies that D_5 is less than D_6, then it can be inferred that X, being the variable that forces Y to be more similar to itself (i.e., the difference between z_{x1} and z_{y2} is less than the difference between z_{y1} and z_{x2}), is the source of influence.

(d) Conversely, is C_3 is greater than or equal to zero, we would infer that Y is the source of influence.

For clarity, the steps involved in making causal inferences are flow-charted as follows:



Note: In analyses reported in later chapters of this dissertation, routes 1 and 2 have consistently yielded the same causal estimates. The C's are the "second order" differences derived from pretest-posttest z scores. See text on pages 52, 53, and 54.

Part correlation

The part correlation model assumes that the causal system is complete in itself. In other words, the variance in the dependent or endogenous variable is assumed to be fully explainable by the other variables included in the system, barring sampling error. In the two-variable situation where each variable is measured at two time points, we wculd assume that the variance in Y_2 is fully accounted for by Y_1, X_1, and X_2 if X is the source of influence and that the variance in X_2 is fully accounted for by X_1, Y_1, and Y_2 if Y is the source of influence. Four part correlations can be computed from the data: (i) $R_{x_1(y_2.y_1)}$ or the correlation between X_1 and Y_2 with the Y_2 variance due to Y_1 partialled out; (ii) $R_{x2(y_2.y_1)}$ or the correlation between X_2 and Y_2 with the Y_2 variance due to Y_1 partialled out; (iii) $R_{y1(x_2.x_1)}$ or the correlation between Y_1 and X_2 with the X_2 variance due to X_1 partialled out; and (iv) $R_{y2(x_2.x_1)}$ or the correlation between Y_2 and X_2 with the X_2 variance due to X_1 partialled out.

On the assumption that the causal system is complete, the following causal inferences may be made: (a) If $R_{x1(y_2.y_1)}$ is greater than $R_{y1(x_2.x_1)}$ we may infer that X is the source of influence; (b) Conversely, if $R_{y1(x_2.x_1)}$ is greater than $R_{x_1(y_2.y_1)}$ then we may infer that Y is the source of influence. In either case, the sign of the greater part correlation will indicate whether the causal influence is congruent or incongruent.

The part correlation model appears to have the advantage of considering both congruent and incongruent causal influences and testing hypotheses regarding source and direction of influence directly. In the FSM, FCP, and MFCP models, summary statements on source and direction of influence can be made only after the chi square test. The part correlation model also avoids the assumption that the variable that changes less is the source of influence.

The model, incidentally, also appears to be consistent with Portwood's (1972) formulation of causal micro-mathematics. His derivation assumes that a micro-variable can send out causal impulse only to an adjacent microvariable. Conversely, it can only receive causal impulse from its adjacent counterpart. The derivation further assumes that the state of a variable at any particular point of time is at least partially caused by the state of the same variable just prior to that point of time. In the context of the part correlation model, it may be said that X_2 and Y_2 are at least partially caused by X_1 and Y_1, respectively. Hence the partialling out of the criterion (X_2, Y_2) variance which can be attributed to the prior state of the variables (X_1, Y_1).

Variance components model:

This model differs from the correlational models in that it eliminates the confounding effects of "outside" factors in estimating common variation between two variables. The procedure may best be illustrated with an example. In the following equations, let X_1, X_2, Y_1, Y_2 be the time-one and time-two measures of variables X and Y, respectively:

$$X_1 = a + b + e_1,$$

$$X_2 = a + b + e_2 + f(c),$$

$$Y_1 = b + c + e_3,$$

$$Y_2 = b + c + e_4.$$

The covariation between X_1 and Y_2 is determined by the magnitude of the common element b. So is the covariation between X_1 and Y_1. The ratio of the two covariances (i.e., cov X_1 Y_2/cov X_1 Y_1) should therefore have approximately an F distribution with an expected value of unity. On the other hand, the covariation between Y_1 and X_2 is determined by the common elements b and c while the covariation between Y_1 and X_1 is determined by the common element b. The ratio of these two covariances (i.e., cov Y_1 X_2/cov Y_1 X_1) should therefore have approximately an F distribution with an expected value of greater than unity.

Thus, the relative magnitudes of the two F values should indicate the source of causal influence with respect to X and Y. The sign of cov X_1 Y_2 or cov Y_1 X_2 (depending on whether X or Y is found to be the source of causal influence) will suggest the direction of the influence. In other words, if the sign is positive, the causal influence is congruent. If the sign is negative, the causal influence is incongruent.

Let us now consider in some detail the difference between the correlational models and the variance components model. Suppose X and Y are made up of elements as follows:

Case 1 $\quad X = a + b + e_1,$

$\quad Y = c + b + e_2.$

It is obvious that the covariation between X and Y will be determined by the common element b. The correlation between X and Y, on the other hand, will be the ratio of the common variance divided by the total variance. To put this in notation, we have:

$$\text{cov}\ (XY) = b,$$

$$r_{xy} = \text{common variance/total variance}_1.$$

Now suppose variables X and Y are made up of elements as follows:

Case 2 $\quad X = a + b + d + e_1,$

$\quad Y = c + b + e_2.$

The covariation between X and Y will be dependent on the common element b. The correlation between the two variables will, however, be slightly changed. The change occurs because the total variance has increased as a result of the inclusion of element d. To put this in notation, we have:

$$\text{cov }(XY) = b,$$

$$r_{xy} = \text{common variance/total variance}_2.$$

Since total variance_2 is greater than total variance_1, r_{xy} in case 1 will be greater than r_{xy} in case 2 while the covariation in actuality remains the same. The change in magnitude of the correlation is due to the presence of variance attributable to the "outside" element d.

The difference between correlation and covariation may also be seen from the following equations:

$$\text{cov }(XY) = r_{xy}\ (S_x)\ (S_y), \qquad (1)$$

$$r_{xy} = \text{cov }(XY)/S_x\ S_y. \qquad (2)$$

The magnitude of the denominator in equation (2) varies with the number of elements included in variables X and Y. It follows that when the covariation between X and Y is held constant the value of r_{xy} will fluctuate with the number of elements included in the variables.

It should be noted that the variance components model, as it is formulated here, will not be applicable to situations where the time-one measures of variables X and Y are uncorrelated. A1, A2, and A3 uncorrelated data sets used in the present study are examples of such situations. This limitation of the variance components model is illustrated as follows:

Suppose we have the following data model:

$$X_1 = a + e_1,$$

$$X_2 = a + f(b) + e_2,$$

$$Y_1 = b + e_3,$$

$$Y_2 = b + f(b) + e_4.$$

Since X_1, Y_1, Y_2 have nothing in common, we have:

$$\text{cov } (X_1\ Y_1) = 0,$$

$$\text{cov } (X_1\ Y_2) = 0,$$

which yields an F value of:

$$\text{cov } (X_1\ Y_2)/\text{cov } (X_1\ Y_1) = 0/0, \text{ which is indeterminate.}$$

On the other hand, since cov $(Y_1\ X_2) \neq 0$, we have:

$$\text{cov } (Y_1\ X_2)/\text{cov } (Y_1\ X_1) = \text{cov } (Y_1\ X_2)/0, \text{ which is infinity.}$$

Due to this limitation, the variance components model is applied to only model B1, B2, and B3 correlated data sets in the present study.

Applying the Causal Models to Simulated Data

Seven causal models are used to analyze the created causal data. These include frequencies-of-change-in-product-moment (FCP), frequencies-of-shift-across-median (FSM), modified-frequencies-of-change-in-product-moment (MFCP), cross-lagged correlation (CLC), part correlation (PC), econometric model (EC), and variance components model (VC). These models are selected because they appear to be distinctly different from each other and are all applicable to panel data on continuous variables.

The application of FCP, FSM, MFCP, CLC, PC, and VC is direct and presents no problems. The use of EC, on the other hand, requires careful formulation of the causal equations. Specifically, the equations should be formulated in such a way that they are identified. We recall that if the equations are not identified, no unique estimates of the parameters can be obtained.

Bearing the identification problem in mind, we consider the following equations:

$$X_2 = \hat{B}_1 Y_1 + e_1$$

$$Y_2 = \hat{B}_2 X_1 + e_2$$

The first equation is derived from the hypothesis that Y is the cause of X and the second equation is derived from the hypothesis that X is the cause of Y. In this causal system, X_2 and Y_2 are endogenous variables, being caused by Y_1 and X_1 respectively. Y_1 and X_1, on the other hand, are regarded as predetermined or exogenous. Thus the system contains two endogenous and two exogenous variables. In either equation, an exogenous variable is excluded and an endogenous variable is included. Applying the rule of thumb for identification (number of exogenous variables excluded $\geq$ number of endogenous variables included minus one), we have $1 > 0$. Both equations are therefore identified and ordinary least squares can then be used to solve the equations for $\hat{B}_1$ and $\hat{B}_2$.

When the various models are used to analyze the same sets of data whose causal characteristics are known a priori, a comparison of the results yielded by the different models should tell something about the relative strengths and weaknesses of the models. It will also provide some insights with respect to the assumptions inherent in the models. Whenever possible (e.g., in FSM, FCP, and MFCP) each hypothetical subject (or causal case) is examined to see if the correct causal inferences are made with regard to that particular subject.

In applying the models to data sets of varying amounts of causal influence, we are in essence testing the efficiency of the various models. In other words, a model that is capable of detecting causal relationship when the amount of influence is relatively small can be regarded as more efficient or powerful than one that does not detect the relationship.

The direction of the causal influence serves a similar purpose. An efficient model should be able to detect causal relationship whether the nature of such relationship is congruent or incongruent or both.

Another major concern of the present investigation pertains to the amount of error built into the data. Apart from being orthogonal to one another, the error and true score vectors are weighted with coefficients. The relative magnitudes of these vectors thus have great implications with respect to the nature of the data set. Specifically, the reliability of the data is directly related to the relative magnitudes of the vectors. By weighting the vectors differentially we have in fact obtained data of varying reliability levels. The effects that data reliability has on the estimation of causal influence are therefore being investigated as well.

Note: All computer programs (with the only exception of GAUSS) employed in the present investigation were written by the author. Listings of the programs will be made available to interested individuals upon request.

CHAPTER V

RESULTS

The major concern of the present study is the validity or utility of the various causal models. In other words, the investigation is centered on the question of whether or not these models are able to provide correct estimates of causal parameters with respect to the source of causal influence as well as the direction of such influence. It is not expected that a single model will emerge as the best with respect to all kinds of data. It is only expected that some models will be more efficient than others in dealing with certain kinds of data.

The discussion of the efficiency of the models is divided into six sections--each dealing with a particular data model included in the present study. These sections are preceded by an introductory section explaining the decision rules used in determining the strengths and weaknesses of the models.

Decision Rules

In determining the correctness of the estimation of causal parameters the following decision rules are applied:

1. Cross-lagged correlation model (CLC). (1) If the correlation between Y_1 and X_2 is significant and the correlation between X_1 and Y_2 is not, the estimate is considered to be correct. (2) If both correlations are significant or non-significant, the estimate is considered to be ambiguous. (3) If the correlation between Y_1 and X_2 is not significant and the correlation between X_1 and Y_2 is significant, the estimate is considered to be erroneous. (4) The alpha level is .05, with df = N-2.

If the sign of the correlation between Y_1 and X_2 coincides with the sign of the influence coefficient, the direction of the causal link is correctly estimated. If the signs run counter to each other, the estimation of the direction of causal link is erroneous.

2. Part correlation model (PC). (1) The F values for $R_{X_1(Y_2.Y_1)}$ and $R_{Y_1}(X_2.X_1)$ are computed (see McNemar p. 322). If the F value for $R_{Y_1}(X_2.X_1)$ is significant and the F value of $R_{X_1}(Y_2.Y_1)$ is not, the estimation of causal influence is considered to be correct. (2) If the F value of $R_{X_1(Y_2.Y_1)}$ is significant and the F value for $R_{Y_1(X_2.X_1)}$ is not, the estimate is considered to be erroneous. (3) If both F values are significant or non-significant, the estimation of the causal parameters is considered to be ambiguous. (4) The alpha level is .05, with df = 1/(N-3).

If the sign of the correlation coefficient coincides with the sign of the influence coefficient, the direction of causal influence is correctly estimated. If the signs run counter to each other, the estimation is erroneous.

3. Econometric model (EC). (1) The t values of beta weights ($\hat{B}$'s) for X_1 and Y_1 are computed (see Wonnacott and Wonnacott, p. 27). If the t value of $\hat{B}_{Y_1}$ is significant and the t value for $\hat{B}_{X_1}$ is not, the estimation of causal parameters is considered to be correct. (2) If the t value of $\hat{B}_{X_1}$ is significant and the t value of $\hat{B}_{Y_1}$ is not, the estimate is considered to be erroneous. (3) If both t values are significant or non-significant, the estimate is considered to be ambiguous. (4) The alpha level is set at .05 with df = N-2.

If the sign of $\hat{B}$ for Y_1 coincides with the sign of the influence coefficient, the direction of the causal link is correctly estimated.

The estimate is erroneous if the signs are different.

4. <u>Frequencies-of-change-in-product-moment (FCP)</u>

Notations: YC is the number of instances where variable Y is inferred as the source of the causal influence and the direction of such influence is congruent. YI is the number of instances where variable Y is inferred as the source of the causal influence and the direction of such influence is incongruent. Conversely, XC is the number of instances where variable X is inferred as the source of the causal influence and the direction of such influence is congruent. XI is the number of instances where variable X is inferred as the source of the causal influence and the direction of such influence is incongruent.

<u>Rules</u>: (1) The chi square test is used to test the following hypotheses: (a) $YC + YI > XC + XI$, and (b) $YC > YI$. (2) If both chi square values are significant and in the expected direction, the causal influence is correctly estimated. (3) If either chi square value is not significant, the estimate is ambiguous. (4) If both chi square values are not significant, the estimate is also considered to be ambiguous. (5) If either chi square value is significant but in the wrong direction (e.g., $YC > YI$ when $YI > YC$ is expected) the estimate is erroneous. (6) If both chi square values are significant but in the wrong direction, the estimate is also erroneous. (7) The alpha level is .05 and all chi square values have $df = 1$.

The above decision rules apply to only data sets where the causal influence is either congruent or incongruent. In testing models with data where both congruent and incongruent influences are present, the decision rules are modified as follows: (1) If the chi square value computed for the first hypothesis (i.e., YC + YI > XC + XI) is significant and the chi square value for the second hypothesis (i.e., YC > YI) is not, the estimate is considered to be correct. (2) If both chi square values are significant or non-significant, the estimate is ambiguous. (3) If the chi square of the first hypothesis is not significant, the estimate is ambiguous--regardless of whether or not the second chi square value is significant. (4) If the chi square for the first hypothesis is significant but in the wrong direction (i.e., XC + XI > YC + YI), the estimate is erroneous--regardless of whether or not the second chi square value is significant. (5) The alpha level is .05 and all chi square values have df = 1.

5. Frequencies-of-shift-across-median (FSM)

Notations: YC, YI, XC, XI have the same meaning as in FCP. In addition, XU is the number of instances where variable X is inferred as the source of the causal influence but the direction of such influence is uncertain. Similarly, YU is the number of instances where variable Y is inferred as the source of the causal influence but the direction of such influence is uncertain. UC is the number of instances where the source of the causal influence is uncertain but the direction of such influence is congruent. UI is the number of instances where the source of the causal influence is uncertain but the direction of such influence is incongruent. UU is the number of instances where both the source and the direction of the causal influence are uncertain.

Rules: All rules are the same as those applied to FCP, except for the first hypothesis, which now takes the form: YC + YI + YU >

XC + XI + XU.

6. Modified-frequencies-of-change-in-product-moment (MFCP)

All notations and rules are the same as those applied to the frequencies-of-change-in-product-moment model (FCP).

7. Variance components model (VC). (1) F ratios are computed for the covariances as follows:

$$F_1 = \text{cov}\ (X_1, Y_2)/\text{cov}\ (X_1, Y_1)$$

$$F_2 = \text{cov}\ (Y_1, X_2)/\text{cov}\ (Y_1, X_1)$$

If F_2 is significant and F_1 is not, the estimate is considered to be correct. (2) If F_1 is significant and F_2 is not, the estimate is considered to be erroneous. (3) If F_2 and F_1 are both significant or non-significant, the estimate is considered to be ambiguous. (4) The alpha level is .05 and the F values have df = (N-2)/(N-2). (5) If the sign of cov (Y_1, X_2) coincideswith that of the influence coefficient, the direction of the causal link is correctly estimated. If the signs run counter to each other, the estimation of causal direction is erroneous. A minus sign is prefixed to the F value when cov (Y_1, X_2) is negative.

MODEL A1 (Uniform-Influence, Uncorrelated Data Sets)

Recall that all model A1 data sets are based on the following data model:

$$X_1 = a + e_1,$$

$$X_2 = a + f(b) + e_2,$$

$$Y_1 = b + e_3,$$

$$Y_2 = b + f(b) + e_4.$$

These data sets have been created in such a way that the time-one measures of variable X and variable Y are uncorrelated. The amount of causal influence as well as the direction of causal influence which Y has on X is uniform (i.e., the same) for all 200 simulated cases within each data set. A total of 30 data sets are created each of which differs from the other with respect to the amount and direction of causal influence as well as the error ratio.

The skewness and kurtosis of the four variables are computed--separately for each data set. These results are tabulated in Appendix A. The indices, suffice it to say, are generally low, suggesting that the variables have an approximately normal distribution.

The decision rules discussed in the previous section are applied when the efficiency of the various causal models with respect to model A1 data sets is examined. The results are presented model by model in the following sections.

Cross-lagged correlation model (CLC)

The overall results show that CLC appears to perform very well with respect to model A1 data sets. All but five data sets are correctly estimated. No erroneous estimates are obtained. Ambiguous estimation occurs with data sets of relatively low reliability and weak causal influence. It is of interest to note that CLC appears to perform better with data sets of incongruent causal influence than with data sets of congruent causal influence. In fact, all the incongruent data sets are correctly estimated (See Table I).

The effects that the amount of influence and data reliability have on the estimation power are depicted in Figure 1. Data reliability has

Table I. Efficiency of Six Causal Models with Respect to Model A1 Data Sets (N=200)

Causal Model	Magnitude of Causal Influence	9/1	8/2	7/3	6/4	5/5
	1.0	+	+	+	+	+
	.50	+	+	+	+	±
CLC	.25	+	±	±	±	±
	- .25	+	+	+	+	+
	- .50	+	+	+	+	+
	-1.0	+	+	+	+	+
	1.0	+	+	+	+	+
	.50	+	+	+	+	+
PC	.25	+	+	+	+	±
	- .25	+	+	+	+	+
	- .50	+	+	+	+	+
	-1.0	+	+	+	+	+
	1.0	+	+	+	+	+
	.50	+	+	+	+	+
EC	.25	+	+	+	±	±
	- .25	+	+	+	+	+
	- .50	+	+	+	+	+
	-1.0	+	+	+	+	+
	1.0	+	+	+	+	+
	.50	+	±	±	±	±
FSM	.25	±	±	±	±	±
	- .25	+	±	±	±	±
	- .50	+	±	±	±	±
	-1.0	+	+	+	+	+
	1.0	+	+	+	±	+
	.50	+	+	±	+	+
FCP	.25	+	±	±	±	±
	- .25	+	±	±	±	±
	- .50	+	+	±	±	±
	-1.0	+	+	+	±	±

Table I. (Continued) Efficiency of Six Causal Models with Respect to Model A1 Data Sets (N=200)

Causal Model	Magnitude of Causal Influence	9/1	8/2	7/3	6/4	5/5
	1.0	+	+	+	+	+
	.50	+	+	+	+	+
MFCP	.25	+	+	±	±	±
	- .25	+	+	±	±	±
	- .50	+	+	+	+	±
	-1.0	+	+	+	+	±

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

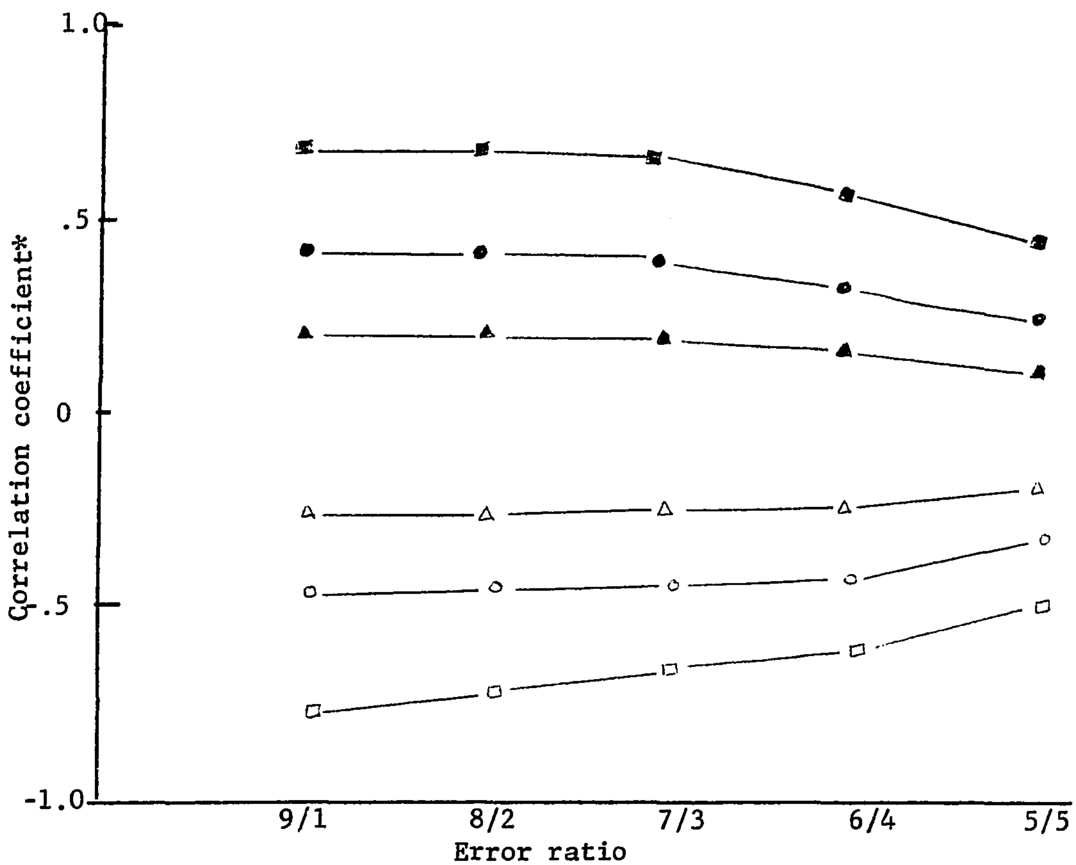


Figure 1. Effects of error and the amount of causal influence on the estimation efficiency of CLC with respect to model A1 (uniform-influence, uncorrelated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-.10
■	●	▲	△	○	□

*indicates that the correlation coefficient pertains to Y as the source of the causal influence.

generally mild effects on the efficiency of estimation. The amount of causal influence, on the other hand, has a very conspicuous effect on the magnitudes of the causal estimates (See Table II).

For A1 data, it may be said that if the data are moderately reliable (i.e., the error ratio is 6/4 or over) and the amount of causal influence is also moderate (i.e., over 25 percent), CLC appears to be an adequate procedure for detecting the source and direction of causal influence.

Part correlation model (PC)

The results appear to overwhelmingly favor this model for A1 data. All but one data sets are correctly estimated. No erroneous estimates are obtained. The only ambiguous estimate occurs with a data set of low reliability and weak causal influence. The magnitudes of the correlations obtained are in general higher than those yielded by CLC (See Tables I and II).

As depicted in Figure 2, data reliability appears to have a greater effect on PC than it does on CLC. The amount of causal influence, while still exerting a conspicuous and consistent impact on the efficiency of the model, appears to have relatively mild effects with highly reliable data. Its effect on data of lower reliability is much more pronounced. It seems evident, then, that when the data are moderately reliable and the magnitude of causal influence is not too low, PC is probably an ideal procedure for detecting causal influence with respect to A1 data.

Econometric model (EC)

The model yields correct estimates for 28 of the 30 A1 data sets. No erroneous estimates are obtained. The ambiguous estimates occur with

Table II. Magnitudes of Causal Estimates Yielded by Six Causal Models with Respect to Model A1 data sets (N=200)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
	1.0	.69	.67	.62	.53	.39	
	.50	.42	.40	.35	.28	.19	
CLC	.25	.20	.19	.16	.11	.05	$r_{Y_1X_2}$
	- .25	-.29	-.29	-.28	-.26	-.22	
	- .50	-.48	-.48	-.45	-.41	-.34	
	- 1.0	-.73	-.71	-.67	-.61	-.50	
	1.0	.98	.92	.80	.62	.43	
	.50	.95	.80	.60	.39	.22	
PC	.25	.85	.57	.35	.19	.08	$R_{Y_1(X_2X_1)}$
	- .25	-.85	-.59	-.41	-.30	-.23	
	- .50	-.95	-.82	-.64	-.48	-.36	
	- 1.0	-.98	-.93	-.83	-.69	-.53	
	1.0	.95	.91	.82	.67	.47	
	.50	.45	.43	.38	.30	.19	
EC	.25	.20	.19	.16	.11	.05	B_{Y_1}
	- .25	-.29	-.29	-.28	-.26	-.23	
	- .50	-.54	-.53	-.50	-.45	-.36	
	- 1.0	-1.04	-1.01	-.94	-.82	-.64	
	1.0	43	46	46	47	49	
	.50	26	30	30	33	39	YC
FSM	.25	16	19	21	20	32	
	- .25	20	17	23	24	29	
	- .50	27	21	27	28	33	YI
	- 1.0	44	35	41	39	38	
	1.0	138	123	102	90	85	
	.50	134	103	90	85	79	
FCP	.25	112	87	80	72	67	YC
	- .25	115	86	72	67	64	
	- .50	136	108	84	75	69	YI
	- 1.0	148	126	102	88	77	

Table II. (Continued) Magnitudes of Causal Estimates Yielded by Six Causal Models with Respect to Model A1 data sets (N=200)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
	1.0	146	122	111	96	88	
	.50	121	101	90	81	77	YC
MFCP	.25	99	83	72	70	68	
	- .25	102	82	73	65	59	YI
	- .50	116	92	76	74	65	
	- 1.0	117	106	92	74	68	

data of relatively low reliability and weak causal influence (See Table I).

Figure 3 depicts the effect of data reliability and the amount of causal influence on estimation power. It will be seen that EC behaves in very much the same way as CLC. That is to say, data reliability appears to have relatively mild effects on estimation while the amount of causal influence appears to be a more dominant factor (See Table II).

The estimation power of EC as suggested by the above results would probably justify applications of the procedure to data of moderate reliability and moderate causal influence. It may also be noted that the model appears to be more efficient with incongruent data sets than with congruent data sets.

Frequencies-of-shift-across-median model (FSM)

The results suggest that the strength of FSM lies only with A1 data sets of high causal influence. The model yields correct estimates for 13

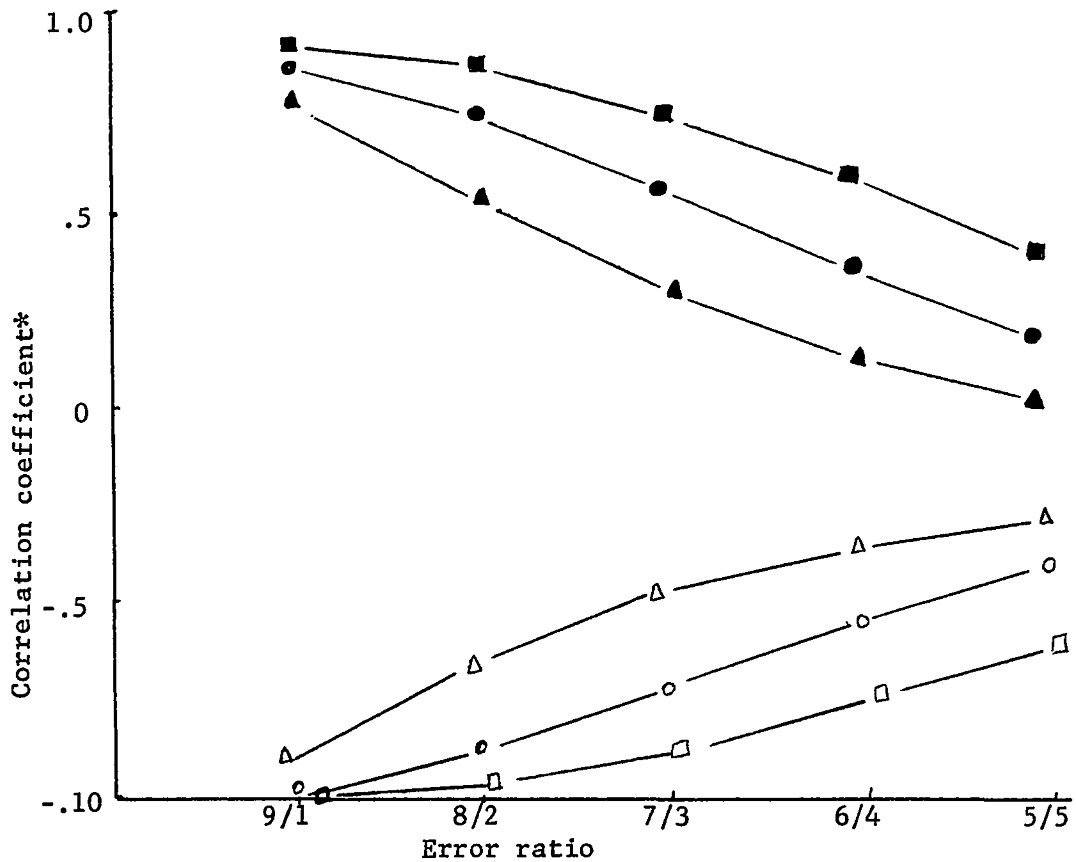


Figure 2. Effects of error and the amount of causal influence on the estimation efficiency of PC with respect to model A1 (uniform-influence, uncorrelated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the correlation coefficient pertains to Y as the source of the causal influence

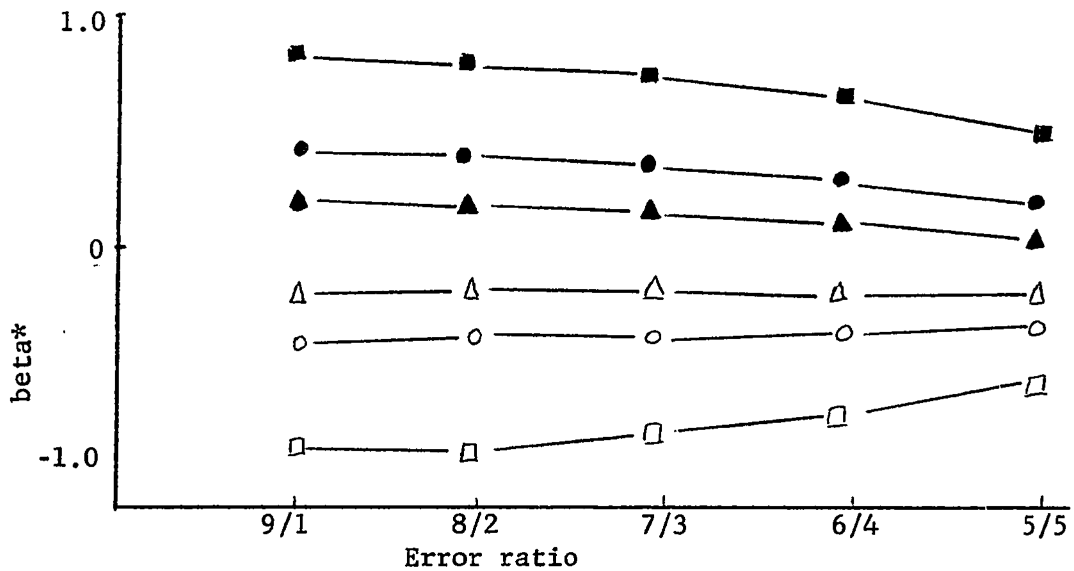


Figure 3. Effects of error and the amount of causal influence on the estimation efficiency of EC with respect to model A1 (uniform-influence, uncorrelated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the beta pertains to Y as the source of the causal influence of the 30 data sets. All these data sets are of high causal influence (i.e., 100 percent) or high reliability. No erroneous estimates are, however, obtained (See Table I).

Figure 4 shows the effects of data reliability and the amount of causal influence on estimation power. The effects of data reliability are shown to be extremely small. The amount of causal influence appears to make a consistent difference in estimation (See Table II).

The above results seem to suggest that the utility of FSM is rather limited with respect to model A1 data sets.

Frequencies-of-change-in-product-moment model (FCP)

In general, the results suggest that FCP may best be used with A1 data of high reliability. The model yields correct estimates for exactly one-half of the 30 data sets. For the other half, the estimates are ambiguous. No erroneous estimates are made (See Table I).

It will be noted that all six data sets of the highest reliability are correctly estimated. Four other data sets which are also of relatively high reliability are also correctly estimated. For the majority of the other data sets, the estimates are ambiguous.

The effects of data reliability and the amount of causal influence are shown in Figure 5. While the amount of causal influence creates a consistent difference in estimation, the magnitude of the difference appears to be small. The effects of data reliability are shown to be consistent and quite pronounced (See Table II).

Modified-frequencies-of-change-in-product-moment model (MFCP)

The model yields correct estimates for 22 of the 30 A1 data sets. Ambiguous estimates are obtained with data sets of weak causal influence

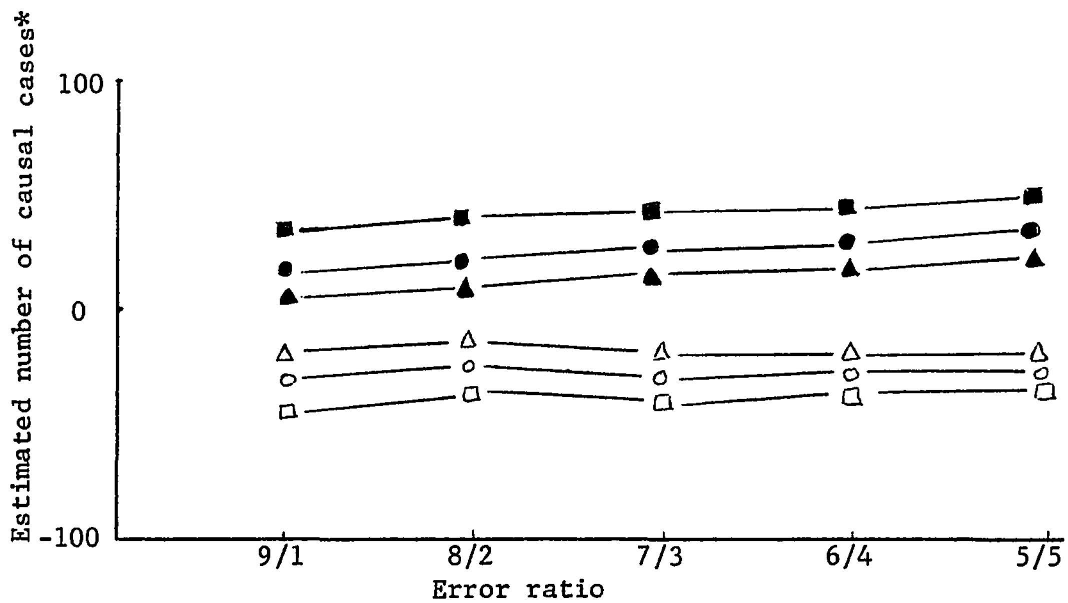


Figure 4. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model A1 (uniform-influence, uncorrelated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

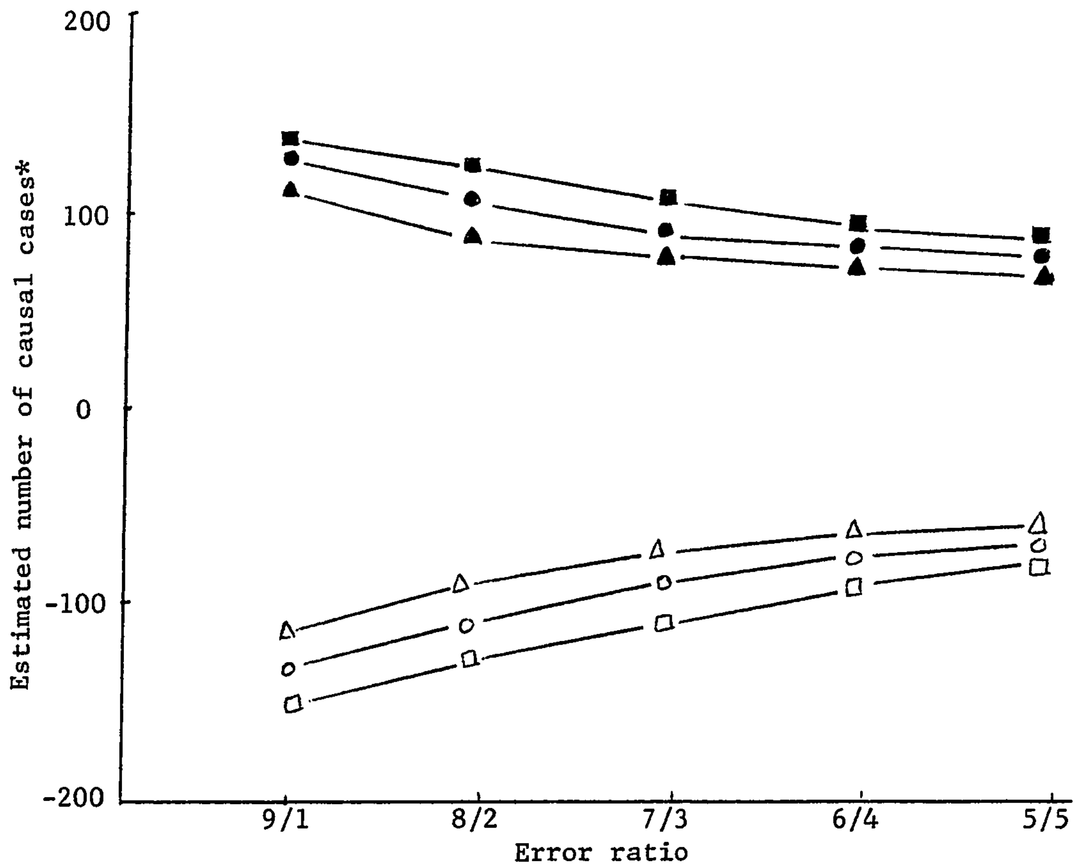


Figure 5. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model A1 (uniform-influence, uncorrelated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

and low reliability. No erroneous estimates are obtained. All data sets of high reliability are correctly estimated (See Table I).

The effects of data reliability and the amount of causal influence on estimation power are depicted in Figure 6. Both factors appear to have moderate and consistent effects on the efficiency of estimation (See Table II).

The above results seem to bear out the expectation that MFCP will be more efficient than FCP with respect to uncorrelated data. As far as model A1 data sets are concerned, causal analysis should probably be conducted by means of MFCP rather than FCP.

Summary

There seems to be little doubt that PC is the most efficient procedure for detecting the source and direction of causal influence with respect to model A1 data sets. In spite of the fact that the procedure is highly sensitive to data reliability, it is the only model that yields practically all correct estimates. EC and CLC appear to have a high enough degree of efficiency to justify their use. MFCP and FCP may be used when the data are known to be at least moderately reliable and the amount of causal influence is substantial. FSM is shown to have very limited utility.

The effects of data reliability and the amount of causal influence on estimation power are consistent and in the expected direction for all causal models. That is to say, error tends to lower the efficiency of estimation and a greater causal influence is invariably followed by a better estimate. It is also evident that not all the causal models are equally sensitive to these two factors. Comparison in this regard can,

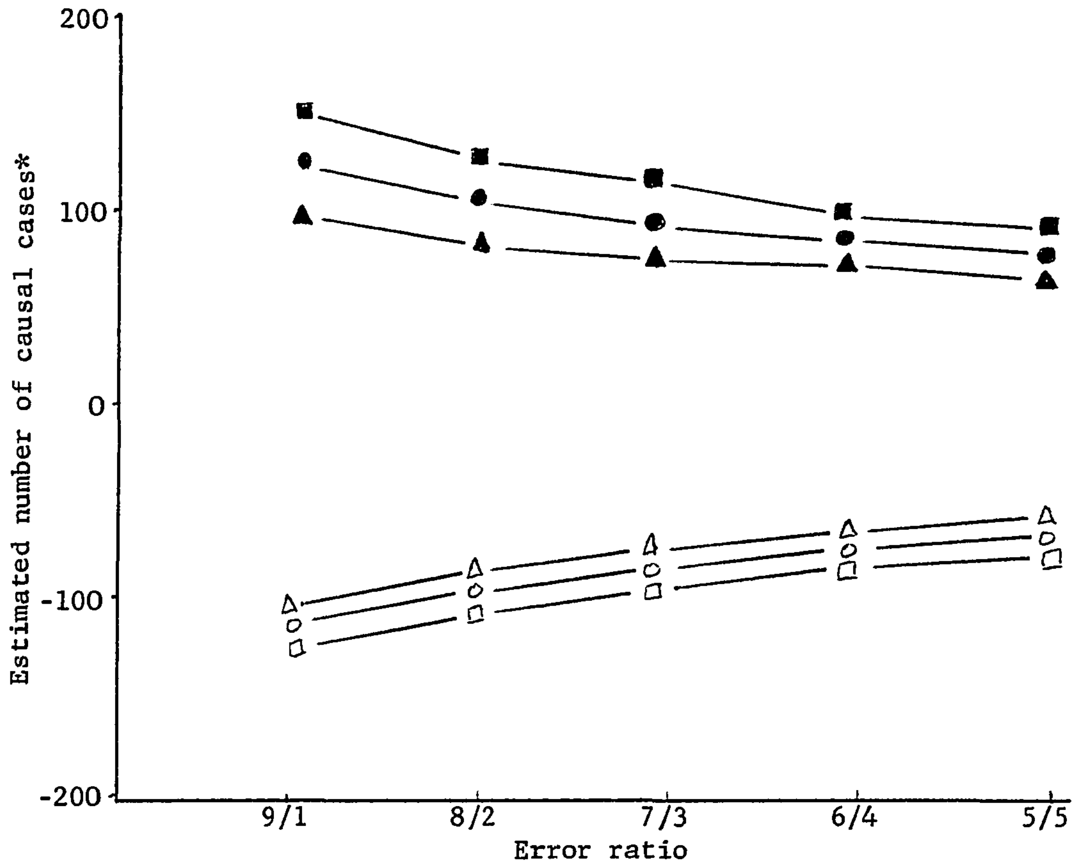


Figure 6. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model A1 (uniform-influence, uncorrelated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence.

however, be made only between models which have the same unit of measurement for causal parameters. For instance, between CLC and PC, the latter appears to be much more sensitive to error, while the former is, in general, more sensitive to the effects of the amount of causal influence. Both data reliability and the amount of causal influence seem to have very similar effects on FCP and MFCP. FSM, on the other hand, appears to be highly insensitive to data reliability and moderately sensitive to the amount of causal influence.

It is important to note that no erroneous causal estimates are yielded by any of the causal models. The ambiguous estimates merely do not provide the basis for reaching a conclusion with regard to the source and direction of causal influence. They do not, however, lead to erroneous causal inferences.

MODEL A2 (Varied-Influence, Uncorrelated Data Sets)

Data sets under model A2 are created from the same data model as model A1. That is, the following equations are used to generate the time-one and time-two measures of the two variables:

$$X_1 = a + e_1,$$

$$X_2 = a + f(b) + e_2,$$

$$Y_1 = b + e_3,$$

$$Y_2 = b + f(b) + e_4.$$

These data sets have been created in such a way that the time-one measures of variable X and variable Y are uncorrelated. While the direction of the causal influence is uniform (i.e., either congruent or incongruent), the amount of causal influence which Y has on X is varied within each data set of 300 simulated cases. Specifically, one third of the cases have been created with an influence coefficient of 1.0 or -1.0, another third of the cases have been created with an influence coefficient of .50 or -.50, and the last third of the cases have been created with an influence coefficient of .25 or -.25. A total of 10 data sets are created, each of which differs from the other with respect to the direction of causal influence and the error ratio.

Coefficients of skewness and kurtosis are computed for each of the data sets. These indices are tabulated in Appendix B. The magnitudes of the coefficients suggest that the distributions of the measures created are approximately normal.

In determining the efficiency of the various causal models, the same decision rules are used. The following sections present results obtained for each of the causal models used in the present study.

Cross-lagged correlation model (CLC)

The results are entirely in favor of the model. All data sets are correctly estimated. See Table III. The effects of data reliability on estimation appear to be consistent but mild (See Table IV and Figure 7).

Part correlation model (PC)

The results suggest that the model is highly efficient for A2 data. All estimates of causal influence are shown to be correct. See Table III. The model is, however, quite sensitive to data reliability and the efficiency of the model tends to be lower when large errors are present (See Table IV and Figure 8).

Table III. Efficiency of Six Causal Models with Respect to Model A2 Data Sets (N=300)

Nature of Causal Influence	Error Ratio	Causal Model					
		CLC	PC	EC	FSM	FCP	MFCP
	9/1	+	+	+	+	+	+
	8/2	+	+	+	+	+	+
Congruent	7/3	+	+	+	+	+	+
	6/4	+	+	+	±	+	+
	5/5	+	+	+	+	+	+
	9/1	+	+	+	+	+	+
	8/2	+	+	+	+	+	+
Incongruent	7/3	+	+	+	+	+	+
	6/4	+	+	+	+	+	±
	5/5	+	+	+	+	+	±

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

Table IV. Magnitudes of Causal Estimates Yielded by Six Causal Models with Respect to Model A2 Data Sets (N=300)

Causal Model	Direction of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
CLC	+	.36	.35	.32	.27	.20	
	-	-.44	-.43	-.41	-.35	-.26	$r_{Y_1X_2}$
PC	+	.62	.57	.47	.35	.23	
	-	-.61	-.56	-.47	-.37	-.26	$R_{Y_1(X_2X_1)}$
EC	+	.51	.49	.43	.34	.24	$\hat{B}_{Y_1}$
	-	-.64	-.62	-.56	-.46	-.32	
FSM	+	50	49	49	51	56	YC
	-	43	47	56	57	53	YI
FCP	+	178	152	140	125	114	YC
	-	171	149	126	109	105	YI
MFCP	+	168	152	135	122	116	YC
	-	171	156	129	104	88	YI

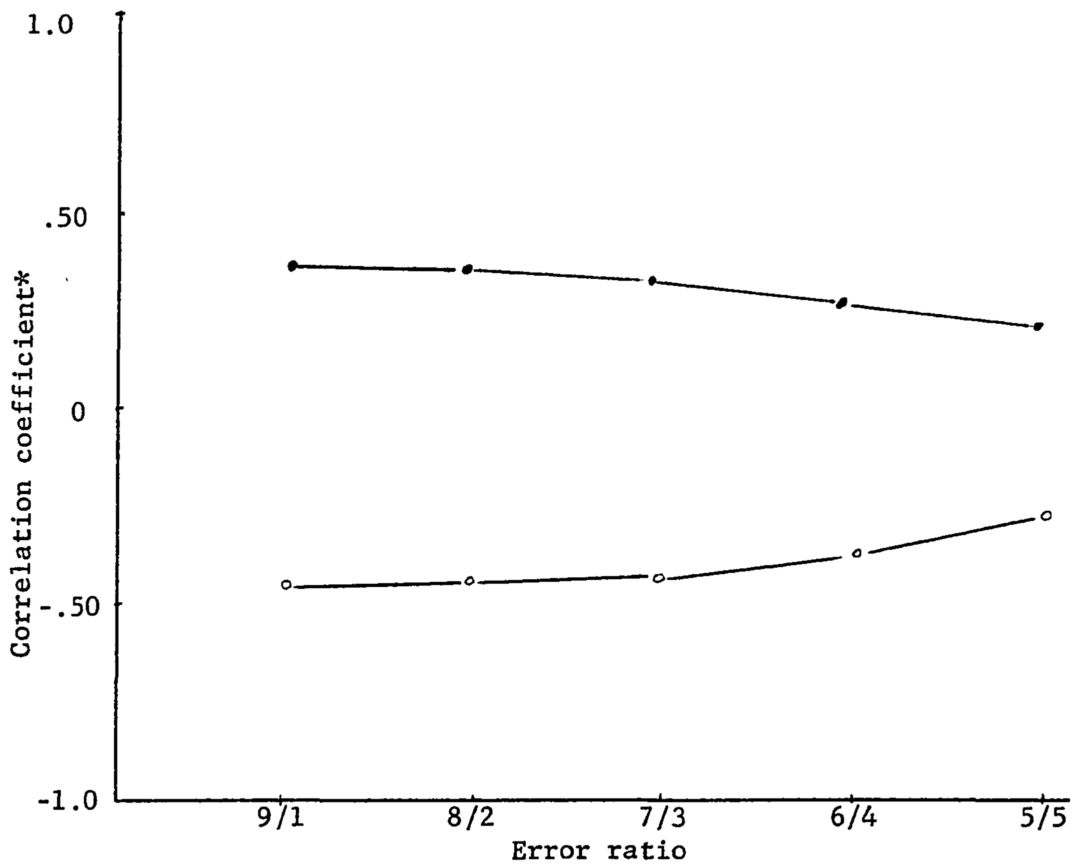


Figure 7. Effects of error and the amount of causal influence on the estimation efficiency of CLC with respect to model A2 (varied-influence, uncorrelated) data sets.

Influence coefficient:

Congruent ●   Incongruent ○

*indicates that the correlation coefficient pertains to Y as the source of the causal influence

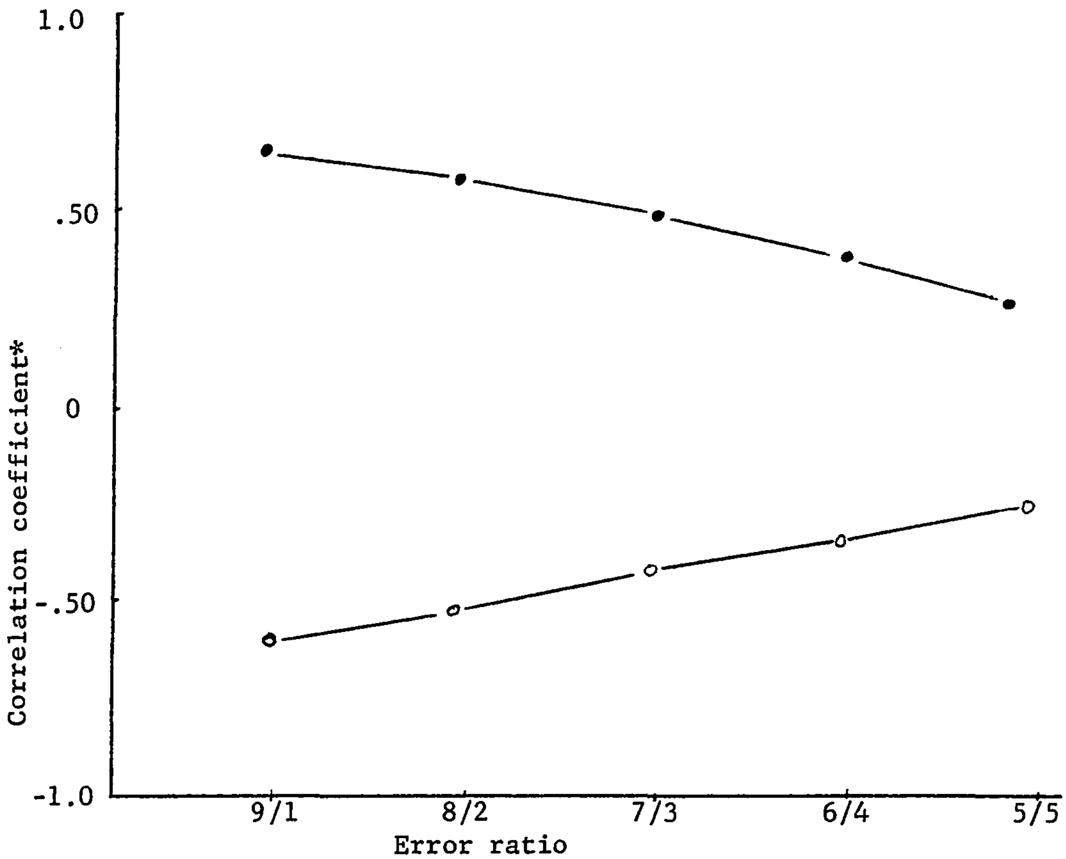


Figure 8. Effects of error and the amount of causal influence on the estimation efficiency of PC with respect to model A2 (varied-influence, uncorrelated) data sets.

Influence coefficient:

Congruent ●   Incongruent ○

*indicates that the correlation coefficient pertains to Y as the source of the causal influence

Econometric model (EC)

The results indicate that the model is highly efficient. All data sets are correctly estimated. See Table III. The effects of error on estimation power are shown to be mild but consistent (See Table IV and Figure 9).

Frequencies-of-shift-across-median model (FSM)

The model correctly estimates all but one data sets. The ambiguous estimate is obtained with a data set of relatively low reliability. See Table III. Data reliability appears, however, to have little effect on the efficiency of the model (See Table IV and Figure 10).

Frequencies-of-change-in-product-moment model (FCP)

The results suggest that the model is highly efficient. All data sets are correctly estimated. The effects of data reliability are shown to be consistent and moderate (See Tables III and IV and Figure 11).

Modified-frequencies-of-change-in-product-moment model (MFCP)

The model yields correct estimates for eight of the 10 data sets. Two ambiguous estimates are obtained with data of low reliability. The effects of data reliability on estimation power appear to be consistent and moderate (See Tables III and IV and Figure 12).

Summary

The results illustrate that most of the causal models are highly efficient with model A2 data sets. Four of the six models (i.e., CLC, PC, EC, and FCP) yield correct estimates for all data sets. FSM and MFCP appear to be slightly less efficient, yielding one and two ambiguous estimates respectively. No erroneous estimates are made by any of the

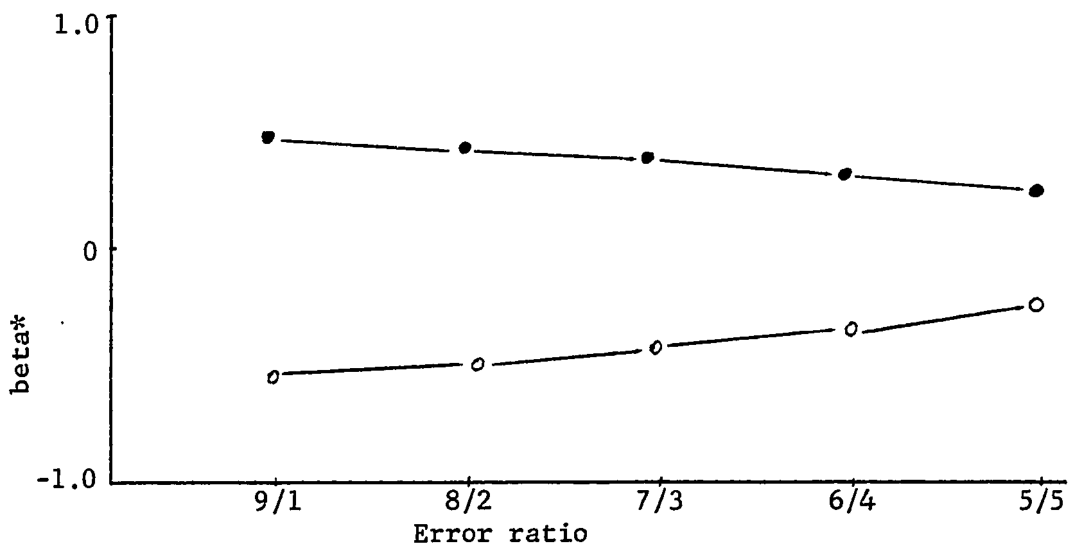


Figure 9. Effects of error and the amount of causal influence on the estimation efficiency of EC with respect to model A2 (varied-influence, uncorrelated) data sets.

Influence coefficient:

Congruent	Incongruent
●	○

*indicates that beta pertains to Y as the source of the causal influence

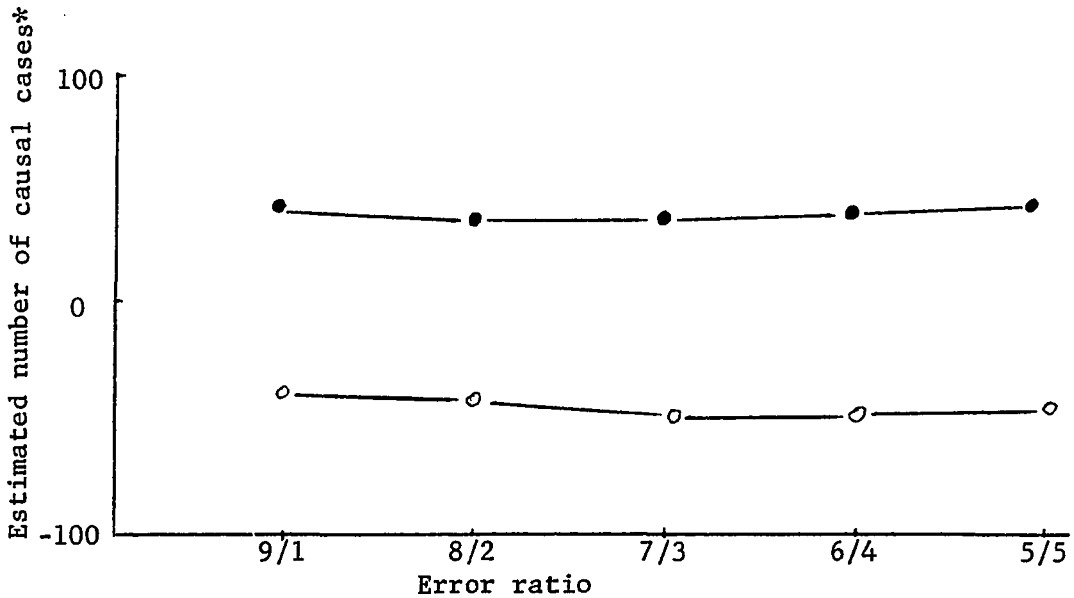


Figure 10. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model A2 (varied-influence, uncorrelated) data sets.

Influence coefficient:

Congruent ●   Incongruent ○

*indicates that the estimated number of causal cases pertains to Y as the source of causal influence

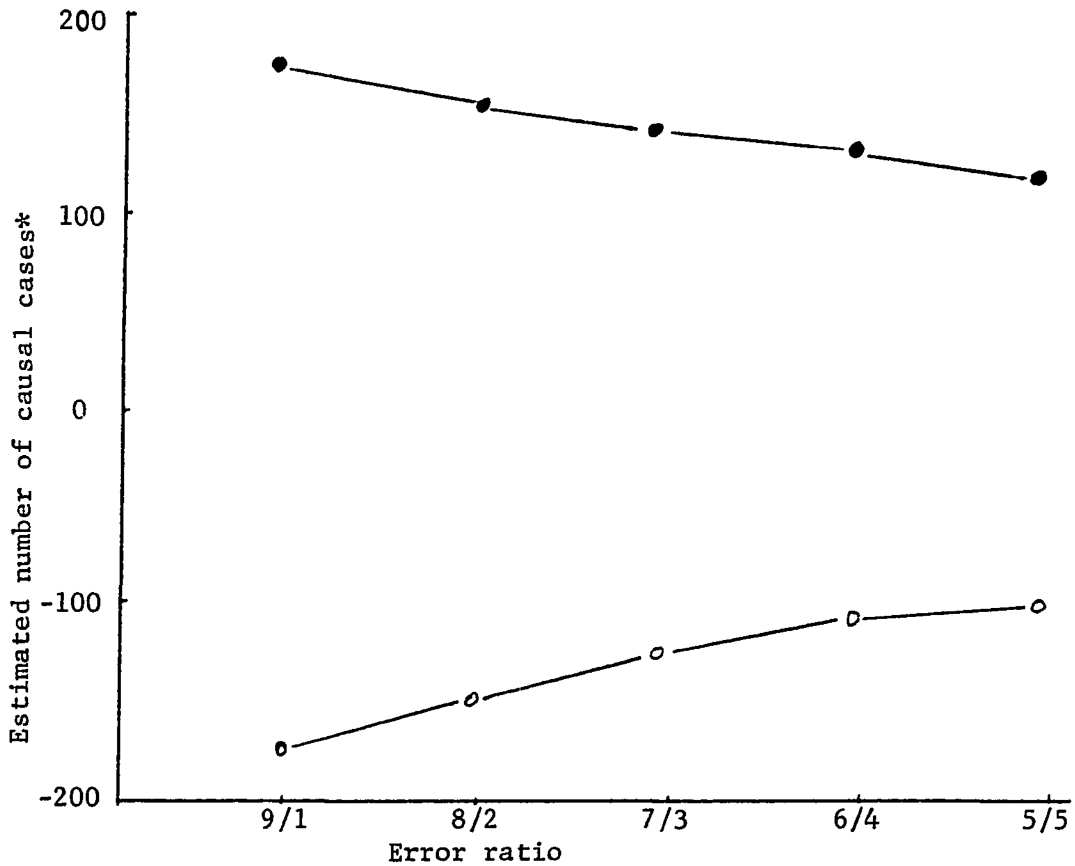


Figure 11. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model A2 (varied-influence, uncorrelated) data sets.

Influence coefficient:

Congruent	Incongruent
●	○

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

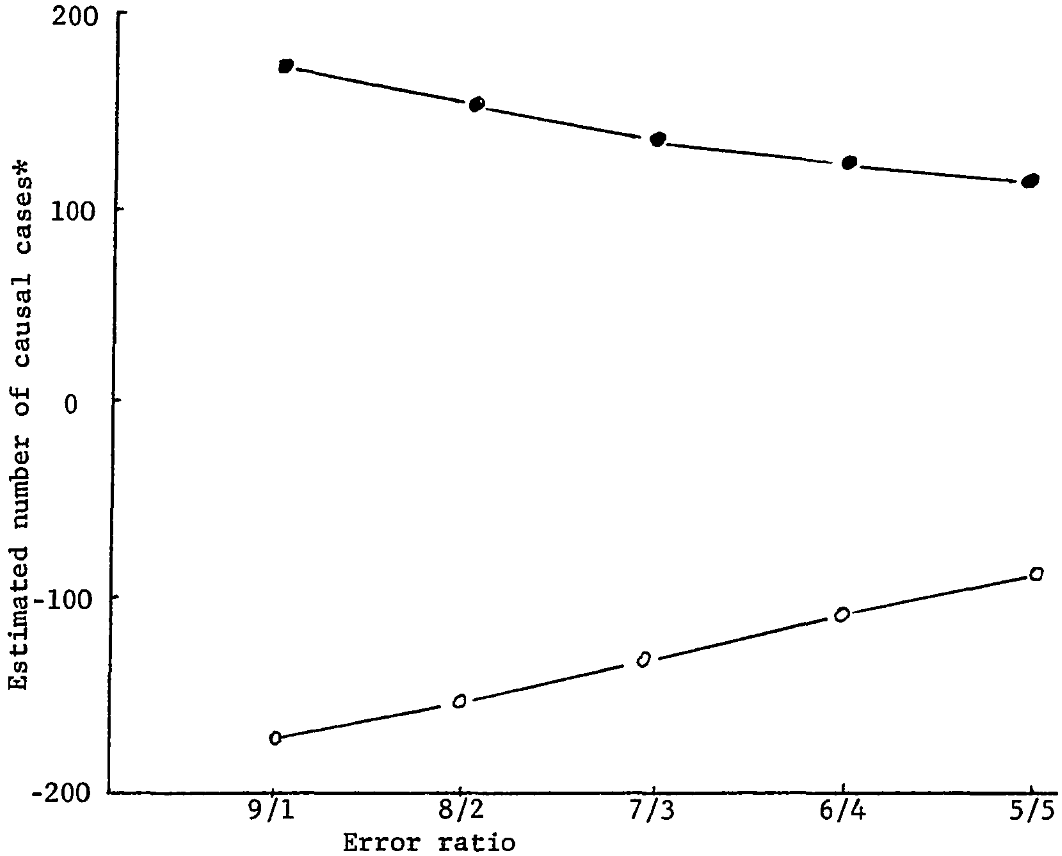


Figure 12. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model A2 (varied-influence, uncorrelated) data sets.

Influence coefficient:

Congruent	Incongruent
●	○

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence.

models.

At first glance, the highly positive results obtained for the causal models may seem unexpected. However, it will be recalled that most models are highly efficient with model A1 data sets when the causal influence is strong. The positive results obtained with model A2 data sets are no doubt partially attributable to the presence of cases of strong causal influence. One-third of the cases are generated with causal influence coefficient of 1.0 or -1.0.

MODEL A3 (Congruent-Incongruent-Influence, Uncorrelated Data Sets)

Data sets under model A3 are generated from the same data model as model A1 data sets. That is to say, time-one and time-two measures of variables X and Y are created by means of the following equations:

$$X_1 = a + e_1,$$
$$X_2 = a + f(b) + e_2,$$
$$Y_1 = b + e_3,$$
$$Y_2 = b + f(b) + e_4.$$

These data sets have been created in such a way that the time-one measures of variable X and variable Y are uncorrelated. While the amount of causal influence is uniform (i.e., the same) within each data set of 400 cases, half of the cases have been created with congruent causal influence and the other half of the cases have been created with incongruent causal influence. A total of 15 data sets are created, each of which differs from the other with respect to the amount of causal influence and the error ratio.

Model A3 data sets differ from model A1 data sets in only one respect:

every model A3 data set consists of two model A1 data sets with the same error ratio and the same amount of causal influence (of opposite causal directions.) For instance, a data set under model A3 may consist of a model A1 data set with error ratio of 9/1 and a causal influence coefficient of 1.0 and another model A1 data set with error ratio of 9/1 and a causal influence coefficient of -1.0.

The skewness and kurtosis coefficients computed for model A3 data sets are presented in Appendix C. These indices suggest that the distributions of the measures created by the simulation model may be regarded as approximately normal.

It will be noted at the outset that inasmuch as their coefficients can bear only one sign (i.e., either positive or negative), several causal models (i.e., CLC, PC, and EC) are not designed to detect both congruent and incongruent causal influence simultaneously. In other words, the best one could expect of these models is the efficiency of detecting the source of the causal influence. The models will inevitably yield ambiguous results with respect to the direction of causal influence.

The decision rules presented earlier are used in determining the estimation power of the causal models. As indicated earlier, with respect to FSM, FCP, and MFCP, the chi square value for YC > YI is expected to be non-significant. The rationale is that since one-half of the cases are congruent and the other one-half incongruent, an efficient model should be able to detect about an equal number of congruent and incongruent cases.

The following sections present results obtained for the various causal models when they are used to analyze model A3 data.

Cross-lagged correlation model (CLC)

The results indicate that CLC fails entirely to detect the source of the causal influence. All causal estimates are shown to be ambiguous. A closer examination reveals that the pertinent cross-lagged correlation (i.e., $r_{y_1x_2}$) not only fails to reach statistical significance but are also of extremely small magnitude (See Tables V and VI).

As the model is shown to have extremely low or near zero efficiency, the effect of data reliability and the amount of causal influence on estimation is not further investigated.

Part correlation model (PC)

The model also appears to fail to detect the source of the causal influence. All estimates are ambiguous. The magnitudes of the part correlations are extremely small: in about half of the cases the actual magnitude is almost zero (See Tables V and VI).

In view of the negative results, no further investigation is made of the effects of data reliability and the amount of influence on estimation.

Econometric model (EC)

Following CLC and PC, the model fails to identify the source of the causal influence. All estimates are ambiguous and the magnitudes of the beta weights are small (See Tables V and VI).

The overall negative results appear to warrant no further investigation of the effects of data reliability and the amount of causal influence on estimation.

Table V. Efficiency of Six Causal Models with Respect to Model A3 Data Sets (N=400)

Magnitude of Causal Influence	Error Ratio	Causal Model					
		CLC	PC	EC	FSM	FCP	MFCP
	9/1	±	±	±	+	+	±
	8/2	±	±	±	+	+	±
±1.0	7/3	±	±	±	+	+	+
	6/4	±	±	±	+	+	+
	5/5	±	±	±	+	+	+
	9/1	±	±	±	+	+	±
	8/2	±	±	±	+	+	±
± .50	7/3	±	±	±	+	+	±
	6/4	±	±	±	+	+	+
	5/5	±	±	±	+	+	+
	9/1	±	±	±	+	+	+
	8/2	±	±	±	+	+	+
± .25	7/3	±	±	±	+	+	+
	6/4	±	±	±	±	±	+
	5/5	±	±	±	+	±	+

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

Table VI. Magnitudes of Causal Estimates Yielded by Six Causal Models with Respect to Model A3 Data Sets (N = 400)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
CLC	± 1.0	-.02	-.02	-.03	-.04	-.05	$r_{Y_1X_2}$
	± .50	-.03	-.04	-.05	-.06	-.07	
	± .25	-.04	-.05	-.06	-.07	-.08	
PC	± 1.0	.00	-.00	-.01	-.02	-.04	$R_{Y_1(X_2X_1)}$
	± .50	.00	-.00	-.01	-.03	-.06	
	± .25	.00	-.00	-.02	-.05	-.07	
EC	± 1.0	-.04	-.05	-.06	-.07	-.08	$\hat{B}_{Y_1}$
	± .50	-.04	-.05	-.06	-.07	-.09	
	± .25	-.04	-.05	-.06	-.07	-.09	
FSM	± 1.0	89/72	76/73	75/68	69/70	64/70	YI/YC
	± .50	60/56	56/59	61/52	56/52	57/62	
	± .25	34/31	31/34	42/36	42/39	50/52	
FCP	± 1.0	179/168	154/161	145/146	135/126	119/121	YI/YC
	± .50	167/168	144/150	142/131	120/121	111/117	
	± .25	158/157	136/132	117/116	108/111	102/118	
MFCP	± 1.0	223/159	206/151	178/143	160/135	124/132	YI/YC
	± .50	221/156	197/146	175/138	131/133	105/119	
	± .25	203/169	170/154	132/130	119/115	110/114	

<u>Frequencies-of-shift-across-median model (FSM)</u>

The model appears to have a surprisingly high degree of efficiency in detecting the source and direction of causal influence in A3 data. All but one data sets are correctly estimated. The only ambiguous estimate is made with data of relatively low reliability and weak causal influence (See Table V).

A distinction should be made between the number of YC and YI cases actually estimated and the number of such cases correctly estimated. The former does not necessarily represent the number of individual YC and YI cases correctly estimated. In other words, the number does not take into account instances where YC is erroneously estimated as YI and vice versa. Many such reversions could occur in the estimation.

To determine the number of correct predictions (i.e., YC is estimated as YC and YI is estimated as YI) each individual case is examined. The results, as expected, show that the number of correct predictions is considerably smaller than the number of obtained predictions. The pattern of the results, or the relative frequencies of YC and YI cases, however, remain almost the same. Consequently, causal estimates based on correct predictions turn out to be the same as those derived from obtained predictions.

For both the correct and obtained causal estimates data reliability seems to have negligible effects on estimation power. The effects of the amount of causal influence, on the other hand, appear to be considerable and consistent (See Tables VII and VIII and Figures 13 and 14).

The above results thus seem to suggest that FSM is highly efficient in detecting the source and direction of causal influence with respect to model A3 data sets.

Table VII. Efficiency of FSM, FCP, and MFCP with Respect to Model A3 Data Sets (N = 400)[1]

Magnitude of Causal Influence	Error Ratio	Causal Model FSM	FCP	MFCP
	9/1	+	+	±
	8/2	+	+	±
± 1.0	7/3	+	+	+
	6/4	+	+	+
	5/5	+	+	+
	9/1	+	+	±
	8/2	+	+	+
± .50	7/3	+	+	+
	6/4	+	+	+
	5/5	+	+	+
	9/1	+	+	+
	8/2	+	+	+
± .25	7/3	+	+	+
	6/4	±	±	+
	5/5	+	±	+

[1] Based on correct predictions

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

Table VIII. Correct Individual Predictions Yielded by FSM, FCP, and MFCP with Respect to Model A3 Data Sets (N = 400)

Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Prediction
FSM	± 1.0	54/46	47/47	47/44	43/45	37/47	YI/YC
	± .50	40/38	39/40	10/35	36/38	34/43	
	± .25	20/19	20/23	27/25	25/25	28/31	
FCP	± 1.0	110/103	97/95	91/89	84/75	70/73	YI/YC
	± .50	100/101	91/88	89/80	75/70	64/71	
	± .25	95/93	87/79	71/69	61/66	58/68	
MFCP	± 1.0	121/86	110/81	101/80	90/77	70/75	YI/YC
	± .50	118/84	114/86	103/80	72/76	57/65	
	± .25	114/93	101/87	74/75	66/66	60/64	

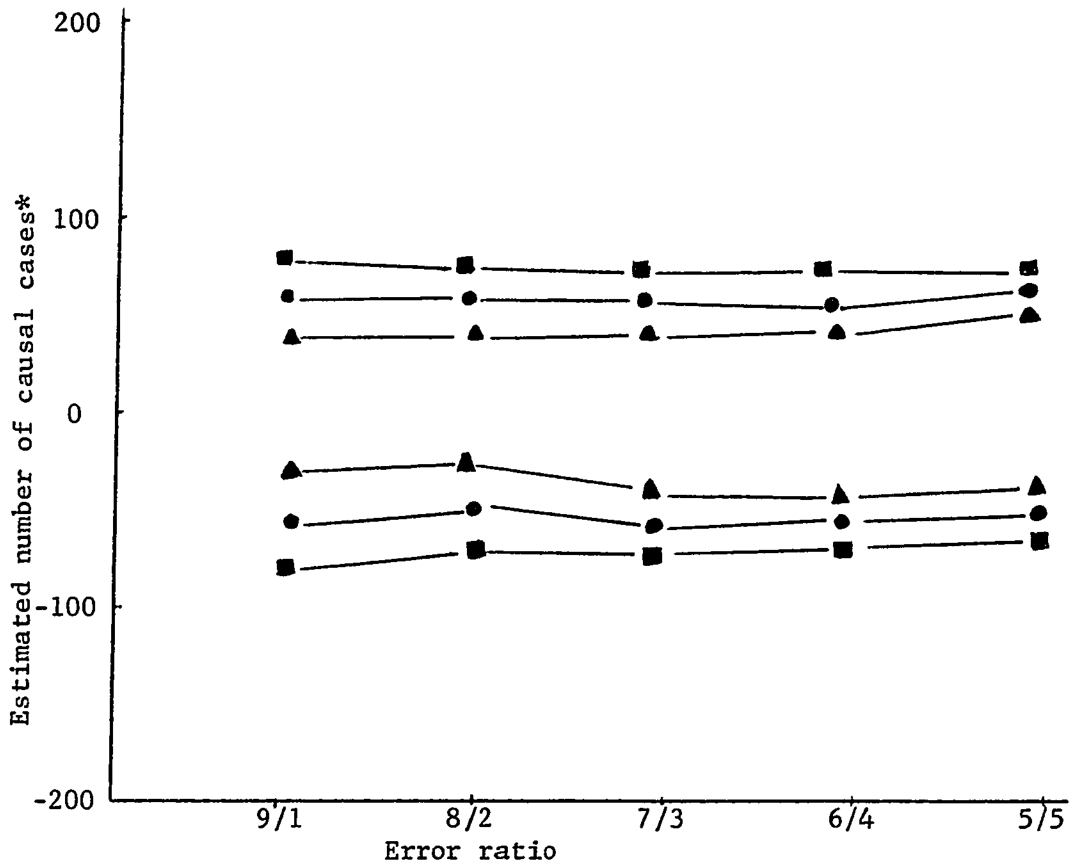


Figure 13. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model A3 (congruent-incongruent-influence, uncorrelated) data sets on the basis of obtained predictions.

Influence coefficient:

± 1.00	± .50	± .25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

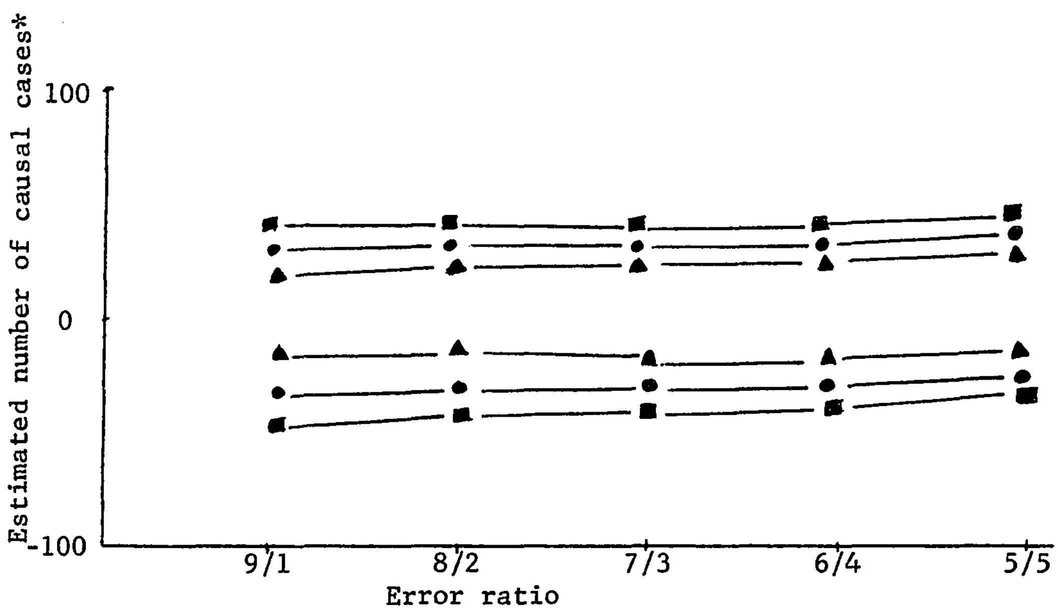


Figure 14. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model A3 (congruent-incongruent-influence, uncorrelated) data sets on the basis of correct predictions.

Influence coefficient:

± 1.00 ■ ± .50 ● ± .25 ▲

*indicates that the number of causal cases pertains to Y as the source of the causal influence

Frequencies-of-change-in-product-moment model (FCP)

The model appears to have a high level of utility for identifying the source and direction of causal influence of A3 data. All but two data sets are correctly estimated. The ambiguous estimates are made with data of low reliability and weak causal influence (See Table V).

Again, there appears to be a discrepancy between obtained predictions and correct predictions. In this case, the difference turns out to be quite remarkable. In general, less than two-thirds of the obtained predictions are correct predictions. The relative frequencies of YC and YI cases, however, remain the same. In other words, causal inference derived from correct predictions in this case are the same as those derived from obtained predictions (See Tables V to VIII).

For both the correct and obtained estimates, data reliability and the amount of causal influence seem to have mild effects on estimation. The effects of both factors are shown to be quite consistent (See Figures 15 and 16).

Modified-frequencies-of-change-in-product-moment model (MFCP)

At first glance, the results appear to be rather unusual. The model yields correct estimates for 10 of the 15 data sets. The five ambiguous estimates are, however, made with data of high reliability and strong causal influence. A closer examination of the results reveals that this appears to be an artifact attributable to the model's differential sensitivity toward congruent and incongruent influence. The model appears to be highly sensitive to incongruent influence and actually outperforms FSM and FCP in identifying YI cases (See Table VIII). Relative to congruent influence, the model appears to be much less sensitive.

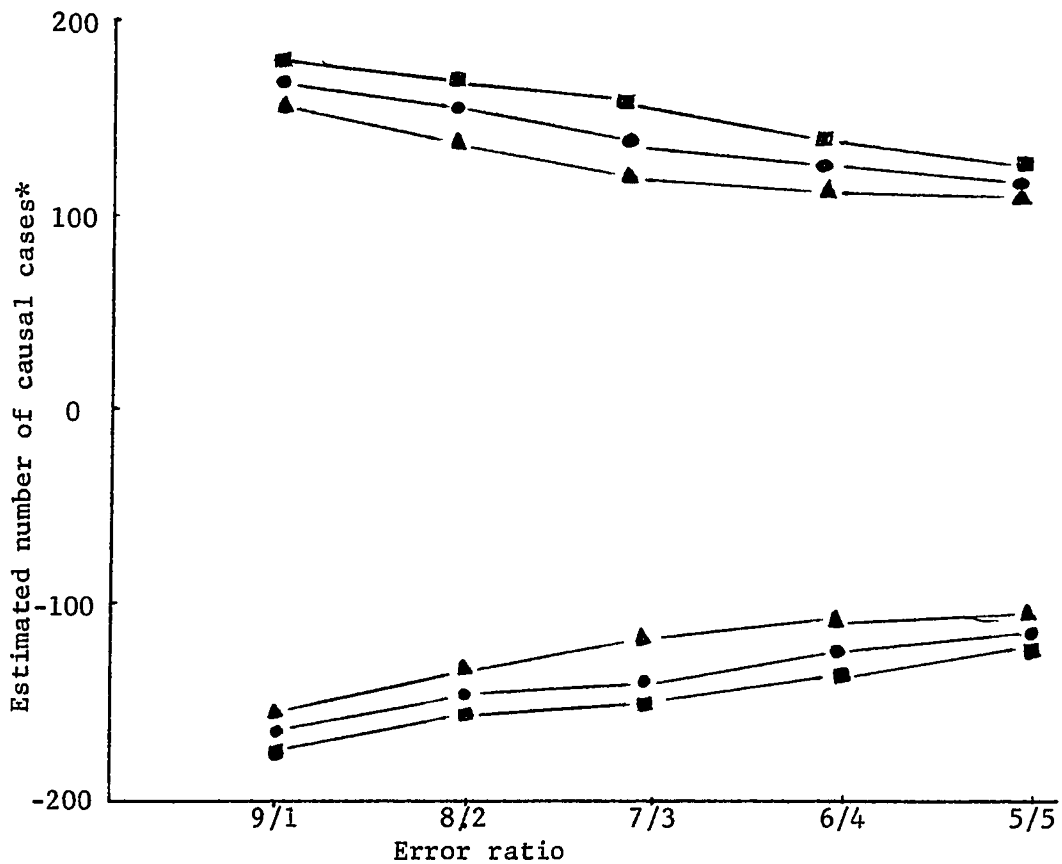


Figure 15. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model A3 (congruent-incongruent-influence, uncorrelated) data sets on the basis of obtained predictions.

Influence coefficient:

± 1.00 ■   ± .50 ●   ± .25 ▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

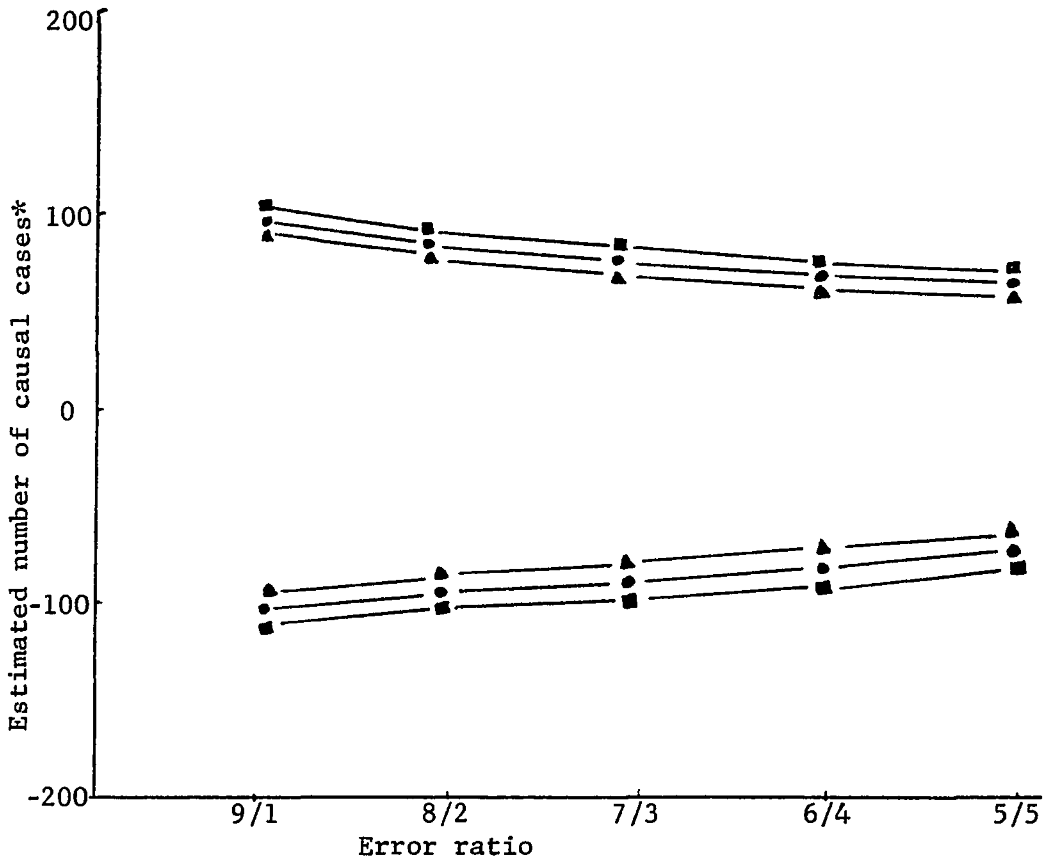


Figure 16. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model A3 (congruent-incongruent-influence, uncorrelated) data sets on the basis of correct predictions.

Influence coefficient:

$\pm$ 1.00	$\pm$ .50	$\pm$ .25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

Consequently, the estimates for YC and YI cases are highly disproportionate, leading to the conclusion that the incongruent influence dominates (See Tables V and VI).

In comparison with FSM and FCP, MFCP appears to create a more remarkable discrepancy between obtained predictions and correct predictions. Causal estimates based on correct predictions are shown to be slightly better than those based on obtained predictions: only three of the estimates based on correct predictions are ambiguous (See Tables V to VIII).

Data reliability and the amount of causal influence appear to have mild effects on estimation power. For both obtained predictions and correct predictions estimates, there seems to be some slight interaction between error and the amount of causal influence (See Figures 17 and 18).

Summary

The results presented in the preceding sections appear to indicate that when both congruent and incongruent influences are present in the data, some causal models (i.e., CLC, PC, and EC) are entirely ineffective in detecting the source and direction of the causal influence. Other models (i.e., FSM, FCP, MFCP) seem to have some promise of being the appropriate procedure for causal analysis. The efficiency of MFCP is however obscured by its differential sensitivity toward congruent and incongruent causal influence.

While the discrepancies between obtained and correct predictions do not seem to be serious enough to render the causal estimates invalid, they clearly suggest that the efficiency of FSM, FCP, and MFCP are highly exaggerated as far as the individual causal cases are concerned. That is to say, the degree of efficiency represented by the obtained predictions

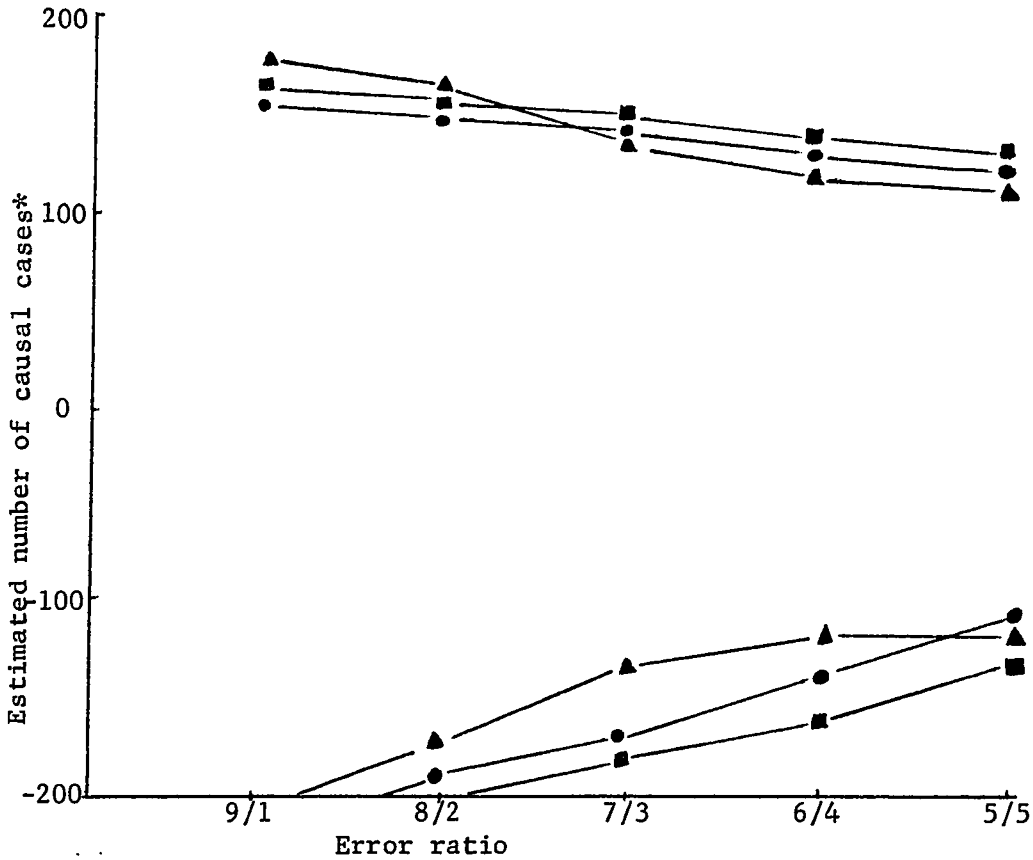


Figure 17. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model A3 (congruent-incongruent-influence, uncorrelated) data sets on the basis of obtained predictions.

Influence coefficient:

± 1.00	± .50	± .25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

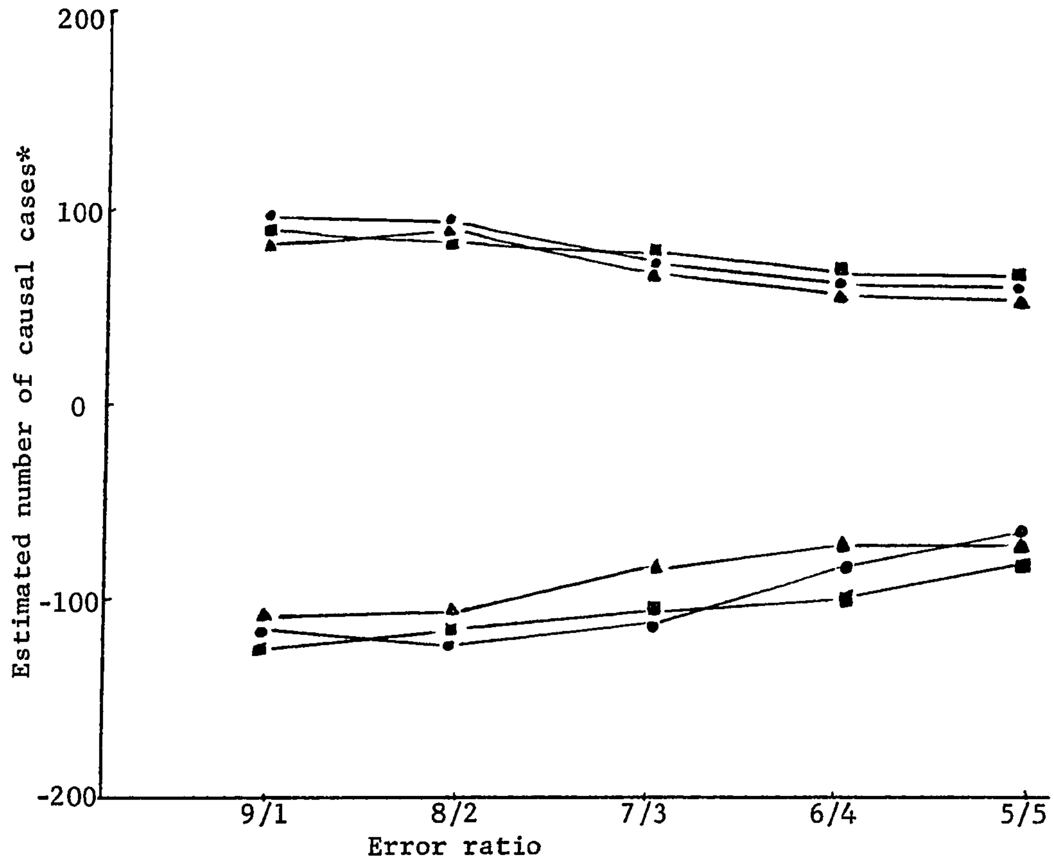


Figure 18. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model A3 (congruent-incongruent-influence, uncorrelated) data sets on the basis of correct predictions.

Influence coefficient:

± 1.00	± .50	± .25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

estimates (which, incidentally, are the only estimates the researcher is able to obtain if he is dealing with real life data) is considerably higher than the true efficiency level of the models. This is clearly reflected in the finding that, in general, less than two-thirds of the obtained predictions are correct predictions (See Tables VI and VIII).

It will also be noted that the decision rules assign more weight to the pattern of results (i.e., the relative frequencies of YC and YI cases) than they do to the number of correct predictions. Thus, the superiority of FSM and FCP over MFCP stems from their capability to provide such patterns rather than from their efficiency in correctly identifying individual YC and YI cases. As a matter of fact, if the power to correctly identify YC and YI cases were used as the basis for determining efficiency, MFCP would appear to be superior to both FCP and FSM.

Low data reliability appears to lower the efficiency of FSM, FCP, and MFCP. The effect of low reliability, however, appears to be mild. In the case of MFCP, there seems to be a slight interaction between data reliability and the amount of causal influence. The latter appears to have moderate and consistent effects on estimation power. In general, such effects appear to be in the expected direction: a greater amount of influence usually results in a better estimate of the source and direction of such influence.

MODEL B1 (Uniform-Influence, Correlated Data Sets)

Recall that model B1 data sets are created by means of the following equations:

$$X_1 = a + b + e_1,$$

$$X_2 = a + b + f(c) + e_2,$$

$$Y_1 = b + c + e_3,$$

$$Y_2 = b + c + e_4.$$

These data sets have been created in such a way that the time-one measures of variable X and variable Y are correlated. The amount of causal influence as well as the direction of causal influence which Y has on X is uniform (i.e., the same) for all 200 simulated cases within each data set. A total of 30 data sets are created each of which differs from the other with respect to the amount and direction of causal influence as well as the error ratio.

Coefficients of skewness and kurtosis of the data sets are presented in Appendix D. These indices appear to be sufficiently low to suggest that the measures created for the data sets have an approximately normal distribution. The following sections present results obtained for the various causal models when they are used to analyze model B1 data sets:

Cross-lagged correlation model (CLC)

The model yields only two correct causal estimates. Seven estimates turn out to be erroneous, leading to wrong conclusions with regard to the source or direction of the causal influence. The rest of the estimates are ambiguous. Unexpectedly, the correct estimates are made with data of low reliability (See Table IX). Data reliability appears to have moderate effects on estimation--especially with data of congruent causal

Table IX. Efficiency of Seven Causal Models with Respect to Model B1 Data Sets (N = 200).

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5
CLC	1.0	±	±	±	±	±
	.50	±	±	±	±	+
	.25	±	±	±	±	±
	- .25	±	±	±	—	±
	- .50	±	±	—	—	±
	-1.0	—	—	—	—	±
PC	1.0	+	+	+	+	+
	.50	+	+	+	+	+
	.25	+	+	+	+	+
	- .25	+	±	±	±	±
	- .50	+	+	±	±	±
	-1.0	+	+	+	+	±
EC	1.0	±	±	±	±	+
	.50	±	±	±	±	+
	.25	±	±	±	±	+
	- .25	±	±	±	±	±
	- .50	±	±	±	—	±
	-1.0	—	—	—	—	±
VC	1.0	+	+	+	+	+
	.50	+	+	+	+	+
	.25	±	±	±	±	±
	- .25	±	±	±	±	±
	- .50	±	±	±	±	±
	-1.0	±	±	±	±	±
FSM	1.0	+	+	±	±	±
	.50	+	±	±	±	±
	.25	±	±	±	±	±
	- .25	±	±	±	±	±
	- .50	+	±	±	±	±
	-1.0	+	+	±	±	±

Table IX. (Continued) Efficiency of Seven Causal Models with Respect to Model B1 Data Sets (N = 200)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5
FCP	1.0	+	+	±	±	±
	.50	+	±	±	±	±
	.25	+	±	±	±	±
	- .25	+	±	±	±	±
	- .50	+	±	±	±	±
	-1.0	+	+	±	±	±
MFCP	1.0	+	+	±	±	±
	.50	+	+	±	±	±
	.25	+	±	±	±	±
	- .25	+	+	+	±	±
	- .50	+	+	+	±	±
	-1.0	+	+	+	+	±

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

influence. The effects of the amount of influence seem to be negligible (See Table X and Figure 19).

The model is thus shown to be an extremely weak procedure for detecting causal influence with model B1 data. In some cases the model may actually lead to erroneous causal inferences.

Part correlation model (PC)

The model appears to be considerably more efficient than CLC in detecting the source and direction of causal influence. The results indicate that 22 of the 30 data sets are correctly estimated. The other eight estimates are ambiguous. It is of interest to note that all data sets of congruent causal influence are correctly estimated. The ambiguous estimates are made with data of incongruent influence and low reliability or weak causal influence (See Table IX).

Both data reliability and the amount of influence appear to have considerable effects on estimation power. The estimation of the source and direction of causal influence appears to be poorer as error increases or the amount of influence decreases (See Table X and Figure 20).

Econometric model (EC)

The model yields results which are highly similar to those obtained for CLC. Only three of the 30 data sets are correctly estimated. Five of the other estimates are in fact erroneous. Again, the correct estimates are made with data of low reliability. It may also be noted that all erroneous estimates are made with data of incongruent influence (See Table IX).

Data reliability appears to have moderate effects on estimation. The amount of influence is shown to have mild but consistent effects on

Table X. Magnitudes of Causal Estimates Yielded by Seven Causal Models with Respect to Model B1 Data Sets (N = 200)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
	1.0	.78	.74	.63	.47	.31	
	.50	.68	.63	.52	.38	.24	
CLC	.25	.59	.54	.45	.32	.19	$r_{Y_1X_2}$
	- .25	.35	.32	.26	.18	.10	
	- .50	.22	.20	.16	.11	.06	
	- 1.0	-.01	-.02	-.02	-.03	-.03	
	1.0	.66	.59	.50	.40	.28	
	.50	.59	.45	.37	.29	.20	
PC	.25	.44	.31	.26	.22	.16	$R_{Y_1(X_2X_1)}$
	- .25	-.37	-.12	.01	.06	.06	
	- .50	-.55	-.29	-.11	-.01	.02	
	- 1.0	-.64	-.49	-.30	-.15	-.07	
	1.0	.92	.85	.71	.52	.33	
	.50	.69	.63	.52	.38	.24	
EC	.25	.57	.52	.43	.31	.20	$\hat{B}_{Y_1}$
	- .25	.34	.30	.25	.17	.10	
	- .50	.22	.19	.15	.11	.06	
	- 1.0	-.01	-.02	-.03	-.03	-.03	
	1.0	34	29	22	21	18	
	.50	23	21	17	13	14	
FSM	.25	12	13	15	13	14	YC
	- .25	17	24	29	32	34	
	- .50	28	28	31	33	35	
	-1.0	42	41	38	38	39	YI
	1.0	105	77	64	52	49	
	.50	93	68	55	48	48	
FCP	.25	79	58	51	44	47	YC
	- .25	90	74	66	61	62	
	- .50	115	87	73	69	64	
	- 1.0	133	104	85	72	72	YI

Table X. (Continued) Magnitudes of Causal Estimates Yielded by Seven Causal Models with Respect to Model B1 Data Sets (N = 200)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
	1.0	108	86	70	62	54	
	.50	100	76	62	52	48	
MFCP	.25	77	56	52	46	48	YC
	- .25	90	76	73	69	61	
	- .50	106	89	76	74	63	
	- 1.0	126	100	88	80	65	YI
	1.0	2.02	2.02	2.00	1.95	1.84	
	.50	1.51	1.50	1.48	1.44	1.33	
VC	.25	1.25	1.24	1.22	1.18	1.08	$\frac{cov\ (Y_1X_2)}{cov\ (Y_1X_1)}$
	- .25	.74	.73	.70	.66	.58	
	- .50	.48	.47	.44	.40	.33	
	- 1.0	-.03	-.05	-.08	-.12	-.18	

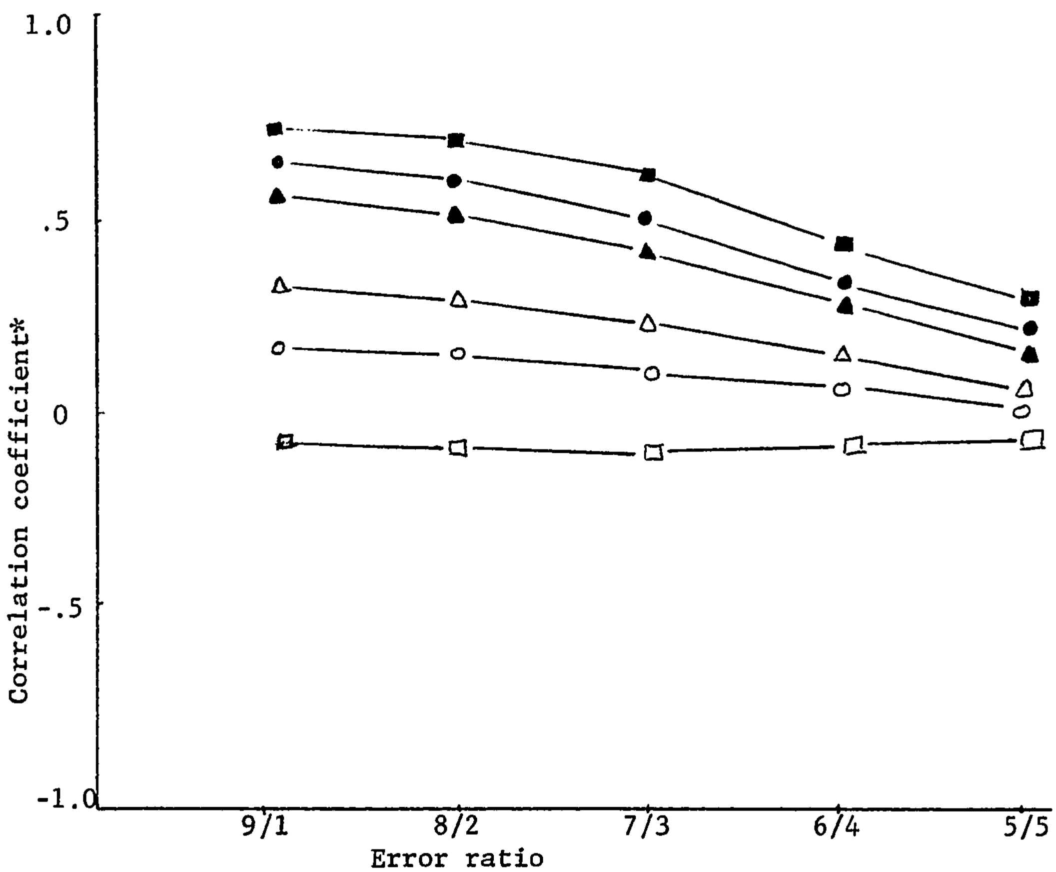


Figure 19. Effects of error and the amount of causal influence on the estimation efficiency of CLC with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the correlation coefficient pertains to Y as the source of the causal influence

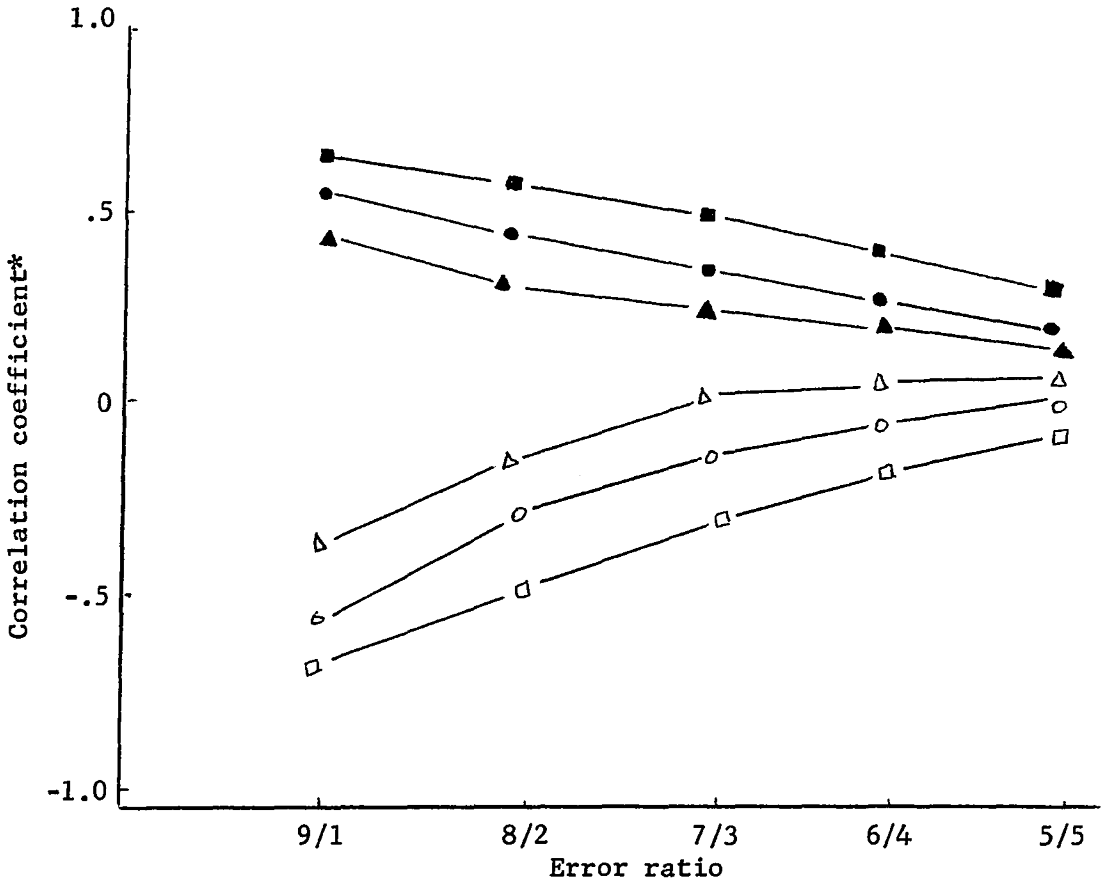


Figure 20. Effects of error and the amount of causal influence on the estimation efficiency of PC with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the correlation coefficient pertains to Y as the source of the causal influence

estimation power (See Table X and Figure 21).

Variance components model (VC)

The model yields a rather interesting pattern of results. All data sets of strong congruent influence are correctly estimated. Ambiguous results are obtained for the rest of the data sets (See Table IX).

It follows that data reliability has very little to do with the efficiency of estimation. The amount of influence, on the other hand, appears to have considerable effect on estimation power. As a matter of fact, a rather neat pattern of results is clearly discernible: there appears to be a consistent decrease in the F value as the influence coefficient declines from 1.0 to -1.0 (See Table X and Figure 22).

Frequencies-of-shift-across-median model (FSM)

The model appears to be an effective procedure for detecting causal influence only when data reliability is high and causal influence is strong. The results show that only six of the 30 data sets are correctly estimated. The other estimates are all ambiguous. All correctly estimated data sets are of high data reliability and strong causal influence (See Table IX).

In general, error appears to have mild effects on estimation--which is typical of FSM. The effects of the amount of causal influence are consistent but small in magnitude (See Table X and Figure 23).

Frequencies-of-change-in product-moment-model (FCP)

The model yields results which are highly similar to those obtained for FSM. In general, correct estimates are made with data of high reliability and strong causal influence. The model, however, appears to be

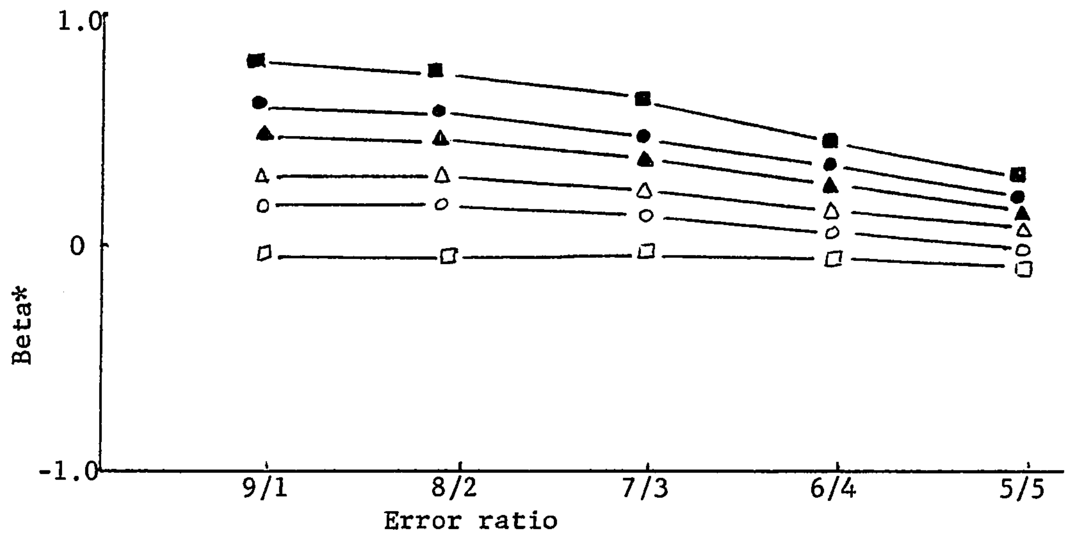


Figure 21. Effects of error and the amount of causal influence on the estimation efficiency of EC with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that beta pertains to Y as the source of the causal influence

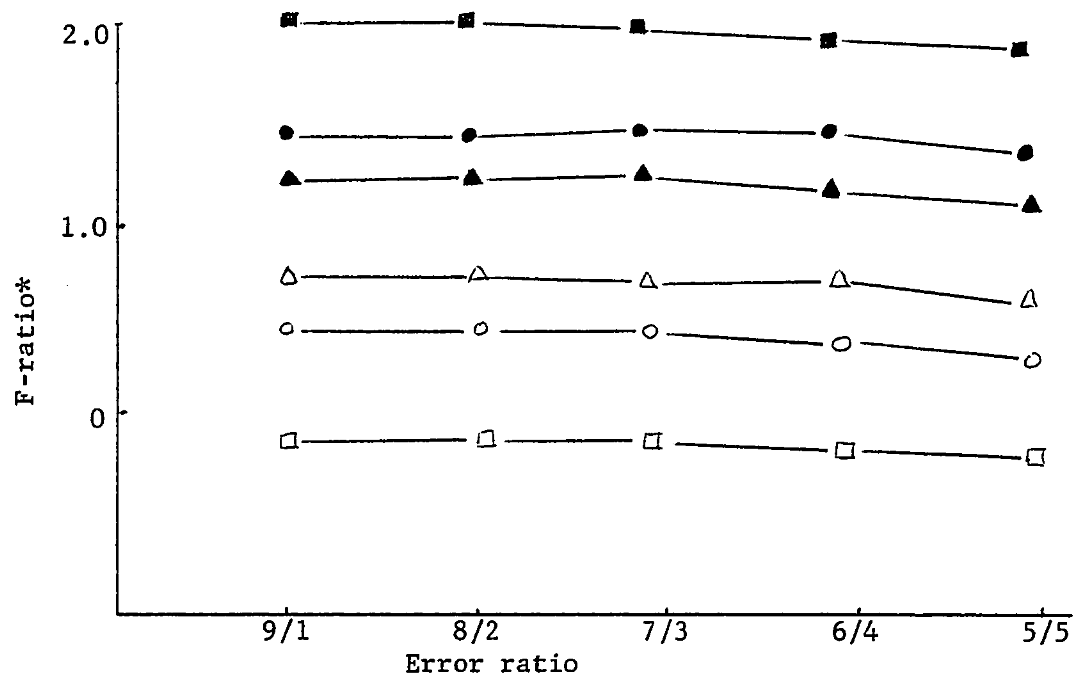


Figure 22. Effects of error and the amount of causal influence on the estimation efficiency of VC with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates F-ratio pertains to Y as the source of the causal influence

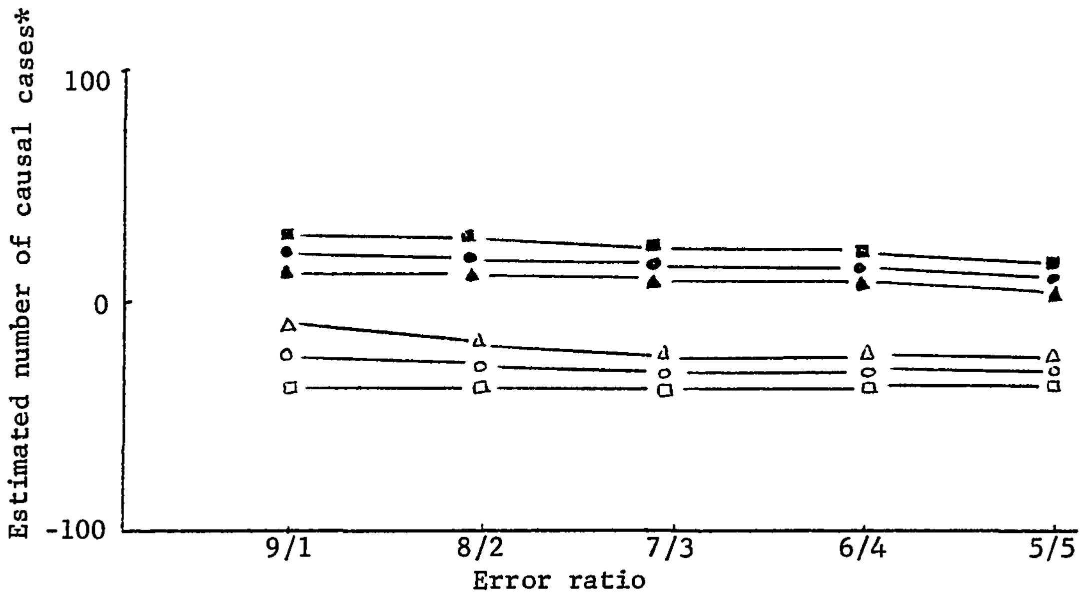


Figure 23. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

superior to FSM: when data reliability is very high, all causal estimates turn out to be correct. When data reliability is low and causal influence is weak, most of the causal estimates are ambiguous. No erroneous estimates are, however, made by the model (See Table IX).

The overall effects of data reliability and the amount of causal influence are in the expected direction. The efficiency of estimation drops quite considerably when the error ratio declines from 9/1 to 8/2. The amount of influence has consistent but small effects on estimation power (See Table X and Figure 24).

Modified-frequencies-of-change-in-product-moment model (MFCP)

The model is shown to be superior to both FCP and FSM. The results show that half of the 30 data sets are correctly estimated. Ambiguous estimates are made with data of low reliability and weak causal influence. It is interesting to note that more incongruent data sets are correctly estimated than congruent data sets. This result appears to be consistent with the fact that the model is more sensitive to incongruent influence than it is to congruent influence (See Table IX).

The effects of data reliability appear to be most pronounced when the error ratio drops from 9/1 to 8/2. Thereafter, the effects can be regarded as mild. The effects of the amount of influence are consistent but small in magnitude (See Table X and Figure 25).

Summary

CLC and EC are shown to be almost entirely ineffective in detecting the source and direction of causal influence with model B1 data sets. In some cases, results yielded by the two models may lead us to faulty causal inferences. FSM, FCP and MFCP appear to be effective only with

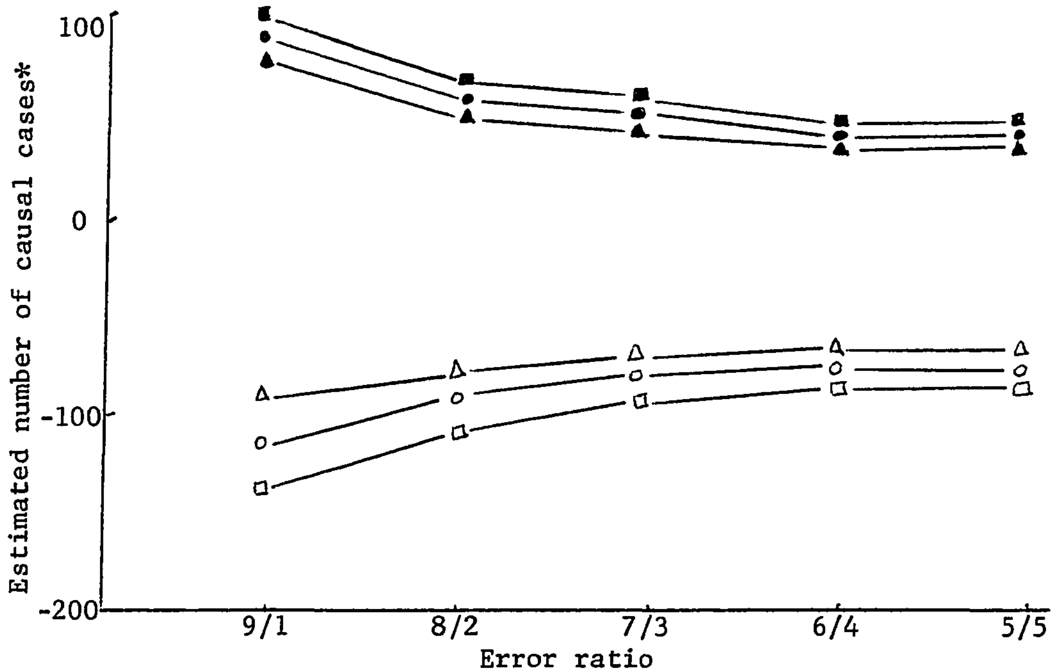


Figure 24. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0
■	●	▲	△	○	□

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

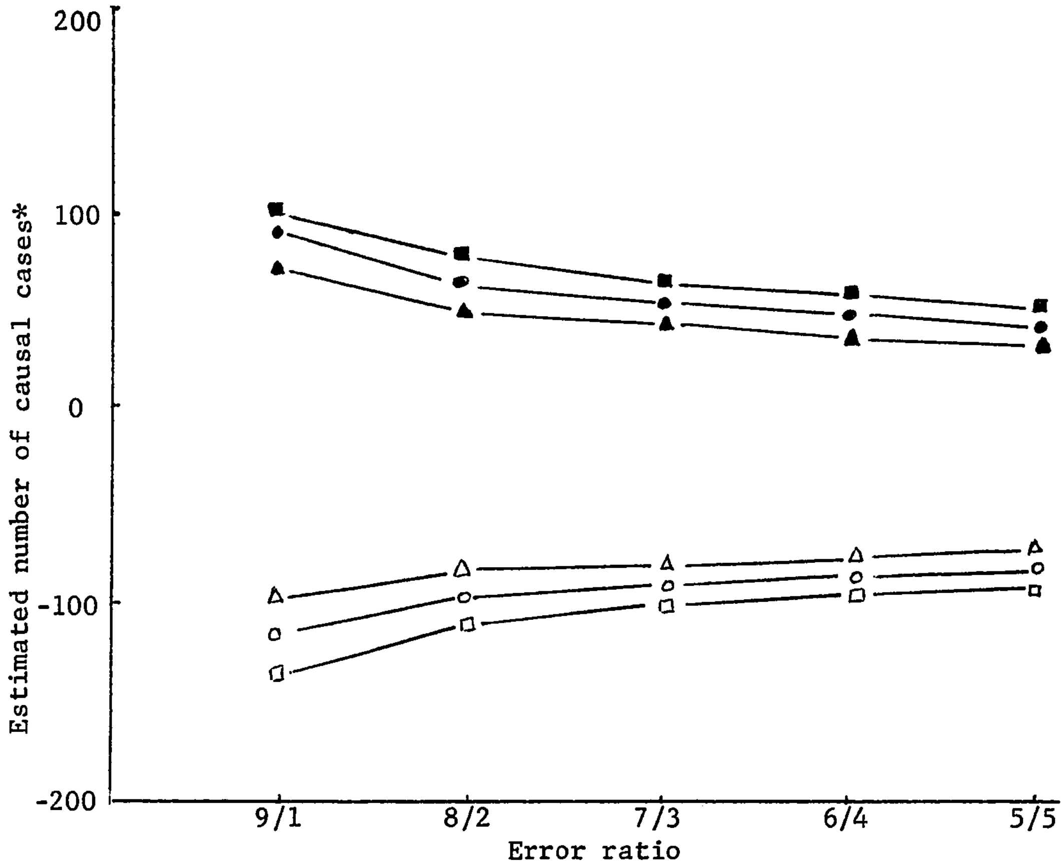


Figure 25. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model B1 (uniform-influence, correlated) data sets.

Influence coefficient:

1.0	.50	.25	-.25	-.50	-1.0

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

highly reliable data or data of strong causal influence. VC is shown to be effective in detecting causal influence with data of at least moderate causal influence. The most versatile model appears to be PC, which yields correct estimates for all the congruent data sets. The model, however, fails to correctly identify the source or direction of causal influence with data of weak incongruent influence and low reliability.

For all causal models data reliability does not seem to be a crucial factor with respect to the efficiency of estimation. Among the seven causal models, error appears to have the most noticeable effects on the efficiency of PC. For FCP and MFCP, the effects of data reliability are more discernible when error ratio declines from 9/1 to 8/2. Error appears to have practically no effects on the estimation power of FSM and VC. The effects of the amount of causal influence appear to be in the expected direction for all models. That is to say, a stronger influence results in a better causal estimate. The difference in estimation attributable to the amount of influence is highly consistent. The magnitude of the difference is, however, generally small.

Again, it should be pointed out that the estimation efficiency of FSM, FCP, and MFCP depends on the relative frequencies of YC, YI, XC, and XI cases rather than on the number of correct predictions. This is particularly true with respect to FSM, which as a rule correctly identifies a very small proportion (i.e., about 10 to 20 percent) of YC and YI cases.

MODEL B2 (Varied-Influence, Correlated Data Set)

Model B2 data sets are created from the same data model as model B1 data sets. Specifically, the measures are derived from the following equations:

$$X_1 = a + b + e_1,$$

$$X_2 = a + b + f(c) + e_2,$$

$$Y_1 = b + c + e_3,$$

$$Y_2 = b + c + e_4.$$

These data sets have been created in such a way that the time-one measures of variable X and variable Y are correlated. While the direction of the causal influence is uniform (i.e., either congruent or incongruent), the amount of causal influence which Y has on X is varied within each data set of 300 simulated cases. Specifically, one third of the cases have been created with an influence coefficient of 1.0 or -1.0, another third of the cases have been created with an influence coefficient of .50 or -.50, and the last third of the cases have been created with an influence coefficient of .25 or -.25. A total of 10 data sets are created each of which differs from the other with respect to the direction of causal influence and the error ratio.

Coefficient of skewness and kurtosis for the model B2 data sets are tabulated in Appendix E. These indices appear to suggest that the measures created can be regarded as having approximately normal distributions. The following sections present results obtained for the various causal models when they are used to analyze the model B2 data sets.

Cross-lagged correlation model (CLC)

The model appears to be almost entirely ineffective in detecting causal influence. The results show that only one of the 10 data sets is correctly estimated. Three of the other estimates turn out to be erroneous, the rest being ambiguous. All erroneous estimates are made with data of incongruent influence (See Table XI).

Table XI. Efficiency of Seven Causal Models with Respect to Model B2 Data Sets (N = 300)

Nature of Causal Influence	Error Ratio	Causal Model CLC	PC	EC	VC	FSM	FCP	MFCP
	9/1	±	±	±	+	+	+	±
	8/2	±	±	±	+	+	+	±
Congruent	7/3	±	±	±	+	+	±	±
	6/4	+	+	+	+	+	+	±
	5/5	±	+	+	+	±	+	±
	9/1	−	±	−	±	+	+	+
	8/2	−	±	−	±	+	+	+
Incongruent	7/3	−	±	±	±	±	+	+
	6/4	±	+	+	+	±	±	±
	5/5	±	+	+	+	±	±	±

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

While data reliability is shown to have a consistently negative effect on the magnitudes of estimates, it is of interest to note that the only correct causal estimate is made with a data set of low reliability (See Table XII and Figure 26).

Part correlation model (PC)

The model yields a rather unusual pattern of results. Only four of the 10 data sets are correctly estimated, the rest of the estimates being ambiguous. Most unexpectedly, all correct estimates are made with data of low reliability. This result may be due to one or both of the following factors: (1) The effects of error are more pronounced for variable X than for variable Y as the source of causal influence. (2) The model may be incompatible with model B2 data and the presence of relatively large errors serves to obscure or reduce this incompatibility (See Table XI).

Data reliability appears to have a consistently negative effect on estimation power. This result, incidentally, tends to rule out the second factor mentioned above as a plausible explanation for the unusual estimates yielded by the model (See Table XII and Figure 27).

Econometric model (EC)

This model yields results which are highly similar to those obtained for PC. The same four low reliability data sets are correctly estimated. Two of the other estimates are erroneous, the rest being ambiguous. The erroneous estimates are made with data of incongruent influence and low reliability. The result pertaining to correct causal estimates for low reliability data may again be due to differential effects of error on

Table XII. Magnitudes of Causal Estimates Yielded by Seven Causal Models with Respect to Model B2 Data Sets (N = 300)

Causal Model	Direction of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
CLC	+	.53	.47	.38	.25	.12	$r_{Y_1X_2}$
	-	-.05	-.09	-.13	-.16	-.18	
PC	+	.45	.39	.32	.23	.11	$R_{Y_1(X_2X_1)}$
	-	-.47	-.38	-.27	-.21	-.19	
EC	+	.52	.46	.35	.23	.11	$\hat{B}_{Y_1}$
	-	-.05	-.09	-.12	-.15	-.18	
VC	+	2.11	2.15	2.27	2.63	5.82	$\frac{\text{cov }(Y_1X_2)}{\text{cov }(Y_1X_1)}$
	-	-.22	-.42	-.80	-1.78	-9.59	
FSM	+	43	45	50	61	67	YC
	-	42	43	43	47	36	YI
FCP	+	128	113	102	107	113	YC
	-	157	126	101	86	73	YI
MFCP	+	132	108	91	96	91	YC
	-	163	129	106	92	83	YI

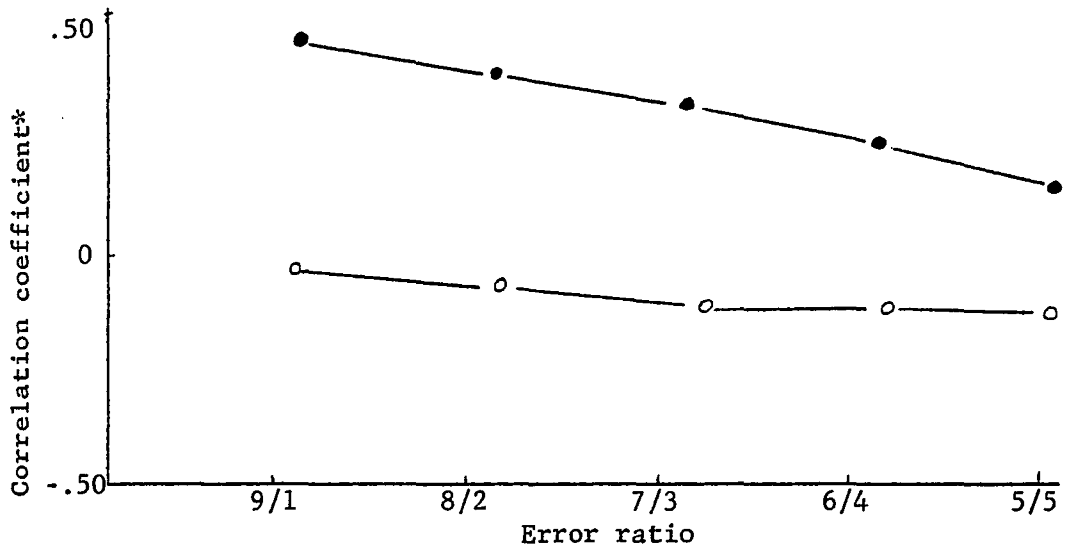


Figure 26. Effects of error and the amount of causal influence on the estimation efficiency of CLC with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent	Incongruent
●	○

*indicates that the correlation coefficient pertains to Y as the source of causal influence

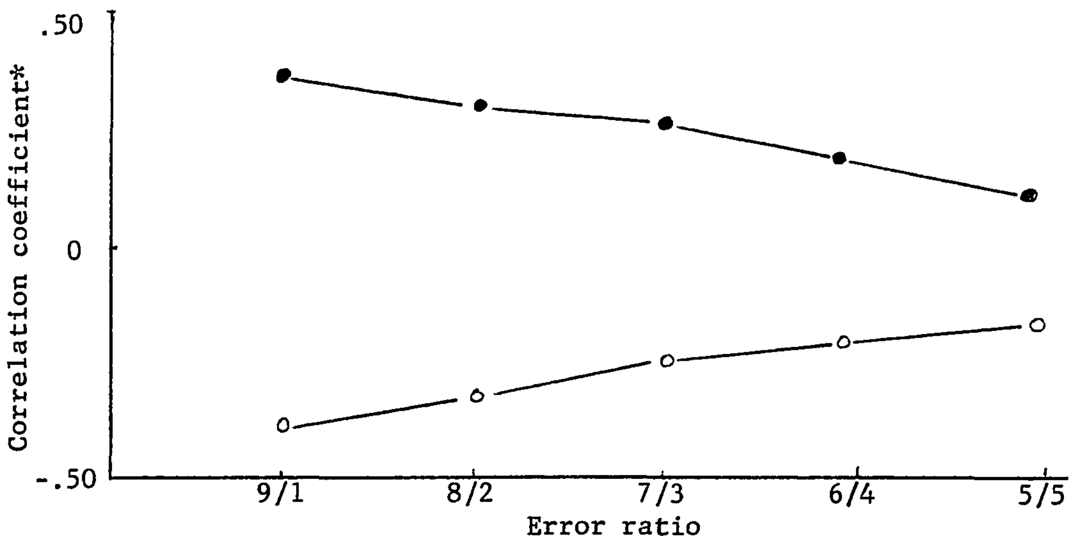


Figure 27. Effects of error and the amount of causal influence on the estimation efficiency of PC with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent ●

Incongruent ○

*indicates that the correlation coefficient pertains to Y as the source of the causal influence

variables X and Y (See Table XI). Data reliability appears to have mild effects on the estimation power of the model (See Table XII and Figure 28).

Variance components model (VC)

The model is shown to have a considerably higher level of efficiency than CLC, PC, and EC. All congruent data sets are correctly estimated. Two incongruent data sets are also correctly estimated, the other three being ambiguous. The correctly estimated incongruent data sets have low reliability, which again suggests that error may have more damaging effects on variable X than on variable Y (See Table XI).

Data reliability appears to have mild effects on estimation power until the error ratio reaches 6/4. Thereafter, the power of estimation appears to increase quite substantially. Thus, error is shown to have a positive effect on estimation--a result that places VC in marked contrast with other models (See Table XII and Figure 29).

Frequencies-of-shift-across-median model (FSM)

The model appears to have a moderately high efficiency in detecting causal influence with model B2 data sets. All but one of the congruent data sets are correctly estimated. The other estimates are ambiguous. The model is shown to be less efficient with incongruent data sets: only two estimates turn out to be correct, the other three being ambiguous. All correct estimates are made with data of high reliability (See Table XI).

Error appears to have mild effects on the power of estimation. It is interesting to note that for congruent data sets the effects of error tend to be positive rather than negative. That is to say, for congruent data sets, error actually appears to increase the power of estimation

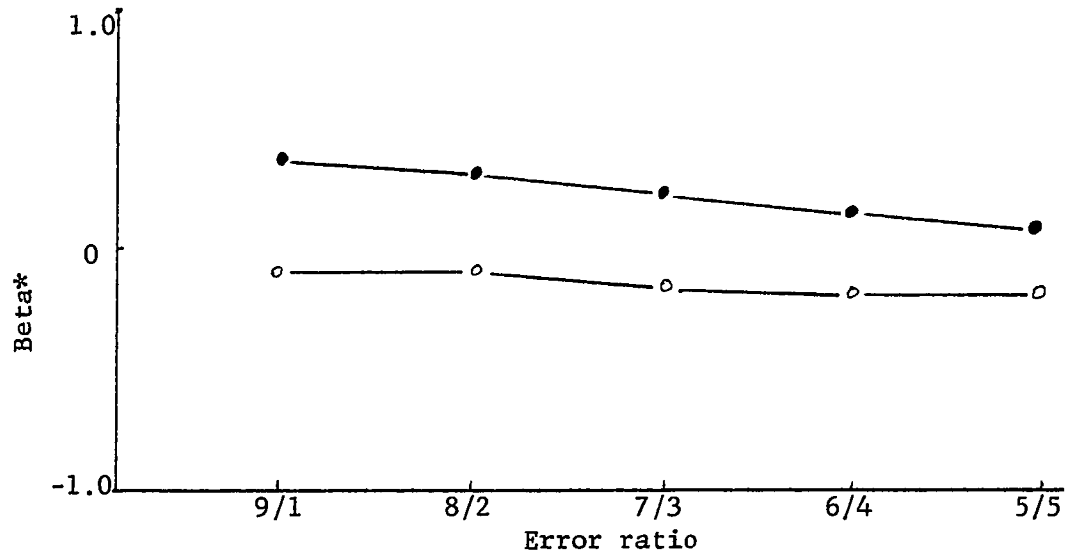


Figure 28. Effects of error and the amount of causal influence on the estimation efficiency of EC with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent ●

Incongruent ○

*indicates that beta pertains to Y as the source of the causal influence

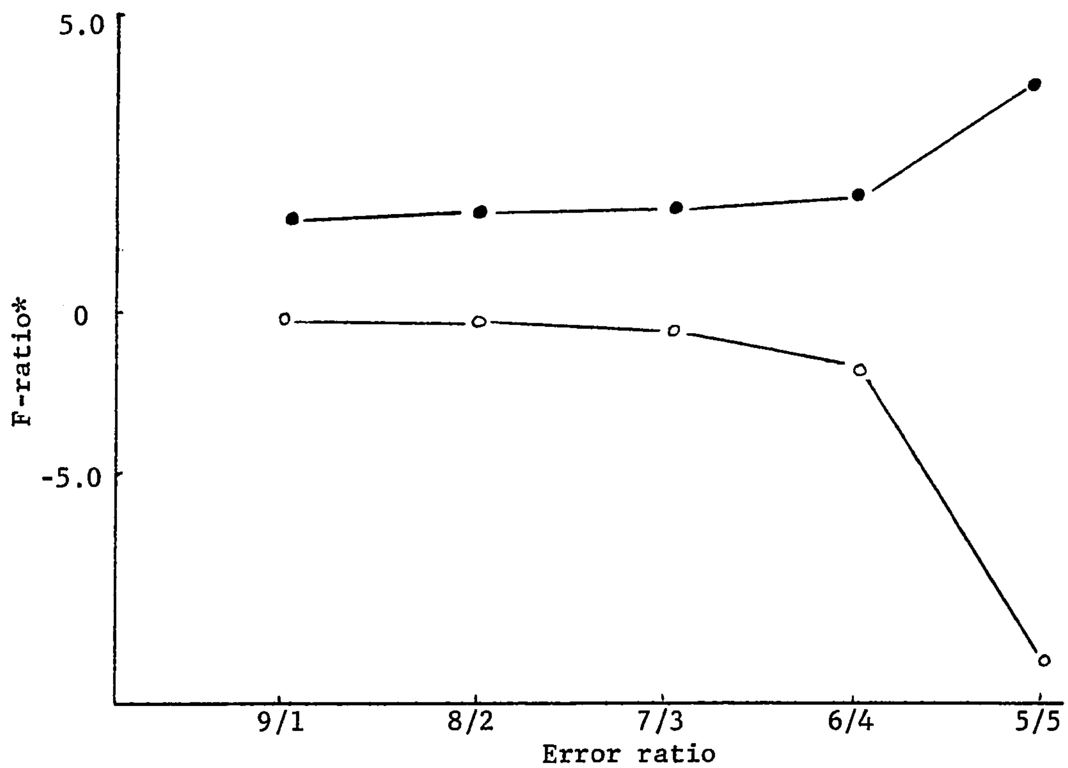


Figure 29. Effects of error and the amount of causal influence on the estimation efficiency of VC with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent	Incongruent
●	○

*indicates that the F-ratio pertains to Y as the source of the causal influence

(See Table XII and Figure 30).

Frequencies-of-change-in-product-moment (FCP)

The model yields results which are somewhat similar to those obtained for FSM. Specifically, all data sets of high reliability are correctly estimated. Moreover, FCP appears to be slightly more efficient with congruent data sets than with incongruent data sets. Only one ambiguous estimate is made with the congruent data sets. Two ambiguous estimates are made with incongruent data sets which also have low reliability (See Table XI).

In general, error seems to lower the efficiency of FCP. The effects appear to be more pronounced for incongruent data than for congruent data. (See Table XII and Figure 31).

Modified-frequencies-of-change-in-product-moment model (MFCP)

The model is shown to be entirely ineffective in detecting causal influence for the congruent data sets. All five estimates turn out to be ambiguous. The model, however, appears to have considerable efficiency in detecting causal influence with respect to the incongruent data sets. Three of the five incongruent data sets are correctly estimated. Two ambiguous estimates are made with data sets of low reliability (See Table XI).

Error seems to have moderate effects on estimation power. In addition, such effects appear to be more pronounced for the incongruent data sets than for the congruent data sets (See Table XII and Figure 32).

Summary

The results presented in the preceding sections suggest that CLC,

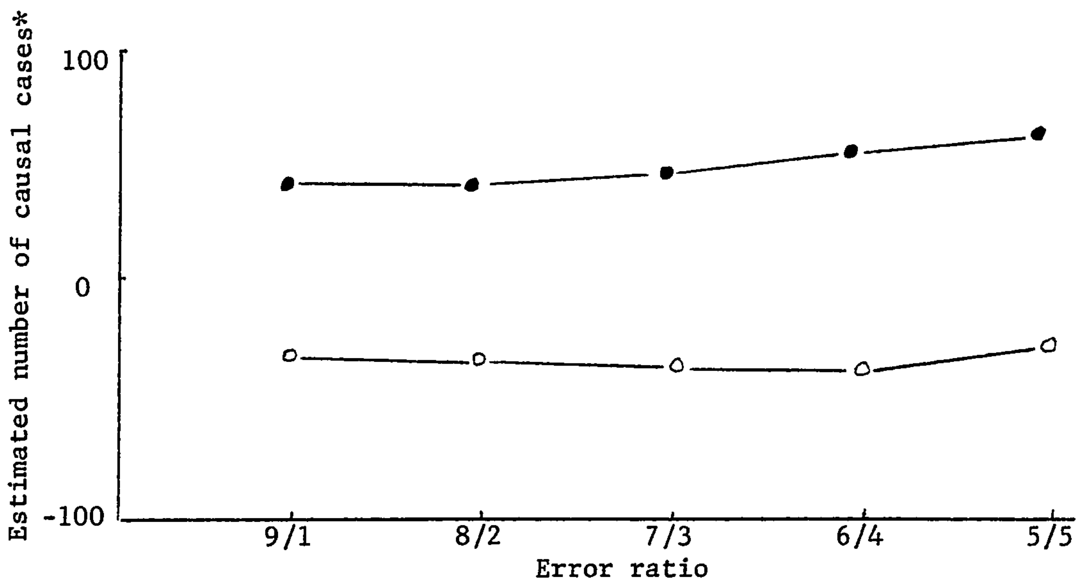


Figure 30. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent ● Incongruent ○

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

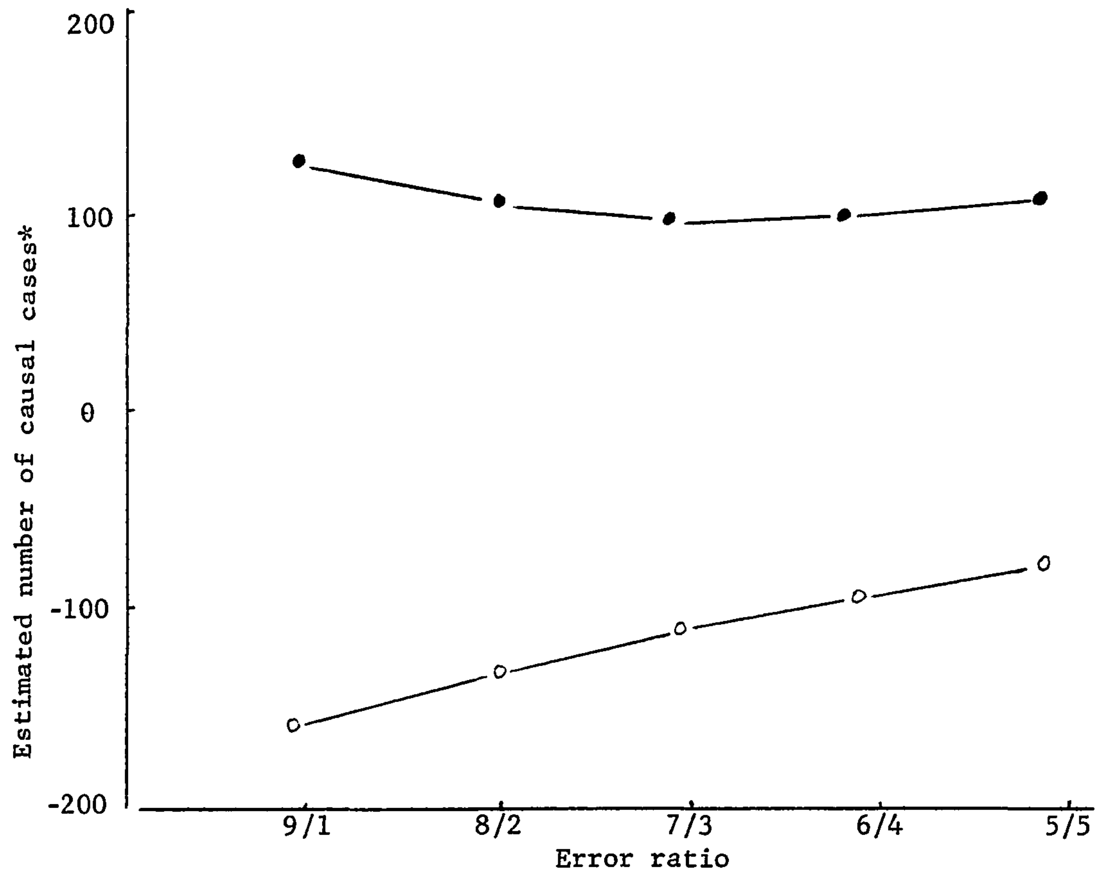


Figure 31. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent ●   Incongruent ○

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

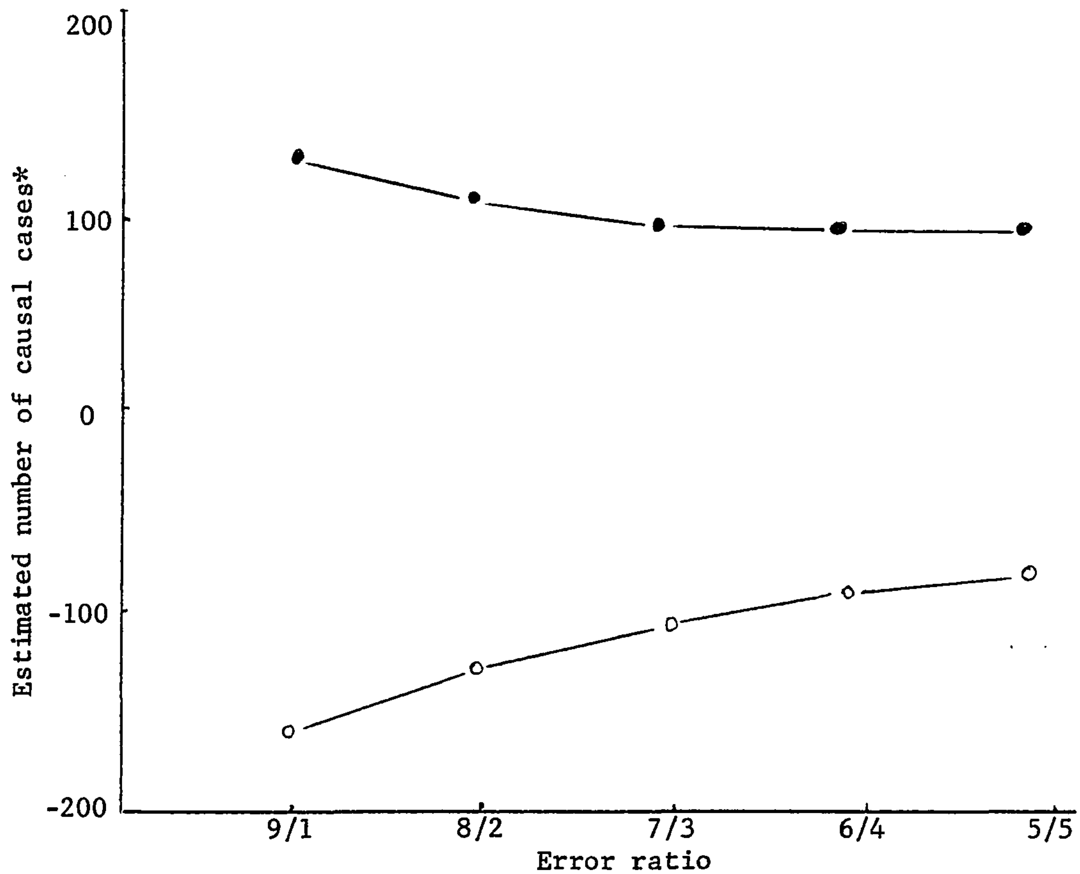


Figure 32. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model B2 (varied-influence, correlated) data sets.

Influence coefficient:

Congruent ●   Incongruent ○

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

PC, and EC have severely limited capability for detecting the source and direction of causal influence. Where the models yield correct causal estimates, the results appear to be partially attributable to the differential effects of error on variable X and variable Y. In some cases, results yielded by the models will actually lead us to erroneous causal inferences.

VC appears to be an effective procedure for detecting causal influence with congruent data sets. The model is, however, shown to have extremely low efficiency with incongruent data sets.

FSM and FCP seem to be satisfactory as procedures for causal analysis when data reliability is high. They are probably the best models to use with model B2 data. MFCP is shown to be able to identify incongruent causal influence when data reliability is high. The model appears to be ineffective with data of congruent influence.

Error poses a puzzling problem for model B2 data. For some causal models (i.e., CLC, EC, VC, and FSM) there appear to be instances where error actually serves to enhance the power of estimation. In other cases, error has the expected effects of decreasing estimation efficiency.

Again, it is to be pointed out that for FSM, FCP, and MFCP, the efficiency of the models is determined by the relative frequencies of YC, YI, XC, and XI cases rather than by the number of correctly identified causal cases. Thus, one model may have correctly identified more YC and YI cases in yielding an ambiguous estimate while another model may have provided a correct estimate on the basis of fewer correctly identified YC and YI cases. This clearly suggests that these models (i.e., FSM, FCP, and MFCP) should be used as group estimating procedures rather than models for predicting individual causal cases.

MODEL B3 (Congruent-Incongruent-Influence, Correlated Data Sets)

Model B3 data sets are created by means of the same data model as model B1 data sets. The time-one and time-two measures of variables X and Y are derived from the following equations:

$$X_1 = a + b + e_1,$$

$$X_2 = a + b + f(c) + e_2,$$

$$Y_1 = c + b + e_3,$$

$$Y_2 = c + b + e_4.$$

These data sets have been created in such a way that the time-one measures of variable X and variable Y are correlated. While the amount of causal influence in uniform (i.e., the same) within each data set of 400 cases, half of the cases have been created with congruent causal influence and the other half of the cases have been created with incongruent causal influence. A total of 15 data sets are created each of which differs from the other with respect to the amount of causal influence and the error ratio.

A crucial difference between model B3 and model B1 data sets lies in the fact that each of the former data sets actually comprises two of the latter data sets. In other words, each model B3 data set consists of two model B1 data sets of the same reliability level and same magnitude of influence--congruent and incongruent. Thus, each model B3 data set contains an equal number of congruent and incongruent cases with the same influence coefficient.

Coefficients of skewness and kurtosis for measures created for model B3 data sets are tabulated in Appendix F. These indices seem to be sufficiently low to suggest that the measures can be regarded as normally distributed.

It will be noted that since each of the model B3 data sets consists of congruent and incongruent causal influence, some causal models (i.e., CLC, PC, EC, and VC) will inevitably yield ambiguous estimates with respect to the direction of causal influence. None of these models can estimate both congruent and incongruent influence simultaneously. They can only be expected to determine the source of causal influence.

The following sections present results obtained for the various causal models when they are used to analyze model B3 data sets.

Cross-lagged correlation model (CLC)

The model appears to be entirely ineffective in detecting causal influence with model B3 data. The results show that no correct estimates are made with any of the data sets. One estimate is in fact erroneous, the rest being ambiguous (See Tables XIII and XIV).

In view of the overwhelmingly negative results, no further examination is conducted with respect to the effects of data reliability and the amount of causal influence on estimation power.

Part correlation model (PC)

The model yields a rather perplexing pattern of results. Only six of the 15 data sets are correctly estimated, the other nine estimates being ambiguous. Quite surprisingly, the correct estimates are made with data of low reliability and weak causal influence (See Tables XIII and XIV).

This unexpected result may partially be due to the fact that error appears to have positive effects on estimation efficiency. That is to say, as error becomes larger, the efficiency of the model increases and

Table XIII. Efficiency of Seven Causal Models with Respect to Model B3 Data Sets (N = 400)

Magnitude of Causal Influence	Error Ratio	Causal Model: VC	CLC	PC	EC	FSM	FCP	MFCP
1.0	9/1	±	±	±	±	±	+	±
	8/2	±	±	±	±	±	±	±
	7/3	±	±	±	±	±	±	±
	6/4	±	-	±	±	±	+	±
	5/5	±	±	±	±	±	±	±
.50	9/1	±	±	±	±	±	+	±
	8/2	±	±	±	±	+	+	±
	7/3	±	±	+	±	±	+	±
	6/4	±	±	+	±	±	±	±
	5/5	±	±	+	±	±	±	±
.25	9/1	±	±	±	±	+	+	±
	8/2	±	±	±	±	±	±	±
	7/3	±	±	+	±	±	±	±
	6/4	±	±	+	±	±	±	±
	5/5	±	±	+	±	±	±	±

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous estimate; and a combination of plus and minus indicates an ambiguous estimate.

Table XIV. Magnitudes of Causal Estimates Yielded by Seven Causal Models with Respect to Model B3 Data Sets (N = 400)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Type of Coefficient
CLC	± 1.0	.25	.23	.20	.16	.11	$r_{Y_1X_2}$
	± .50	.37	.34	.29	.21	.14	
	± .25	.44	.41	.33	.24	.15	
PC	± 1.0	.01	.03	.06	.09	.08	$R_{Y_1(X_2X_1)}$
	± .50	.01	.05	.10	.12	.10	
	± .25	.02	.08	.12	.14	.11	
EC	± 1.0	.45	.41	.34	.24	.15	$\hat{B}_{Y_1}$
	± .50	.45	.41	.34	.24	.15	
	± .25	.45	.41	.34	.24	.15	
VC	± 1.0	.99	.98	.96	.92	.83	$\frac{\text{cov }(Y_1X_2)}{\text{cov }(Y_1X_1)}$
	± .50	.99	.98	.96	.92	.83	
	± .25	.99	.98	.96	.92	.83	
FSM	± 1.0	95/59	90/48	79/38	72/40	63/32	YI/YC
	± .50	59/34	60/40	65/28	65/31	62/29	
	± .25	37/26	47/28	54/26	61/29	64/26	
FCP	± 1.0	181/148	155/115	141/102	129/100	125/94	YI/YC
	± .50	171/146	142/112	132/102	129/91	128/91	
	± .25	144/134	133/101	125/90	123/86	122/92	
MFCP	± 1.0	219/141	195/123	159/104	143/92	130/95	YI/YC
	± .50	203/148	171/111	140/102	132/96	126/93	
	± .25	176/134	140/99	135/97	129/89	119/96	

better causal estimates are obtained.

It will also be noted that the amount of causal influence appears to have similar effects on estimation: the weaker the causal influence the better the causal estimates. This result suggests that PC may be incompatible with model B3 data. Consequently, stronger causal influence and higher data reliability tend to aggravate this incompatibility while weaker causal influence and lower reliability serve to attenuate such incompatibility.

Econometric model (EC)

The model is shown to be entirely ineffective with model B3 data. All estimates yielded by the model turn out to be ambiguous. A closer look at the results reveals that the ambiguous estimates are due to the fact that all beta weights for variables X and Y are statistically non-significant (See Tables XIII and XIV).

In view of the overall negative results obtained for the model, no further analysis is made of the effects of error and the amount of causal influence on estimation power.

Variance components model (VC)

The general pattern of results yielded by the model is identical with that obtained for EC. All estimates turn out to be ambiguous. A closer examination of the results reveals that F_2 is in actuality consistently greater than F_1, as expected. Both F values are, however, statistically non-significant in all instances (See Tables XIII and XIV).

In view of the overall negative results obtained for the model, no further analysis is made of the effect of error and the amount of causal influence on estimation power.

Frequencies-of-shift-across-median model (FSM)

The model is shown to have little or no utility in detecting the source and direction of causal influence with model B3 data. The results show that only two of the 15 data sets are correctly estimated, the rest being ambiguous. Moreover, the two correct estimates do not seem to fall into any meaningful pattern relative to either data reliability or the amount of causal influence (See Tables XIII and XIV).

It will be recalled that when both congruent and incongruent influences are present in the same data set, a discrepancy exists between obtained predictions and correct predictions. As it turns out, when only the correct predictions are considered, the model appears to have a higher level of efficiency relative to results based on actual predictions: five of the six data sets of high reliability are correctly estimated (See Table XV).

Data reliability appears to have mild effects on the estimation power of the model. The effects of the amount of influence are shown to be slightly more pronounced. In some cases, there appears to be a slight interaction between error and the amount of influence (See Tables XIV and XVI; Figures 33 and 34).

Frequencies-of-change-in-product-moment model (FCP)

The model appears to be slightly more efficient than FSM in detecting causal influence with model B3 data. On the basis of obtained predictions, six of the 15 data sets are correctly estimated. Three of the six correct estimates are made with data sets of high reliability (See Table XIII).

The results based on correct predictions are indicative of a higher level of efficiency. Nine of the 15 data sets are correctly estimated.

Table XV. Efficiency of FSM, FCP, and MFCP with Respect to Model B3 Data Sets (N = 400)[1]

Magnitude of Causal Influence	Error Ratio	Causal Model		
		FSM	FCP	MFCP
1.0	9/1	+	+	±
	8/2	±	+	±
	7/3	±	+	±
	6/4	±	+	+
	5/5	±	±	+
.50	9/1	+	+	+
	8/2	+	+	±
	7/3	±	+	+
	6/4	±	±	+
	5/5	±	±	±
.25	9/1	+	+	+
	8/2	+	+	+
	7/3	±	±	+
	6/4	±	±	±
	5/5	±	±	±

[1]Based on correct predictions.

Note: A plus sign indicates a correct causal estimate; a minus sign indicates an erroneous causal estimate; and a combination of plus and minus indicates an ambiguous estimate.

Table XVI. Correct Individual Predictions Yielded by FSM, FCP, and MFCP with Respect to Model B3 Data Sets (N = 400)

Causal Model	Magnitude of Causal Influence	Error Ratio 9/1	8/2	7/3	6/4	5/5	Prediction
	± 1.0	51/33	51/27	45/24	41/24	36/22	
FSM	± .50	35/23	37/25	40/17	39/19	37/17	YI/YC
	± .25	21/17	27/18	32/16	35/16	34/14	
	± 1.0	99/85	86/68	80/60	70/52	66/51	
FCP	± .50	94/88	81/66	73/56	68/51	69/50	YI/YC
	± .25	80/81	70/58	69/49	67/47	64/49	
	± 1.0	114/74	107/70	83/56	76/54	71/56	
MFCP	± .50	104/79	92/59	73/58	69/56	68/57	YI/YC
	± .25	94/71	72/56	69/58	68/51	65/56	

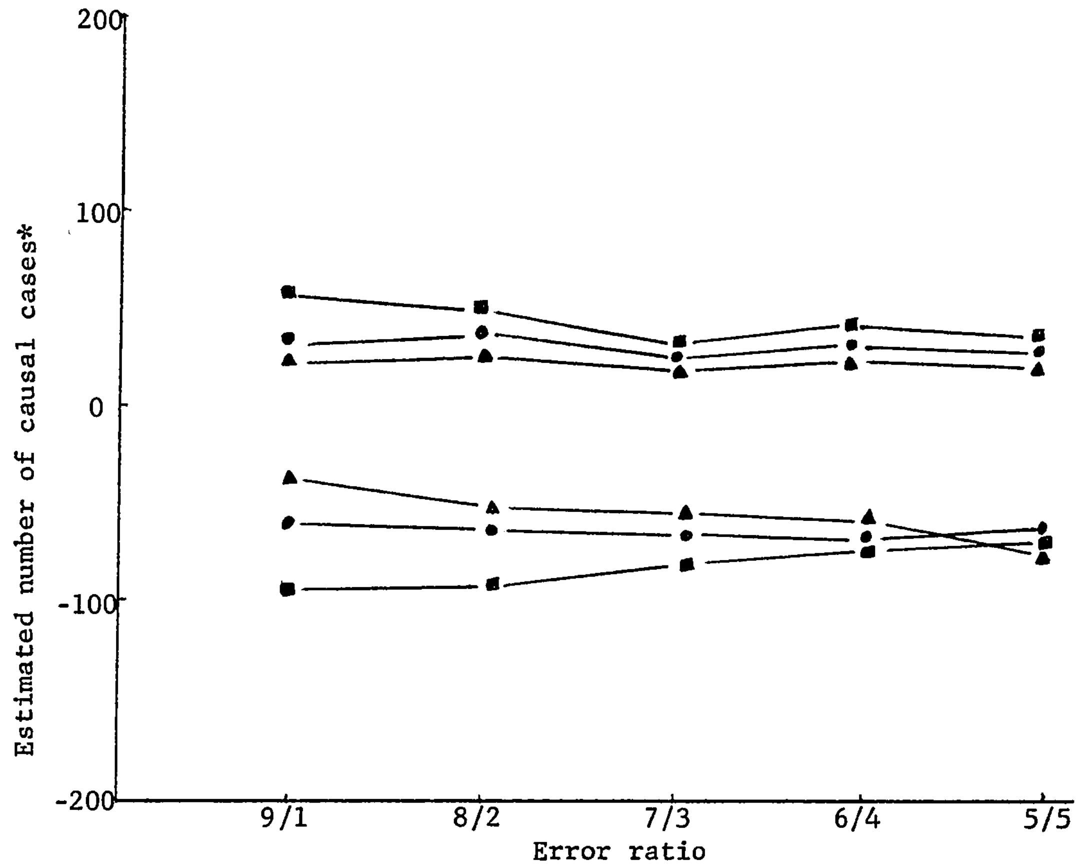


Figure 33. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model B3 (congruent-incongruent-influence, correlated) data sets on the basis of obtained predictions.

Influence coefficient:

± 1.0 ■ ± .50 ● ± .25 ▲

*indicates that the number of causal cases pertains to Y as the source of the causal influence

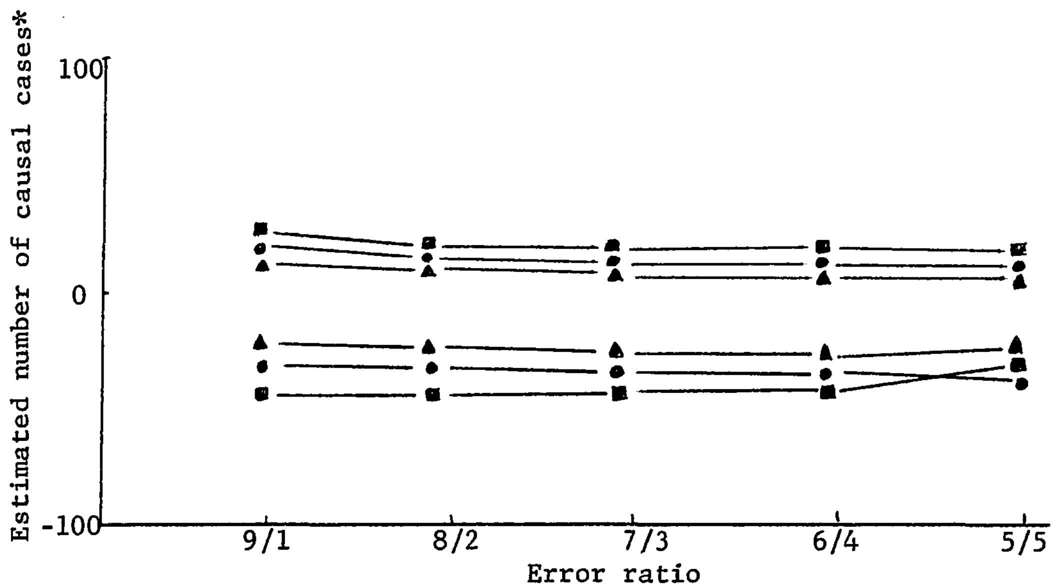


Figure 34. Effects of error and the amount of causal influence on the estimation efficiency of FSM with respect to model B3 (congruent-incongruent-influence, correlated) data sets on the basis of correct predictions.

Influence coefficient:

$\pm$ 1.0	$\pm$ .50	$\pm$ .25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

The correct estimates are made with data of high reliability or strong causal influence (See Table XV).

It should be noted again that a researcher who deals with real life data will not have the advantage of knowing the number of correct predictions. Causal inferences with regard to the source and direction of causal influence will have to be made on the basis of obtained predictions.

In general, the effects of error and the amount of influence on estimation power appear to be in the expected direction: error tends to lower the model's estimation efficiency and so does weak causal influence. In some instances, there appears to be some slight interaction between error and the amount of causal influence (See Tables XIV, XVI and Figures 35, 36).

Modified-frequencies-of-change-in-product-moment model (MFCP)

On the basis of obtained predictions, the model appears to be an entirely ineffective procedure for causal analysis with model B3 data. The results show that all estimates yielded by the model are ambiguous. A closer examination of the results reveals that most of the ambiguous estimates are derived as a result of the model's high sensitivity toward YI cases. In other words, the ambiguity pertains only to the direction of the causal influence. The source of the causal influence is in general correctly identified (See Table XIII).

Causal estimates derived from correct predictions indicate a higher level of efficiency relative to obtained predictions. The pattern of results is, however, perplexing: highly reliable data sets with weak causal influence are correctly estimated. So are data sets of low reliability but strong causal influence. It appears possible that the model's

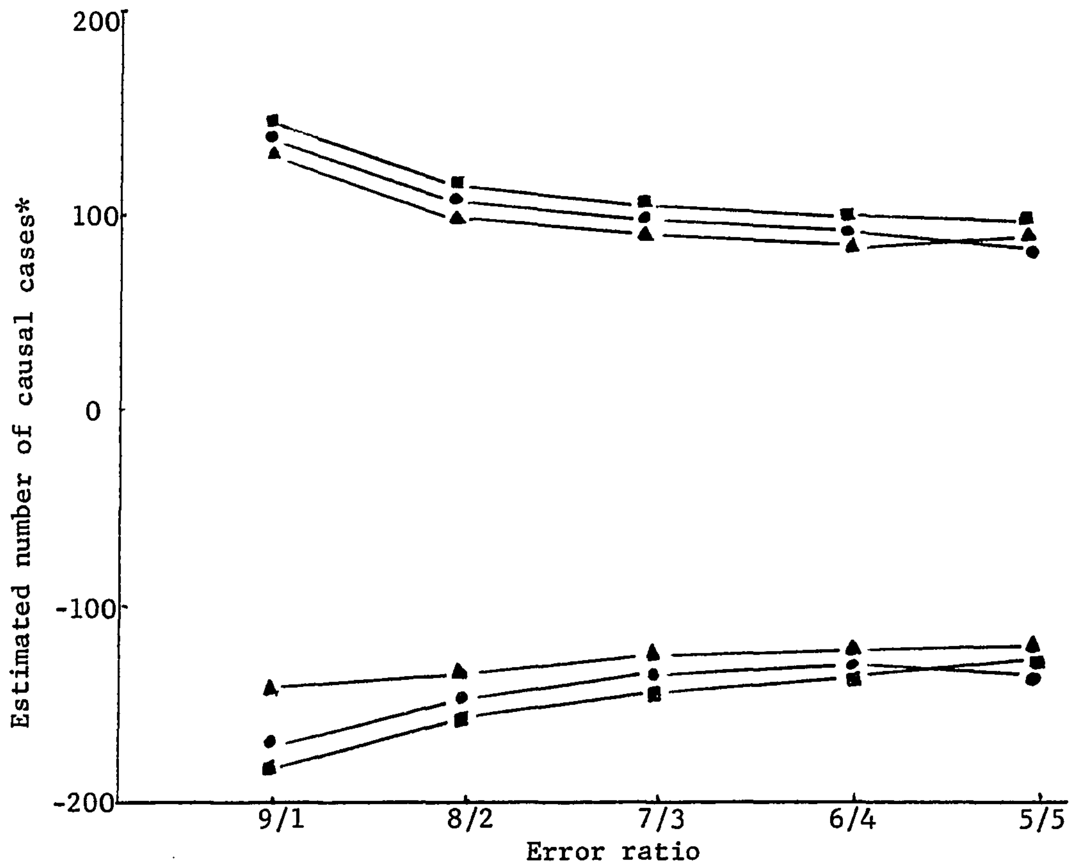


Figure 35. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model B3 (congruent-incongruent-influence, correlated) data sets on the basis of obtained predictions.

Influence coefficient:

± 1.0 ■   ± .50 ●   ± .25 ▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

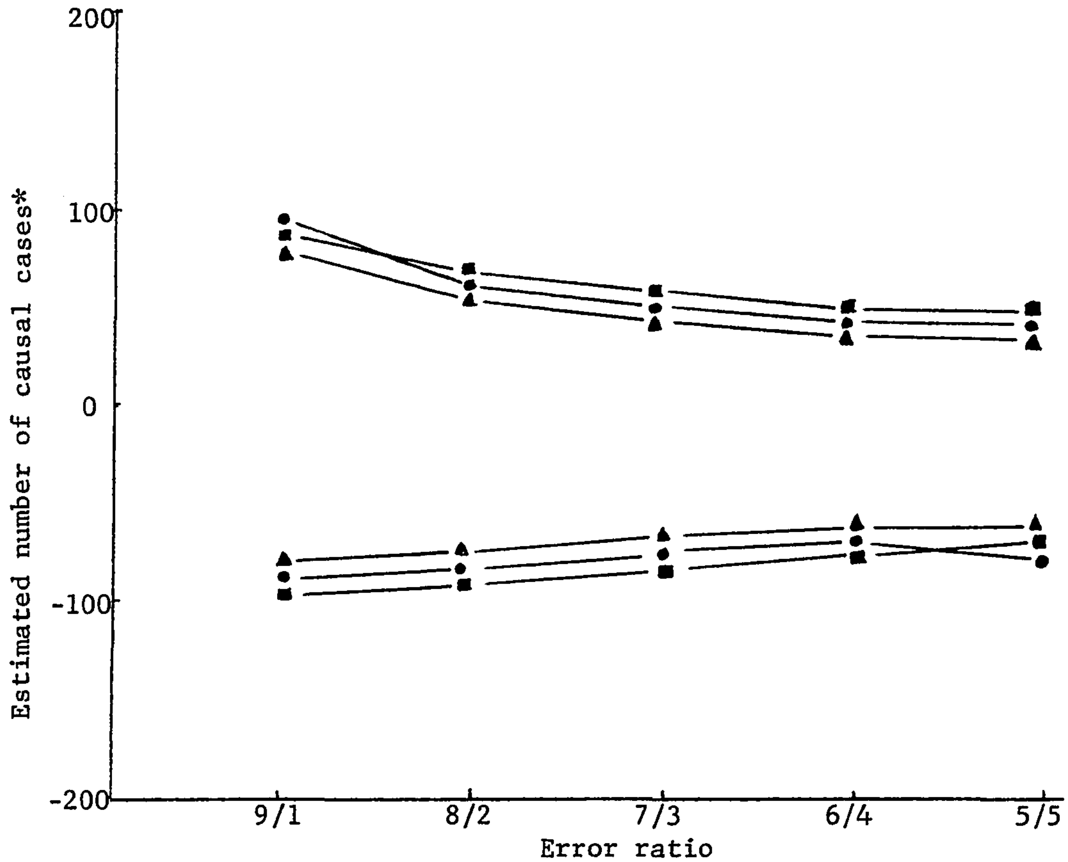


Figure 36. Effects of error and the amount of causal influence on the estimation efficiency of FCP with respect to model B3 (congruent-incongruent-influence, correlated) data sets on the basis of correct predictions.

Influence coefficient:

± 1.0	± .50	±.25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

sensitivity toward YI cases is attenuated with either error or weak causal influence. High reliability or strong causal influence, on the other hand, tends to aggravate this sensitivity, resulting in ambiguous causal estimates. The attenuative effects, however, seem to disappear when data reliability is low and causal influence is weak (See Table XV).

Again, it should be pointed out that when real life data are used, the researcher will have to make causal inferences on the basis of obtained predictions. He will not have the advantage of knowing the correct predictions.

While the overall effects of error on estimation power can be regarded as mild, they appear to be quite substantial when the error ratio drops from 9/1 to 8/2. The effects of the amount of influence are shown to be moderate but consistent. In some instances, slight interaction appears to occur between error and the amount of causal influence (See Tables XIV and XVI; Figures 37 and 38).

Summary

CLC, PC, EC, and VC are shown to be generally ineffective procedures for identifying the source of causal influence with model B3 data. Most of the estimates yielded by these models turn out to be ambiguous. Where data sets are correctly estimated, the pattern of results appears to be inconsistent and sometimes baffling.

Causal estimates yielded by FSM, FCP, and MFCP are based on the relative frequencies of YC, YI, XC, and XI cases rather than the number of correctly identified individual YC and YI cases. In other words, when a YC case is erroneously identified as YI, the results may or may not be detrimental to the efficiency of the model. In some instances such

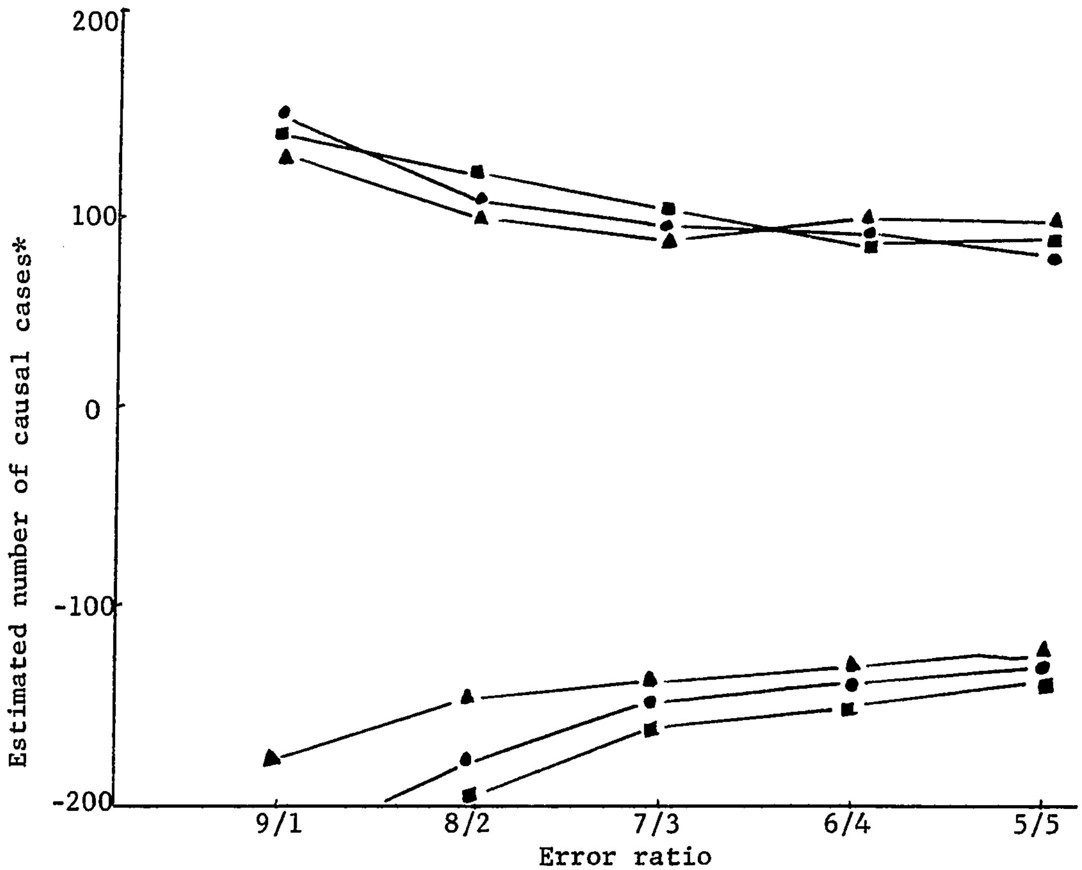


Figure 37. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model B3 (congruent-incongruent-influence, correlated) data sets on the basis of obtained predictions.

Influence coefficient:

$\pm$ 1.0	$\pm$ .50	$\pm$ .25
■	●	▲

*indicates that the estimated number of causal cases pertains to Y as the source of the causal influence

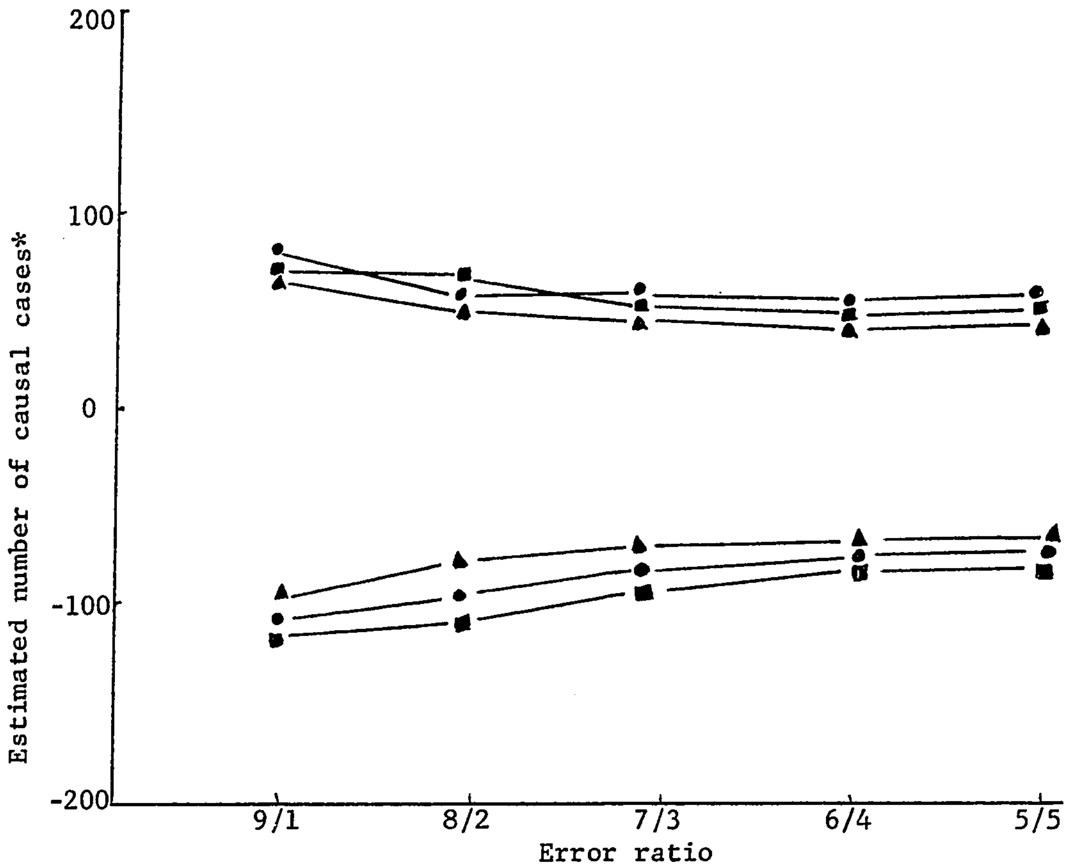


Figure 38. Effects of error and the amount of causal influence on the estimation efficiency of MFCP with respect to model B3 (congruent-incongruent-influence, correlated) data sets on the basis of correct predictions.

Influence coefficient:

$\pm$ 1.0 ■ $\pm$ .50 ● $\pm$ .25 ▲

*indicates that the number of causal cases pertains to Y as the source of the causal influence

misclassifications may actually serve to enhance the estimation power of the model. It follows that there is almost always a discrepancy between obtained predictions and correct predictions. The former include all estimated YC and YI cases while the latter consist of only YC and YI cases correctly identified. Such discrepancies, however, may or may not create a difference in terms of the overall estimation power of the model.

On the basis of obtained predictions, FSM and MFCP appear to have little or no efficiency in detecting the source and direction of causal influence with model B3 data. Of the seven causal models, FCP appears to be the most efficient but its efficiency is limited to highly reliable data.

The discrepancies between obtained predictions and correct predictions clearly suggest that the efficiency of FSM, FCP, and MFCP is in a sense highly exaggerated. As far as individual predictions (i.e., the correct predictions of individual YC and YI cases) are concerned, the true power of the model is much lower than it would appear. This is indicated by the fact that, in most cases, less than two-thirds of the obtained predictions turn out to be correct predictions (See Tables XIV and XVI).

In general, error appears to have rather mild effects on the estimation power of the models. The same can be said about the effects of the amount of influence. In some instances, there appears to be a slight interaction between the two factors.

All in all, it seems evident that none of the causal models have a satisfactorily high level of efficiency to justify their general use with model B3 data. If a causal procedure has to be used, the most logical choice would appear to be FCP. It should, however, be borne in mind that the efficiency of FCP is probably limited to highly reliable data.

CHAPTER VI

DISCUSSION

General considerations

An important part of the present study concerns the definition of cause. The definition provides the rationale for the simulation model. From this point of view it seems important to recall that cause as it is used in the present study is directly related to change. A change implies a force and the source of the force is defined as the cause of the change. All data sets used in the study are created on the basis of this notion of cause.

None of the simulated data sets will exactly parallel a data set in real life. On the other hand, every simulated data set may have its "real life" counterpart. The degree of correspondence in any particular case appears to be a judgment that can best be made by the empirical researcher.

Another important aspect of the present study is the set of decision rules used to determine the efficiency of the causal models. These rules are in a sense arbitrary. If the criterion of statistical significance appears to be crude or one-sided, it might be defended on the grounds that in all data sets the source of causal influence lies with only one variable (Y). In other words, the question of dominant causal influence does not seem to be strictly applicable to the simulated data sets. In addition, it should be borne in mind that the decision rules are formulated as guidelines for making decisions with respect to each data set as a whole rather than individual YC or YI cases. All estimates are to be considered as group estimates rather than individual predictions.

The following sections present a general review of the efficiency of the seven causal models included in the present study.

Cross-lagged correlation model (CLC)

The model is by far the most widely used procedure for causal analysis. (See Campbell and Stanley, 1963; Campbell, 1963; Rozelle and Campbell, 1969). The results of the present study appear to suggest that the model is highly effective with model A1 and model A2 data sets but with model A3, B1, B2, and B3 data sets the model appears to be completely ineffective.

Relative to model B1, B2, and B3 data sets, two factors seem to complicate the identification of causal influence. First, since the time-one measures are correlated, none of the cross-lagged correlations can be expected to be zero. When both correlation coefficients reach statistical significance, an ambiguous estimate results. Second, when the causal influence is incongruent the magnitude of $r_{y_1x_2}$ is found to be extremely low. When this correlation loses its statistical significance and $r_{x_1y_2}$ remains significant, an erroneous estimate results. The first problem can be resolved by testing the difference between the two correlation coefficients as has been suggested by various writers (e.g., Campbell, 1963). Solution to the second problem appears to be more difficult to obtain.

It thus seems obvious that the use of CLC should be confined to certain types of data (i.e., model A1 and model A2). This implies, of course, that the researcher should have some prior knowledge about his empirical data. This finding appears to be in full agreement with the contention (Polk, 1962) that causal inferences can rarely be made solely on the basis of empirical data. Some prior knowledge about the phenomenon under investigation appears to be desirable if not absolutely necessary.

Part correlation model (PC)

The model appears to be highly efficient with model A1 and model A2 data sets. Relative to model B1 data sets, its power of estimation seems to be much lower. The model is shown to be entirely ineffective with model B2, model A3, and model B3 data sets. In other words, the model is shown to be highly effective with data sets where the time-one measures are uncorrelated and the direction of the causal influence is either congruent or incongruent. When both congruent and incongruent influences co-exist, the model appears to be quite ineffective. Relative to data where the time-one measures are correlated, the model seems to be more efficient with congruent than with incongruent data. When varying amounts of causal influence or both congruent and incongruent causal influences are present, causal analysis again becomes difficult. All this suggests that PC is not without its limitations and should therefore be used with caution.

Econometric model (EC)

The model appears to be highly effective with model A1 and model A2 data sets. However, when the model is used to analyze model A3 data sets, its efficiency declines sharply. Both beta weights for variable X and variable Y are found to be extremely low. In no instances is either beta weight statistically significant. Causal estimates are consequently ambiguous.

EC also appears to have low efficiency with model B1, B2, and B3 data where the time-one measures are correlated. In such cases both beta weights tend to be of considerable magnitude and statistically significant, yielding ambiguous causal estimates.

Thus, a widely used technique for causal analysis (Blalock, 1969; Wonnacott and Wonnacott, 1970) appears to have much less utility than generally expected. The model is shown to be effective only with data where the time-one measures are uncorrelated and the direction of causal influence is either congruent or incongruent. When the time-one measures are correlated or when congruent and incongruent causal influences are present simultaneously, EC appears to have little or no efficiency in estimating causal influence.

Variance components model (VC)

As indicated earlier, the model is appropriate for data sets where the time-one measures are correlated. If used on data where the time-one measures are uncorrelated, the results of causal analysis will be entirely dependent on sampling fluctuation.

VC appears to be efficient with data where the causal influence is congruent and strong. It yields ambiguous estimates with data of incongruent or weak causal influence. Relative to data where both congruent and incongruent causal influences are present, the efficiency of the model also appears to be low.

In general, it may be said that the model has utility only for certain kinds of data. Its general efficiency appears to be less than desirable--especially when the causal influence is incongruent or when both congruent and incongruent causal influences are present.

Frequencies-of-shift-across-median model (FSM)

As indicated in earlier sections, the efficiency of FSM rests heavily on the relative frequencies of YC and YI cases rather than the identification of individual YC or YI cases. Should the latter be used as the criterion of efficiency, the model would probably turn out to be the least effective causal procedure.

In general, strong causal influence is required for the model to detect the source and direction of such influence. Quite surprisingly, however, the model is shown to be highly effective with model A3 data where congruent and incongruent causal influences are present.

With data sets where the time-one measures are correlated, the efficiency of the model appears to decline quite considerably. The model is shown to have low efficiency in detecting causal influence when both congruent and incongruent influences are present in such data sets. With these data sets the model fails to detect the direction of the causal influence. The source of the influence is in most cases correctly identified.

Relative to the overall efficiency of the model it may be said that despite the fact that FSM almost always classifies more than half of the cases as "uncertain" (i.e., XU, YU, UC, UI, UU) correct causal estimates are often made on the basis of the relatively few "certain" cases (i.e., YC, XC, YI, XI). Again, this is a result of the fact that the causal estimates are made on the basis of the relative frequencies of YC, YI, XC, and XI cases rather than the number of correctly identified YC and YI cases. It follows that the efficiency of FSM applies only to situations where the model is used as a group estimating procedure. If FSM is used to identify individual YC, YI, XC, XI cases, it will probably yield more misclassifications than correct classifications.

Frequencies-of-change-in-product-moment model (FCP)

The model has been regarded as the most efficient causal analysis procedure presently available to researchers (Yee, 1966). The results of the present study indicate that while the model certainly appears to be more versatile than most causal procedures, its efficiency is limited to certain types of data.

When the time-one measures are uncorrelated, the model is shown to be effective with data of high reliability. Ambiguous estimates are made with low reliability. In some cases, only the source of causal influence is correctly identified, estimates of the direction of the influence being ambiguous. With data where both congruent and incongruent influences are present, the model yields correct estimates only when data reliability is high and the causal influence is strong.

It will be noted that following FSM the efficiency of the model is dependent on the relative frequencies of YC, YI, XC, and XI cases rather than the number of correct individual predictions. Thus, despite the fact that the discrepancies between obtained predictions and correct predictions appear to be quite remarkable, the final causal estimates remain the same.

When the time-one measures are correlated, high data reliability appears to be a necessary condition for correct causal estimation. When both congruent and incongruent influences are present (i.e., B3 data) the efficiency of the model is more severely limited. Only the highly reliable data sets are correctly estimated.

While the results based on a few correct predictions appear to be better than those based on all obtained predictions, it should be realized that the researcher using "real life" data will simply have no knowledge of the correct predictions.

One more point needs to be stressed. The versatility of FCP is limited to situations where the model is used as a group procedure for detecting the source and direction of causal influence. If it is used as a technique for identifying individual YC, YI, XC, XI cases, the efficiency of the model would be very low. As a matter of fact, in most cases it will yield more erroneous classifications than correct classifications.

<u>Modified-frequencies-of-change-in-product-moment model (MFCP)</u>

The model is designed as an improvement on FCP and is shown to be so with respect to model A1, A2, and A3 data. In such cases the model fails to identify causal influence only when the data reliability is very low or the causal influence is weak.

It is to be noted, however, that when both congruent and incongruent influences are present, the pattern of results appears to be quite peculiar; ambiguous causal estimates are made with data of high reliability and where the causal influence is strong. A closer examination of the results reveals that this failure appears to be a result of the model's tendency to overclassify YI cases. As data reliability or the amount of causal influence decreases, this tendency appears to be weakened.

MFCP is also shown to be slightly better than FCP when used to analyze model B1 data. The model yields more correct causal estimates than FCP with high reliability data. The model is, however, shown to be less efficient than FCP when varying amounts of causal influence are present in the data.

When both congruent and incongruent influences are present, MFCP appears to be completely ineffective. All causal estimates turn out to be ambiguous. Again, much of the failure can be attributed to the model's tendency to overclassify YI cases. This gives rise to a difference between YI and YC cases which indicates that the causal influence is congruent. The source of the causal influence is correctly identified in most cases.

Following FSM and FCP, the efficiency of MFCP pertains only to situations where the model is used as a group procedure. As a technique for classifying individuals on the basis of source and direction of causal influence, the model would appear to have very limited utility. In most clases, more than half of the cases will be misclassified.

CHAPTER VII

SUMMARY OF PART I

The present investigation derives a numerical definition of cause and on the basis of this definition develops a computer simulation model to generate six types of data with different causal characteristics. These characteristics pertain to (1) the direction of causal influence, (2) the amount of causal influence, and (3) data reliability. A total of 110 data sets were created.

Seven causal models are briefly described and their apparent strengths and weaknesses discussed. Five of these models are generally known. The other two are developed in the course of the present study. All seven causal models are applied to the 110 data sets. The results are tabulated and a set of decision rules are formulated to determine the relative efficiency of the models.

In general, the results appear to suggest that the estimation power of the causal models has been exaggerated by authors of causal studies reported in the literature. That is not to say, however, that causal models should not be used at all in psychological and educational research. Rather, the point is that the use of these models should be made with full awareness of their strengths and limitations. The major contribution of the present investigation may well lie in its efforts to make these limitations explicit to researchers who need to use the models in their research.

The relative efficiency of the models is summarized in the table on the following page:

	Causal Model						
Causal Data	CLC	PC	EC	VC	FSM	FCP	MFCP
A1 (uniform-influence, uncorrelated)	+	+	+		-	-	+
A2 (varied-influence, uncorrelated)	+	+	+		+	+	+
A3 (congruent-incongruent-influence, uncorrelated)	-	-	-		+	+	-
B1 (uniform-influence, correlated)	-	+	-	-	-	-	-
B2 (varied-influence, correlated)	-	-	-	-	+	+	-
B3 (congruent-incongruent-influence, correlated)	-	-	-	-	-	-	-

A plus sign indicates that the causal model may generally be used to analyze the data. A minus sign suggests that the estimation power of the model is limited and the model should probably not be applied to the data.

It should be realized that the above summary entails some degree of subjectivity. VC, for instance, appears to be highly effective with model B1 and model B2 data when the causal influence is congruent and strong. The model is, however, considered as "minus" because of its limitations with respect to incongruent causal data.

Several interesting findings are noted. First, data types listed in the lower half of the table appear to be much more difficult to handle than data types listed in the upper half of the table. Second, with respect to model B3 data, none of the causal models appear to be efficient enough to justify general use. Third, none of the models appear to be efficient with all data types included in the present investigation. The most efficient models (i.e., PC, FSM, FCP) are shown to be effective with only approximately half of the data sets. In other words, the indiscriminate use of the causal models would probably result in causal interpretations which more often than not are erroneous.

PART TWO

EMPIRICAL APPLICATION OF SEVEN CAUSAL MODELS

CHAPTER VIII

INTRODUCTION TO PART II

The findings reported in the previous chapters appear to suggest that the various causal models are to a great extent independent of each other not only in terms of the rationale upon which they are formulated but also in terms of the causal estimates yielded by the models. It is clear that the models should not be regarded as comparable to each other. It is also clear that even where the models yield the same results, such results may still be erroneous. In other words, similarity of results yielded by the models does not necessarily suggest convergent validity.

All this renders the causal interpretation of real data (i.e., non-simulated data) very difficult. Specifically, without any prior knowledge about the data, no single model can be depended upon to unravel the causal relations between variables. Nor can similar results yielded by the majority of the models be regarded as the correct causal estimates.

This is not to say, however, that the situation is hopeless. A closer examination of the results reported in the previous chapters indicates that the various models appear to yield similar results with respect to only certain types of data and not others. Thus, the comparability of the models appears to be contingent upon data types. This means that it is possible to classify data on the basis of the results yielded by the models.

Furthermore, the agreement pattern among the causal models is not the only source of information to which the researcher can turn in attempting to resolve the difficulty. Prior knowledge of data reliability and the magnitude of the correlation between the time-one measures will also help the researcher to identify data types.

Once the data type is identified on the basis of "agreement pattern" among the causal models and prior knowledge about the data, it becomes possible for the researcher to determine the most efficient causal model for the particular data type. This is accomplished by examining the efficiency of the various models with respect to data types as reported in Tables I, III, V, IX, XI, and XIII in Chapter V.

In light of what has been said in the previous sections, it appears appropriate to formulate some decision rules to serve as a guideline for making causal inferences with real data. These are listed as follows:

1. Examine the correlation coefficient between the time-one measures (i.e., $r_{x_1y_1}$). If the correlation is significant ($p < .05$), the data type belongs to the model B category. If the correlation is not significant, the data type belongs to the model A category.
2. Obtain the approximate reliability of the data. The reliability coefficient is matched against the error ratios to identify the data type. As mentioned earlier, the error ratios should be considered as lower bound reliabilities.
3. The "agreement pattern" of the causal models is examined and matched against the patterns presented in Tables I, III, V, IX, XI, XIII in Chapter V. It is important to note here that the "agreement pattern" has nothing to do with the efficiency of the models. It is merely an indication of comparability among the causal models. In other words, high agreement among the models does not necessarily suggest convergent validity. The agreement pattern is examined solely for the purpose of identifying

the data type. It will also be noted that not all agreement patterns will exactly match one (or more) of those presented in Tables I, III, V, IX, XI, and XIII. When a perfect match is not possible, a near-perfect match may suffice.

Using the above criteria, the researcher is likely to be able to identify the data type that he is dealing with. Once the data type is identified, he will be able to determine the most efficient causal model for the data and subsequently to make a causal interpretation which is most likely to be valid. The following chapters describe an application of this procedure to two sets of real (non-simulated) data.

CHAPTER IX

CAUSAL ANALYSIS OF VOCABULARY AND COMPREHENSION ABILITIES

Introduction

Most standardized reading tests consist of two sections: vocabulary and comprehension. This seems to attest to the fact that vocabulary and comprehension are generally regarded as two different types of mental ability. It has also been commonly observed that these two types of mental ability are somehow related to each other. The question thus arises: Does vocabulary ability contribute to comprehension ability? If not, is the truth the other way around? Or are the two types of mental ability independent of each other? The study reported in the following sections attempts to provide answers to the above questions.

Procedures

The Gates-MacGinitie Reading Tests were administered to 55 second-graders and 55 third-graders in Hawaii during the 1971-72 school year. The Primary B Form 1 was given to the second-graders and the Primary C Form 1 was given to the third-graders. Pretest and posttest measures were obtained in October and April, respectively. The second- and third-graders were students randomly selected from public schools in Hawaii to participate in an evaluation study conducted for the Hawaii English Project.

The test data were analyzed by means of the various causal models separately for each grade level. The following specific hypotheses were tested:

(1) Vocabulary is the cause of comprehension and the causal influence is congruent.

(2) Vocabulary is the cause of comprehension and the causal influence is incongruent.

(3) Comprehension is the cause of vocabulary and the causal influence is congruent.

(4) Comprehension is the cause of vocabulary and the causal influence is incongruent.

Results

The results obtained for the second graders are reported in Table XVII. CLC appears to suggest that the causal influence is mutual and congruent. In other words, high vocabulary ability tends to contribute to high comprehension ability, which in turn contributes to high vocabulary ability.

PC indicates that the predominant causal influence is from vocabulary to comprehension. High vocabulary ability tends to produce high comprehension ability. This is entirely in agreement with the common expectation that the student has to know the meaning of the words before he can understand the sentence or the passage.

EC yields results which are in full agreement with those obtained for CLC. That is to say, the causal influence between vocabulary ability and comprehension ability appears to be mutual.

VC, FCP, and MFCP yields ambiguous results with respect to the source and direction of causal influence. FSM identifies vocabulary ability as the cause of comprehension ability. The model, however, fails to determine the direction of the causal influence.

The chi square test results for FSM, FCP, and MFCP are reported in Table XVIII.

The results obtained for the third-graders fall into the same general

Table XVII. Estimation of Causal Relationship Between Vocabulary Ability and Comprehension Ability of Second-graders (N = 55)

Causal Influence	CLC	PC	EC	VC	FSM	FCP	MFCP
	Magnitude of Causal Estimate						
Voc $\xrightarrow{+}$ Comp	.81*	.33*	.68*	1.22	13	23	19
Voc $\xrightarrow{-}$ Comp					6	12	15
Comp $\xrightarrow{+}$ Voc	.65*		1.06*	.96	4	8	8
Comp $\xrightarrow{-}$ Voc		-.01			2	12	13

* $p < .05$

Note: In the above table, cross lagged-correlation coefficients, part correlation coefficients, beta weights, and F ratios are reported for CLC, PC, EC, and VC, respectively. For FSM, FCP, and MFCP, XC, YC, YI, XI cases are reported. In all instances, no YU or XU cases are found for FSM.

Table XVIII. Significance of Chi Square Values for FSM, FCP, and MFCP in the Estimation of Causal Relationship Between Vocabulary Ability and Comprehension Ability of Second-graders (N = 55) and Third-graders (N = 55)

Grade Level	Variable	Test	Significance of X^2 FSM	FCP	MFCP
	Vocabulary	S	*	-	-
Second	Comprehension	D1	-	-	-
		D2	-	-	-
	Vocabulary	S	-	*	*
Third	Comprehension	D1	-	-	-
		D2	-	-	-

* $p < .05$

Note: The first direction test (D1) assumes that the first variable is the source of causal influence; the second direction test (D2) assumes that the second variable is the source of causal influence. When the source test (S) is non-significant, the direction tests become irrelevant.

pattern. These are shown in Table XIX. CLC, again, suggests that the causal influence is mutual. High vocabulary ability contributes to high comprehension ability and vice versa.

PC, on the other hand, indicates that the causal influence comes from vocabulary ability. High vocabulary ability is followed by high comprehension ability.

EC is in agreement with CLC in suggesting that the causal influence is probably mutual. High vocabulary ability contributes to high comprehension ability which in turn contributes to high vocabulary ability.

VC and FSM yield ambiguous results. Neither the source nor the direction of the causal influence is identified.

Both FCP and MFCP suggest that the causal influence lies with vocabulary. The models, however, fail to determine the direction of the causal influence.

The chi square test results for FSM, FCP, and MFCP are shown in Table XVIII.

<u>Application of decision rules</u>

The correlation between the time-one measures of vocabulary and comprehension was highly significant ($r = .78$, $p < .01$, for second-graders and $r = .84$, $p < .01$, for third-graders). The data type is thus confined to the model B category.

Buros (1972) reported reliabilities in the vicinity of .80 for the Gates-McGinitie Reading Tests. In terms of error ratio, it appears appropriate to classify the data type to be in the category of 8/2.

To determine the "agreement pattern" among the causal models, the overall results obtained for the vocabulary-comprehension data may be

Table XIX. Estimation of Causal Relationship Between Vocabulary Ability and Comprehension Ability of Third-graders (N = 55)

Causal Influence	CLC	PC	EC	VC	FSM	FCP	MFCP
Voc + Comp	.79*	.26*	.78*	.94	8	23	18
Voc - Comp					4	17	18
Comp + Voc	.74*	.07	.76*	.90	5	8	10
Comp - Voc					0	7	9

*p < .05

Note: In the above table, cross-lagged correlation coefficients, part correlation coefficients, beta weights, and F ratios are reported for CLC, PC, EC, and VC, respectively. For FSM, FCP, and MFCP, YC YI, XC, XI cases are reported. In all instances, no YU or XU cases are found for FSM.

diagrammatically presented as follows:

Model	CLC	PC	EC	VC	FSM	FCP	MFCP
Estimation	±	+	±	±	±	±	±

The diagram shows that six of the seven causal models yield the same causal estimates. The part correlation model (PC) is the only model that suggests a different interpretation. It is important to realize that the plus sign and the combination of plus and minus are here used as symbols to describe causal estimates yielded by the causal models. The symbols merely denote similarity and dissimilarity of results. They do not--as they did in previous chapters--indicate the validity of the causal models. The symbols merely describe an "agreement pattern" among the causal models.

Putting the three criteria (correlation of time-one measures, data reliability, and "agreement pattern" among causal models) together, one finds that the data may best be described as a B1 data type with an error ratio of 8/2 and an influence coefficient of .25. An examination of Table IX in Chapter V suggests that only the part correlation model (PC) is likely to yield the correct causal estimates with this particular data type.

Discussion

In view of the results reported in the previous sections it may be said that while some of the causal estimates are ambiguous, the overall indications appear to be clear: vocabulary may be regarded as the cause of comprehension ability. This finding is supported by the results yielded by the part correlation model (PC) which is identified as the most efficient model with this particular data type. The results of PC suggest that the direction of causal influence is congruent. This finding also appears to be consistent with what is generally known about the relationship between vocabulary ability and comprehension ability. That is to say, the two variables are generally known to be positively correlated.

A notable feature of the study is the close similarity between results obtained for the second- and third-graders. Although the two sets of results are not identical, the general findings are the same. Thus, each set of results serves as a cross-sectional replication of the other.

Relative to the efficiency of the causal models the results yielded by PC are considered to be most likely to be valid. Moreover, they seem to be in agreement with the overall trend indicated by other models. The results yielded by CLC and EC are ambiguous with respect to the source of the causal influence. VC fails to determine the source as well as the direction of the causal influence. FSM, FCP and MFCP, in general, identify vocabulary ability as the source of causal influence but fail to determine the direction of such influence.

Summary and conclusions

Pretest and posttest data were obtained from 110 second- and third-graders in Hawaii with respect to their vocabulary and comprehension abilities as measured by the Gates-MacGinitie Reading Tests. Seven causal models were used to analyze the data. The results of these analyses appear to support the following conclusions:

(1) For both second- and third-graders, vocabulary ability is shown to be the cause of comprehension ability.

(2) The causal relationship between vocabulary ability and comprehension ability appears to be congruent. That is to say, high vocabulary ability contributes to high comprehension ability.

(3) Relative to the efficiency of the causal models, PC appears to have yielded the most conclusive results.

(4) Although ambiguous estimates are yielded by some models, none of the estimates are in contradiction with the overall findings.

CHAPTER X

CAUSAL ANALYSIS OF HEIGHT, WEIGHT, AND STRENGTH

Introduction

The interrelationships of height, weight, and strength are intriguing but little investigated. These three variables are obviously highly related but the causal influence among them might be complex and difficult to detect. Such questions as "does height cause strength? Or does strength cause height?" are of great interest to researchers in diverse fields related to human development. Yet, in the absence of appropriate statistical techniques to disentangle the complex interrelationships among the variables, little is known about their causal relations.

Procedures

Data used in the present study were taken from Tuddenham and Synder (1954). Measures of height, weight, and strength were obtained for a total of 136 subjects (66 boys and 70 girls) at ages 9, 11, 13, and 15. Weight and strength were measured in kilograms. Height was measured in centimeters.

Two sets of data were obtained from the original data: one set consisted of data for ages 9 and 15 and the other set comprised data for ages 11 and 13. Two causal intervals of 6 and 2 years, respectively, were thus obtained. The seven causal models were applied to the data to determine the source and direction of the causal influence. In analyzing each set of data, separate analyses were conducted for the total sample, the male sample, and the female sample.

For each set of analyses, the following hypotheses were tested:

(i) Height causes weight and the causal influence is congruent.

(ii) Height causes strength and the causal influence is congruent.

(iii) Weight causes strength and the causal influence is congruent.

<u>Results for the first set of data</u>

Table XX reports results obtained for measures of height, weight, and strength taken at ages 9 and 15. Overall, these appear to indicate that the causal interrelationships among the three variables are complex and subtle. For the total sample, CLC and EC suggest that the relationship between weight and height is probably one of mutual causal influence. That is to say, weight causes height, which in turn causes weight. PC seems to indicate that weight is probably the cause of height--the causal influence being incongruent. In other words, the more a person weighs, the shorter he tends to become. VC appears to support the hypothesis that height causes weight and the causal influence is congruent.

As shown in Table XXI, most of the chi square values obtained for FSM, FCP, and MFCP are non-significant, yielding ambiguous estimates of causal influence. MFCP suggests that weight is probably the cause of height. The direction of such influence, however, remains uncertain.

With respect to weight and strength, both CLC and EC suggest that the causal influence is probably mutual. In other words, weight causes strength, which in turn causes weight. The results obtained for PC and VC are ambiguous.

In general, FSM, PCP, and MFCP indicate that weight is probably the cause of strength although the direction of the causal influence cannot be determined. Table XXI reports the significance of the chi square values

Table XX. Estimation of Causal Relationships Between Weight, Height, and Strength for the Total Sample (N = 136)

Causal Influence			Magnitude of Causal Estimate				
	CLC	PC	EC	VC	FSM	FCP	MFCP
$W_{11} \xrightarrow{+} H_{13}$	.59*		.56*	.97	5	27	31
$W_{11} \xrightarrow{-} H_{13}$		-.20*			6	44	56
$H_{11} \xrightarrow{+} W_{13}$	.71*	.10	1.09*	1.22	12	43	29
$H_{11} \xrightarrow{-} W_{13}$					6	22	20
$W_{11} \xrightarrow{+} S_{13}$	.49*		1.39*	1.15	14	34	45
$W_{11} \xrightarrow{-} S_{13}$		-.05			8	47	38
$S_{11} \xrightarrow{+} W_{13}$	.53*		.32*	1.09	5	31	26
$S_{11} \xrightarrow{-} W_{13}$		-.08			13	24	27
$H_{11} \xrightarrow{+} S_{13}$	.56*	.09	1.96*	1.31	15	44	39
$H_{11} \xrightarrow{-} S_{13}$					9	27	33
$S_{11} \xrightarrow{+} H_{13}$	.51*		.24*	1.02	2	31	33
$S_{11} \xrightarrow{-} H_{13}$		-.09			10	34	31
$W_9 \xrightarrow{+} H_{15}$	.41*		.55*	.86	13	23	29
$W_9 \xrightarrow{-} H_{15}$		-.20*			15	44	46
$H_9 \xrightarrow{+} W_{15}$	.59*	.10	1.12*	1.49*	8	43	34
$H_9 \xrightarrow{-} W_{15}$					6	26	27
$W_9 \xrightarrow{+} S_{15}$	.30*	.01	1.65*	1.28	19	39	44
$W_9 \xrightarrow{-} S_{15}$					15	38	51
$S_9 \xrightarrow{+} W_{15}$	.36*		.24*	1.27	13	37	21
$S_9 \xrightarrow{-} W_{15}$		-.02			4	22	20
$H_9 \xrightarrow{+} S_{15}$	.40*	.09	2.38*	1.53*	15	42	44
$H_9 \xrightarrow{-} S_{15}$					11	34	45
$S_9 \xrightarrow{+} H_{15}$	.45*	.03	.23*	1.18	14	39	32
$S_9 \xrightarrow{-} H_{15}$					9	21	15

* $p < .05$

Note: In the above table, cross-lagged correlation coefficients, part correlation coefficients, beta weights, and F ratios are reported for CLC, PC, EC, and VC, respectively. For FSM, FCP, and MFCP, YC, YI, XC, XI are reported. In all instances, no YU or XU cases are found for FSM.

Table XXI. Significance of Chi-square Values for FSM, FCP, and MFCP in the Estimation of Causal Relationships Between Weight, Height, and Strength for the Total Sample (N = 136)

Variables	Test	Significance of X^2		
		FSM	FCP	MFCP
Height and weight at ages 11 and 13	S	--	--	*
	D_1	--	*	--
	D_2	--	--	*
Weight and strength at ages 11 and 13	S	--	*	*
	D_1	--	--	--
	D_2	--	--	--
Height and strength at ages 11 and 13	S	--	--	--
	D_1	--	--	--
	D_2	*	--	--
Height and weight at ages 9 and 15	S	--	--	*
	D_1	--	--	--
	D_2	--	*	--
Weight and strength at ages 9 and 15	S	*	--	*
	D_1	--	--	--
	D_2	--	--	--
Height and strength at ages 9 and 15	S	--	--	*
	D_1	--	--	--
	D_2	--	*	*

* $p < .05$

Note: The first direction test (D_1) assumes that the first variable is the source of causal influence; the second direction test (D_2) assumes that the second variable is the source of causal influence. When the source test (S) is non-significant, the direction tests become irrelevant.

yielded by the three models. Relative to the causal relationship between height and strength, CLC and EC indicate mutual causal influence between the two variables. The results yielded by PC are ambiguous, while VC suggests that height is probably the cause of strength.

The significance of the chi square values computed for FSM, FCP, and MFCP is shown in Table XXI. MFCP indicates that height is probably the cause of strength. The model, however, fails to determine the direction of the causal influence. The results yielded by FCP and FSM are essentially ambiguous.

All in all, the causal estimates presented above reflect a rather inconsistent pattern of relations among weight, height, and strength. To determine if this result can partially be attributed to the fact that the data were obtained from both male and female samples, the data were subdivided on the basis of sex. The following sections describe results obtained for the male and female samples using sex as a possible "qualifier."

Results for the male sample

Relative to the causal relations between weight and height, CLC and EC indicate that the relationship is probably one of mutual causal influence. PC, on the other hand, suggests that weight is the cause of height and the causal influence is incongruent. VC, FSM, FCP, and MFCP yield ambiguous estimates. Neither the source nor the direction of the causal influence is determined.

With respect to the relationship between weight and strength, CLC and EC show that the causal influence between the two variables is probably mutual. MFCP suggests that weight is probably the cause of strength.

The direction of the influence is, however, not determined. The results obtained for VC, FSM, and FCP are ambiguous.

Relatively more conclusive results were obtained for the relationship between height and strength. CLC and EC suggest mutual congruent influence between the two variables. PC and VC suggest that height is probably the cause of strength and the direction of the influence is congruent. FSM, FCP, and MFCP indicate that height is probably the cause of strength, but the direction of the causal influence cannot be determined.

The results described above are shown in Tables XXII and XXIII.

Results for female sample

With respect to the causal relations between weight and strength, the results obtained for girls were found to be in full agreement with those obtained for boys. Specifically, CLC and EC indicate mutual congruent influence between the two variables. PC, on the other hand, suggests that weight is probably the cause of height, the direction of the causal influence being incongruent. Causal estimates yielded by VC, FSM, FCP, and MFCP are ambiguous.

The relationship between weight and strength also appears to be similar for both boys and girls. CLC and EC again indicate mutual congruent causal influence between the two variables. PC and VC yield ambiguous results. FCP and MFCP appear to indicate that weight is probably the cause of strength. The direction of such causal influence is, however, uncertain. The results obtained for FSM are ambiguous.

Relative to the relations between height and strength, the results obtained for girls are less conclusive than the results obtained for

Table XXII. Estimation of Causal Relationships Between Weight, Height, and Strength for Boys (N = 66)

Causal Influence	Magnitude of Causal Estimate CLC	PC	EC	VC	FSM	FCP	MFCP
$W_{11} \xrightarrow{+} H_{13}$	.52*		.62*	1.18	5	16	14
$W_{11} \xrightarrow{-} H_{13}$		-.22*			2	11	20
$H_{11} \xrightarrow{+} W_{13}$	.67*	.22*	1.11*	1.46	6	30	24
$H_{11} \xrightarrow{-} W_{13}$					4	9	8
$W_{11} \xrightarrow{+} S_{13}$	.55*	.03	1.81*	1.31	9	26	26
$W_{11} \xrightarrow{-} S_{13}$					4	10	18
$S_{11} \xrightarrow{+} W_{13}$	.57*		.35*	1.26	3	17	15
$S_{11} \xrightarrow{-} W_{13}$		-.04			7	13	7
$H_{11} \xrightarrow{+} S_{13}$	.65*	.31*	2.59*	1.68*	5	28	31
$H_{11} \xrightarrow{-} S_{13}$					6	16	14
$S_{11} \xrightarrow{+} H_{13}$	.52*		.28*	1.28	3	14	15
$S_{11} \xrightarrow{-} H_{13}$		-.08			2	8	6
$W_{9} \xrightarrow{+} H_{15}$	.43*		.58*	1.02	6	13	14
$W_{9} \xrightarrow{-} H_{15}$		-.20*			1	14	17
$H_{9} \xrightarrow{+} W_{15}$	.56*	.08	1.13*	1.51	7	24	18
$H_{9} \xrightarrow{-} W_{15}$					6	15	17
$W_{9} \xrightarrow{+} S_{15}$	.39*	.12	1.98*	1.43	15	23	23
$W_{9} \xrightarrow{-} S_{15}$					8	13	22
$S_{9} \xrightarrow{+} W_{15}$	.41*		.30*	1.23	8	20	14
$S_{9} \xrightarrow{-} W_{15}$		-.08			3	10	7
$H_{9} \xrightarrow{+} S_{15}$	.49*	.28*	2.84*	2.06*	17	28	28
$H_{9} \xrightarrow{-} S_{15}$					8	17	15
$S_{9} \xrightarrow{+} H_{15}$	.34*		.19*	1.03	5	12	14
$S_{9} \xrightarrow{-} H_{15}$		-.15			1	9	9

* $p < .05$

Note: In the above table, cross-lagged correlation coefficients, part correlation coefficients, beta weights, and F ratios are reported for CLC, PC, EC, and VC, respectively. For FSM, FCP, and MFCP, YC, YI, XC, XI are reported. In all instances, no YU or XU cases are found for FSM.

Table XXIII. Significance of Chi-square Values for FSM, FCP, and MFCP in the Estimation of Causal Relationships Between Weight, Height, and Strength for Boys (N = 66)

Variables	Test	Significance of X^2 FSM	FCP	MFCP
Height and weight at ages 11 and 13	S	--	--	--
	D_1	--	*	*
	D_2	--	--	--
Weight and strength at ages 11 and 13	S	--	--	*
	D_1	--	*	--
	D_2	--	--	--
Height and strength at ages 11 and 13	S	--	*	*
	D_1	--	--	*
	D_2	--	--	--
Height and weight at ages 9 and 15	S	--	--	--
	D_1	--	--	--
	D_2	--	--	--
Weight and strength at ages 9 and 15	S	--	--	*
	D_1	--	--	--
	D_2	--	--	--
Height and strength at ages 9 and 15	S	*	*	*
	D_1	--	--	--
	D_2	--	--	--

* $p < .05$

Note: The first direction test (D_1) assumes that the first variable is the source of causal influence; the second direction test (D_2) assumes that the second variable is the source of causal influence. When the source test (S) is non-significant, the direction tests become irrelevant.

boys. While CLC and EC indicate mutual congruent causal influence between the two variables, PC and VC fail to detect any causal influence. FSM, FCP, and MFCP appear to be in full agreement with one another in identifying height as the cause of strength. The direction of the influence is, however, uncertain. See Tables XXIV and XXV.

Results for the second set of data

In general, the shorter causal interval does not seem to alter the general pattern of results. CLC and EC again indicate mutual congruent causal influence between weight and height. PC suggests that weight is probably the cause of height, the causal influence being incongruent. The results obtained for VC and FSM turn out to be ambiguous. MFCP identifies weight as the cause of height, the direction of such influence being incongruent. FCP fails to determine the source of the causal influence.

With respect to the relations between weight and strength, most of the causal models (i.e., CLC, PC, EC, VC, and FSM) yield ambiguous results. FCP and MFCP indicate that weight is probably the cause of strength, although the direction of the influence remains uncertain.

The relationship between height and strength appears to be even more complex and subtle: none of the causal estimates indicate a definite source or direction of causal influence. The above results are shown in Tables XX and XXI.

Results for the male sample

The performance of CLC and EC remains the same: both suggesting

Table XXIV. Estimation of Causal Relationships Between Weight, Height, and Strength for Girls (N = 70)

Causal Influence	Magnitude of Causal Estimate CLC	PC	EC	VC	FSM	FCP	MFCP
$W_{11} \xrightarrow{+} H_{13}$	.69*		.52*	.85	2	14	13
$W_{11} \xrightarrow{-} H_{13}$		-.14			4	27	28
$H_{11} \xrightarrow{+} W_{13}$	.74*	.04	1.06*	1.08	2	13	14
$H_{11} \xrightarrow{-} W_{13}$					2	16	15
$W_{11} \xrightarrow{+} S_{13}$	.53*		1.26*	1.04	3	21	21
$W_{11} \xrightarrow{-} S_{13}$		-.08			6	28	26
$S_{11} \xrightarrow{+} W_{13}$	.58*		.36*	.98	1	10	9
$S_{11} \xrightarrow{-} W_{13}$		-.12			5	11	14
$H_{11} \xrightarrow{+} S_{13}$	.54*		1.63*	1.05	2	18	17
$H_{11} \xrightarrow{-} S_{13}$		-.08			6	25	26
$S_{11} \xrightarrow{+} H_{13}$	.58*		.24*	.83	0	11	11
$S_{11} \xrightarrow{-} H_{13}$		-.17			7	16	16
$W_{9} \xrightarrow{+} H_{15}$	.49*		.51*	.73	3	10	11
$W_{9} \xrightarrow{-} H_{15}$		-.28*			6	21	19
$H_{9} \xrightarrow{+} W_{15}$	.62*	.11	1.08*	1.43	7	20	21
$H_{9} \xrightarrow{-} W_{15}$					5	19	19
$W_{9} \xrightarrow{+} S_{15}$	.38*	.12	1.29*	1.09	10	22	20
$W_{9} \xrightarrow{-} S_{15}$					10	25	25
$S_{9} \xrightarrow{+} W_{15}$	.28*		.18*	1.02	6	11	13
$S_{9} \xrightarrow{-} W_{15}$		-.10			4	12	12
$H_{9} \xrightarrow{+} S_{15}$	.43*	.06	1.51*	.93	9	19	18
$H_{9} \xrightarrow{-} S_{15}$					11	29	32
$S_{9} \xrightarrow{+} H_{15}$	.44*		.18*	.80	0	5	6
$S_{9} \xrightarrow{-} H_{15}$		-.15			6	17	14

* $p < .05$

Note: In the above table, cross-lagged correlation coefficients, part correlation coefficients, beta weights, and F ratios are reported for CLC, PC, EC, and VC, respectively. For FSM, FCP, and MFCP, YC, YI, XC, XI cases are reported. In all instances, no YU or XU cases are found for FSM.

Table XXV. Significance of Chi-square Values for FSM, FCP, and MFCP in the Estimation of Causal Relationships Between Weight, Height, and Strength for Girls (N = 70)

Variables	Test	Significance of X^2		
		FSM	FCP	MFCP
Height and weight at ages 11 and 13	S	--	--	--
	D_1	--	--	--
	D_2	--	--	*
Weight and strength at ages 11 and 13	S	--	*	*
	D_1	--	--	--
	D_2	--	--	--
Height and strength at ages 11 and 13	S	--	--	--
	D_1	*	--	--
	D_2	--	--	--
Height and weight at ages 9 and 15	S	--	--	--
	D_1	--	--	--
	D_2	--	--	--
Weight and strength at ages 9 and 15	S	--	*	*
	D_1	--	--	--
	D_2	--	--	--
Height and strength at ages 9 and 15	S	*	*	*
	D_1	--	--	--
	D_2	*	*	--

* $p < .05$

Note: The first direction test (D_1) assumes that the first variable is the source of causal influence; the second direction test (D_2) assumes that the second variable is the source of causal influence. When the source test (S) is non-significant, the direction tests become irrelevant.

mutual congruent causal influence between weight and height. PC suggests that weight may have an incongruent influence on height while height may have a congruent influence on weight. The results obtained for VC and FSM are ambiguous. FCP and MFCP also fail to detect the source of the causal influence.

The causal relationship between weight and strength appears to be ambiguous. The results obtained for CLC and EC indicate ambiguous causal influence. PC, VC, FCP did not detect any causal influence. MFCP identifies weight as the cause of strength but fails to determine the direction of the influence.

Relatively more conclusive results are obtained with respect to the relationship between height and strength. PC, VC, and MFCP indicate that height is probably the cause of strength and the causal influence is congruent. FCP also identifies height as the source of causal influence but fails to determine the direction of such influence. While CLC and EC yield ambiguous causal estimates, the relative magnitudes of the causal estimates appear to be in agreement with the results obtained for PC, VC, and MFCP. Causal estimates yielded by FSM turn out to be ambiguous. See Tables XXII and XIII.

Results for the female sample

No major differences appear to exist between boys and girls with respect to the relationship of weight and height. The overall pattern of results suggests that the causal relationship is complex and inconsistent. Specifically, CLC and EC indicate that if there is any causal relationship between the two variables, the influence is congruent and mutual. The results yielded by the other models are ambiguous.

Pertaining to the relations between weight and strength, CLC and EC yield ambiguous estimates. So do PC and VC. FCP and MFCP, on the other hand, identify weight as the cause of strength. The direction of such influence is, however, uncertain. FSM fails to detect any causal influence.

Relative to height and strength, no clear pattern of causal relationships is revealed. CLC and EC yield ambiguous estimates. PC and VC fail to detect any causal influence. FSM, FCP, and MFCP also fail to identify the source of the causal influence.

The results described above are presented in Tables XXIV and XXV.

Application of decision rules

The correlations among the time-one measures of weight, height, and strength ranged from .78 ($p < .01$) to .45 ($p < .01$). All correlations are highly significant and the data can be described as a model B type.

The reliability of the data, which consist of measures of physical growth can be expected to be high--especially with respect to the weight and height measures. In terms of error ratio, the reliability level should be equal to or better than 9/1.

In general, three "agreement patterns" are obtained among the causal models. These are diagrammatically presented as follows:

Pattern A:

Model	CLC	PC	EC	VC	FSM	FCP	MFCP
Estimation	±	+	±	±	±	±	+

Pattern B:

Model	CLC	PC	EC	VC	FSM	FCP	MFCP
Estimation	±	±	±	±	±	±	±

Pattern C:

Model	CLC	PC	EC	VC	FSM	FCP	MFCP
Estimation	±	+	±	+	±	±	+

The pattern A results are highly similar to those obtained for the model B1 data set with an error ratio of 9/1 and an influence coefficient of .25 or -.25. As shown in Table IX in Chapter V the part correlation model (PC) and the modified-frequencies-of-change-in-product-moment model (MFCP) are the most efficient models for such data types.

Pattern B suggests that the data may best be described as a model B3 type with an error ratio of 9/1 and an influence coefficient of 1.0 or .50. For such data types, none of the causal models included in the present study appear to have a high degree of validity. See Table XIII in Chapter V.

The pattern C results are indicative of a B1 data type with an error ratio of 8/2 and an influence coefficient of .50. A perusal of Table IX in Chapter V suggests that the part correlation model (PC), the variance components model (VC), and the modified-frequencies-of-change-in-product moment model (MFCP) are most likely to yield correct estimates with this particular data type. It will be recalled that we have earlier assumed the data reliability to be higher than 9/1 in terms of error ratio. The use of an error ratio of 8/2 to describe the data may seem inconsistent with the assumption. It should, however, be remembered that all error ratios are indications of lower bound reliabilities. The difference between 9/1 and 8/1 in terms of error ratio may therefore be regarded as negligible.

Discussion

While the interrelationships among weight, height, and strength are no doubt complex and subtle, it appears possible to make some tentative causal interpretation on the basis of the results reported in the previous sections. Specifically, a number of causal models (i.e., PC, MFCP, VC) have been identified as the most efficient procedures for causal analysis with respect to the various data types under investigation. The conclusions reported in the following section are made on the basis of causal estimates yielded by these models.

A notable feature of the study is that the data were divided into two subsets with different causal intervals. This was done essentially for the purpose of determining the consistency of the results and to find out if different causal intervals would in fact provide different causal estimation. In this respect, the results show that the two different causal intervals do not make any remarkable difference in causal estimation. Another way to say this is, of course, that the overall pattern of results appears to be the same regardless of the magnitude of the causal intervals.

The data were also divided on the basis of sex. This was done to test the plausibility of sex being a "qualifier." As the results turn out, highly similar causal influences are detected for the male and female samples.

Relative to the efficiency of the various causal models, several interesting findings should be noted. First, there appears to be a remarkable similarity between CLC and EC. In all instances, results yielded by the two models lead to the same causal inferences. Second, in the majority of the cases FCP, FSM, and VC fail to detect any causal influence among the variables. This may or may not be a limitation of the models depending on whether or not causal influence actually exists among the variables in question. Third, relative to FSM, FCP, and MFCP, whenever the direction of the causal influence is said to be ambiguous, another interpretation appears plausible: the direction is both congruent and incongruent. Fourth, relative to CLC, EC, PC, and VC, some of the ambiguity with respect to the source of causal influence may be resolved by further investigating the difference between two corresponding causal estimates (e.g., $\hat{B}$ for X and $\hat{B}$ for Y). Finally, where causal influence is detected among the variables, it has been shown that PC, VC, and MFCP are most likely to yield the correct estimates.

Summary and conclusions

Measures of height, weight, and strength were obtained for a total of 136 boys and girls at ages 9, 11, 13, and 15. Weight and strength were measured in kilograms and height was measured in centimeters. Seven causal models were applied to the data to estimate causal influence that might exist among the variables. Decision rules formulated in Chapter VIII were used to interpret the results with respect to the source and direction of causal influence. Despite the differences in estimation yielded by the different models, the overall results appear to support the following tentative conclusions:

1. Relative to weight and height, the most plausible hypothesis appears to be that weight is the cause of height and the causal influence is incongruent.[1]
2. Relative to weight and strength, the most plausible hypothesis appears to be that regardless of the source of the causal influence (for which no conclusive evidence is obtained in the present study) the direction of such influence is likely to be both congruent and incongruent.
3. Relative to height and strength, the most plausible hypothesis appears to be that height is the cause of strength and the causal influence is congruent.
4. The causal interrelationships of weight, height, and strength appear to be the same for both boys and girls.
5. The above findings seem to be replicated by two sets of data involving different causal intervals.

Although the present study is not concerned with the possibility that the source of the causal influence may lie in variables not included in the study, it will be noted here that the general level of physical health may be the cause of all three variables (i.e., weight, height, and strength) under investigation. The amount of exercise and physical

---

[1]At first glance, this finding may seem perplexing. It is, however, conceivable that given the same genetic potentiality to grow tall, two individuals may in the end of their physical growth differ from each other with respect to height because of the weight factor. It is possible that the one who tends to gain weight faster than the average rate will end up being the shorter of the two. This finding does not preclude the possibility that height might have some congruent influence on weight. Such influence is perhaps less dominant than the incongruent influence that weight has on height.

training may be another cause. Incidentally, it appears possible to set up an experimental situation where the amount of exercise or training can be systematically manipulated. The results obtained from such an experiment can then serve as an empirical validation of the relevant causal hypotheses.

With respect to the performance of the various causal models, the following conclusions appear to be supported:

1. CLC and EC appear to be remarkably similar to each other. The two models consistently yield causal estimates which lead us to the same conclusions.
2. In general, FCP and FSM yield similar results. The former is, however, shown to be more sensitive than the latter.
3. PC, MFCP, and VC are shown to be models which are most likely to yield the correct causal interpretation with respect to the data types under investigation in this study.

CHAPTER XI

SUMMARY OF PART II

Part II of the dissertation discusses some of the problems that the researcher is likely to encounter in making causal interpretation when he is dealing with "real-life" data. An approach to resolving these problems is presented and applied to two sets of real data.

A crucial problem appears to be that the efficiency of the various causal models presently available to the researcher is relative to the kinds of data to which the models are applied. That is to say, some models are more efficient than others with certain types of data. The reverse may be true when other data types are involved. In addition, the concept of convergent validity (i.e., similarity of results yielded by different methods) does not seem to be an appropriate criterion on which to judge the validity of causal estimates.

In formulating a procedure to resolve these problems, the results obtained from the simulative evaluation of causal models reported in Part I are used as one of the criteria for identifying data types. This is described as the "agreement pattern" criterion. The other two criteria are data reliability and the correlation coefficient of the time-one measures. Once the data type is identified, the most efficient causal model with respect to the particular data type is determined on the basis of results obtained in Part I.

The results yielded by the procedure clearly show that while causal interpretation of real data is undoubtedly difficult, it is probably not beyond the reach of the researcher. The application of the procedure to the vocabulary-comprehension data appears to be highly effective in providing a means of making causal interpretation of the data. The method also seems to work very well with the weight-height-strength data.

From another point of view, it may be said that Part II of the dissertation serves to demonstrate the importance of prior knowledge about the data in conducting causal analysis. Specifically, knowledge of data reliability and the correlation of time-one measures contributes a great deal to the researcher's ability to identify the data type and consequently to determine the most efficient causal model for the data.

It should be pointed out, however, that the "agreement pattern" method for identifying data types is probably not entirely "foolproof." Inasmuch as an infinite number of data types may, and probably do, exist in the empirical world, it is highly possible that some data sets, while differing from each other with regard to the source and direction of causal influence, may yield the same "agreement pattern" for the causal models. An erroneous identification of data types may thus result. This problem can, however, be overcome by future research efforts aimed at creating additional data types to investigate the similarity of "agreement patterns" in relation to the data types.

PART THREE

CONCLUDING REMARKS

CHAPTER XII

GENERAL DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Part III concludes the dissertation with (1) a general discussion of the salient features of the results reported in Part I and Part II, (2) conclusions drawn on the basis of these results, and (3) recommendations with respect to causal analysis and possible directions for future research.

(1) General Discussion

Similarity among models

There appears to be a rather low degree of similarity among the various causal models. CLC, PC, and EC are similar only with respect to data where the time-one measures are uncorrelated. The results yielded by these models with respect to model B1 and model B2 data are quite different. Relative to model B3, CLC and EC yield almost the same results. The results obtained for PC, however, appear to be quite different.

Except for model B3 data, where the model appears to be highly similar to CLC and EC, VC appears to be independent of the other models.

This finding supports the expectation that the causal models would probably have different levels of efficiency with different types of data. Some models are more efficient with data where congruent and incongruent influences co-exist, while others appear to yield better estimates when the causal influence is either congruent or incongruent. In fact, a primary purpose of the present study is to determine the differential efficiency of the models and to provide appropriate guidelines for the proper use of these models.

Yee (1966) claims that the FSM and FCP are highly similar to each other and in most instances may be used interchangeably. The claim is only partially supported by the results of the present investigation. With respect to model A1 data, the patterns of results yielded by the two models are by no means similar. Dissimilarities between the two models are also shown with respect to model B1 data and to a lesser extent, model B2 data. The results obtained for model B3 data also appear to suggest that the two models are in fact different. Only with respect to model A2 and model A3 data do the two models show a relatively high degree of similarity. Even here, it will be noted, the results are not identical.

It will also be noted that where individual YI, YC, XI, XC cases are concerned, the two models are vastly different. In the first place, FCP does not make allowance for YU and XU cases while FSM does. The number of YC, YI, XC, or XI cases estimated by FSM is invariably much smaller than that estimated by FCP. In other words, a great proportion of the individual cases are put into entire different causal categories by the two models.

As expected, MFCP bears little similarity to FCP. The highest similarity between the two models is seen in the results yielded for model A2 data, where eight out of 10 causal estimates are identical. For the rest of the data, the two models appear to be vastly different. Needless to say, where individual cases are concerned, MFCP and FCP are also shown to be different from each other.

All this, again, points to the need for some guidelines with respect to the use of FSM, MFCP, and FCP as procedures for causal analysis.

Congruent and incongruent causal influence

Yee (1966) claims that FSM and FCP are capable of identifying congruent and incongruent causal influences. The results of the present investigation seem to suggest that when both congruent and incongruent influences are present in the same data set, the source of the causal influence is, in general, identified by the two models. In many cases, however, the models fail to detect the direction of such influences.

It should also be noted that for data sets where either congruent or incongruent causal influence is present (e.g., models A1, A2, B1, B2), erroneous YC or YI cases are estimated. Thus, in most instances, results yielded by FSM and FCP actually suggest that both congruent and incongruent influences exist to some extent.

It may be added, incidentally, that when both congruent and incongruent causal influences are present in the same data set, two general laws or principles are indicated: one congruent and the other incongruent --and should be identified separately. To cite a concrete example, it may be said that when positive teacher attitude is followed by positive student attitude in some cases and negative student attitude in others, it seems obvious that other factors (e.g., family background, socio-economic status) have entered into the picture to produce such results. These factors should be identified and used as "qualifiers" to divide the data into two subsets each of which contains either congruent or incongruent causal influence.

Data reliability

On the basis of causal estimates pertaining to the source and direction of causal influence, the behavior of error may appear to be quite inconsistent: while the overall effects appear to be negative, large errors, in some cases, may actually enhance the estimation efficiency of some models. However, if we consider only the magnitudes of individual causal estimates (e.g., beta weights, F ratios) the effects of error appear to be quite consistent and in the expected direction. Error is consistently shown to lower the magnitudes of these estimates.

The seeming contradiction is readily reconciled if we realize that a correct causal estimate (with respect to the source and direction of causal influence) depends not only on the magnitude of the estimate pertaining to variable Y as the source variable but also on the magnitude of the estimate pertaining to variable X as the source variable. Where error appears to have differential effects on X and Y, quite unexpected results may be obtained with respect to the estimation of the source and direction of causal influence.

With respect to the magnitudes of the causal estimates, FSM appears to be least affected by error while PC in some instances is very substantially affected by error. The effects of error on other models range from mild to moderate.

<u>Amount of causal influence</u>

The amount of causal influence is an important factor of causal estimation. Its effects, however, differ with each type of data. With model A1 data, strong causal influence usually results in better estimates of the source and direction of causal influence. For some models (e.g., MFCP, FCP) incongruent causal influence seems to be more difficult to detect than congruent influence.

The tendency for causal models to yield better estimates with data of strong causal influence is also found with respect to model A3 data. Here, MFCP appears to be the only exception; the model yields ambiguous estimates for data sets with strong causal influence and correct estimates with data of weak causal influence. As noted earlier, this unexpected pattern of results is probably due to the model's tendency to overclassify YI cases.

The results obtained for model B1 data show little consistency with respect to the effects of the amount of causal influence on estimation. Strong causal influence does seem to enhance the estimation power of FSM, FCP, and MFCP. For other models, the difference in estimation power appears to be attributable to the direction of causal influence rather than the amount of influence. These effects are most pronounced with respect to PC and VC. To a lesser degree the effects are also found with respect to CLC and EC.

Inasmuch as most of the causal models yield ambiguous estimates with respect to model B3 data sets, little need be said about the effects of the amount of influence on estimation power. It is of interest to note, however, that for the two models which do yield some correct causal estimates (i.e., PC and FCP) strong causal influence is actually shown to have negative effects on estimation power.

It should be pointed out that the above discussion pertains only to estimates of the source and direction of causal influence. The effects of the amount of influence on the magnitudes of estimates (e.g., beta weights, F ratios) appear to be quite consistent and in the expected direction. That is to say, the magnitudes of the causal estimates decline as the causal influence weakens. Conversely, the magnitudes of the causal estimates increase as the causal influence increases. There also seems to be a tendency for the magnitudes to drop more substantially when the influence coefficient declines from 1.0 to .50 than when the coefficient decreases from .50 to .25.

In some instances (e.g., when MFCP is applied to model A3 data) there appears to be some interaction between the amount of causal influence and error. However, the positive effects of strong causal influence on the magnitude of causal estimate are still clearly discernible.

Ambiguous and erroneous causal estimates

A word should be said about what has been described as ambiguous causal estimates. In some cases, the ambiguity lies with the source of the causal influence. In others, it pertains to the direction of the influence. At any rate, these ambiguous estimates are actually equivalent to erroneous estimates in the sense that they are different from the true causal parameters.

The decision rules used in the present study are rather liberal. Only causal estimates which lead us to conclusions opposite to the truth are considered erroneous.

Identification of data types

A crucial part of making causal interpretation with real data is the identification of the type of data with which the researcher is dealing. Only when the data type is identified in terms of the simulated data sets is the researcher able to determine the most efficient causal model for the data. The identification of data types is based on the following three criteria: (1) correlation of the time-one measures, (2) data reliability, and (3) the "agreement pattern" yielded by the causal models. This procedure is shown to have worked very effectively with the two sets of real data included in the present study. It should be cautioned, however, that the "agreement pattern" criterion may not be entirely "foolproof" and probably warrants further investigation and refinement.

This aspect of causal analysis (i.e., the identification of data types) fully demonstrates the importance of prior knowledge about the data. Part of this knowledge (e.g., data reliability) can be gained directly from the empirical data. Other prior knowledge may be obtained through previous research findings or what is generally known about the phenomenon being studied. It is, for instance, difficult to conceive of vocabulary ability having incongruent influence on comprehension ability. One may thus conclude that the data type in question is probably not an incongruent data type. At any rate, such knowledge, gained directly from the data or otherwise, contributes a great deal to the researcher's ability to identify the data type and consequently to determine the most efficient causal model for the data type.

(2) Conclusions

In general, the results obtained in the present study appear to support the following conclusions:

I. General conclusions

1. The estimation power of the various causal models reported in the literature appears to be highly exaggerated.
2. The efficiency of FSM, FCP, and MFCP is limited to situations where these models are used as group estimation procedures. If

they are used to identify individual causal cases, the efficiency level will decline sharply.

3. Causal data where the time-one measures consist of more than one element and are correlated appear to be more difficult to handle than causal data where the time-one measures consist of only one element and are uncorrelated.

4. Relative to data where both congruent and incongruent causal influences are present, causal analysis appears to be extremely difficult.

II. Specific conclusions

A. Relative to the efficiency of the models

1. With respect to model A1 data, the most efficient causal model appears to be PC. CLC and EC may also be applied to this type of data. FSM, FCP, and MFCP are much less effective.

2. With respect to model A2 data, all causal models are shown to have a high level of efficiency.

3. With respect to model A3 data, only FSM and FCP appear to be effective. The rest of the models are almost entirely ineffective.

4. With respect to model B1 data, the most efficient model appears to be PC. The use of other causal models on this type of data may result in erroneous causal inferences.

5. With respect to model B2 data, only FSM and FCP are shown to have moderately high efficiency in identifying causal influence.

6. With respect to model B3 data, none of the causal models appear to have the capability of identifying the causal influence.
7. PC, FSM, and FCP appear to have the highest level of overall efficiency.
8. It seems abundantly clear that the efficiency of the causal models is relative to the nature of the data to which the causal models are applied.
9. It follows from the above that the best causal estimates can be obtained only if the researcher has some prior knowledge of the nature of the data.

B. Relative to data reliability

1. Although error appears to have consistently negative effects on the magnitudes of causal estimates, its effects on the estimates of the source and direction of causal influence appear to be much less consistent. In some cases, error may serve to enhance such estimates.
2. The effects of error appear to be most noticeable with respect to PC and least noticeable with respect to FSM.
3. While the effects of error are mild in general, there appear to be instances where such effects turn out to be a decisive factor in correct causal estimation.

C. Relative to the amount of causal influence

1. The amount of causal influence appears to be positively related to the magnitudes of causal estimates. That is to say, a greater amount of causal influence generally results in a causal estimate of greater magnitude.

2. The estimation of the source and direction of causal influence is also affected by the amount of causal influence. Such effects are, however, rather inconsistent. In some cases, a strong causal influence may actually have negative effects on estimation.

3. In general, incongruent causal influence appears to be more difficult to detect than congruent causal influence.

D. Relative to real data

1. Causal interpretation of real data has been shown to be difficult but probably not beyond the reach of the researcher.

2. Inasmuch as the causal models are shown to have differential efficiency with various types of data, a crucial aspect of making causal interpretation with real data is the identification of data types and thence the determination of the most efficient causal model for the data.

3. The procedure proposed in this dissertation for the identification of data types appears to work very well with real data and probably deserves general application.

4. In the absence of any prior knowledge, PC and FCP appear to be the best "risks" for conducting causal analysis with panel data.

(3) Recommendations

1. The researcher should strive to gain some knowledge of the phenomenon under investigation before conducting any causal analysis.

2. More than one causal procedure should be used to analyze the data to

attain validity of causal interpretation.

3. If both congruent and incongruent causal influences can reasonably be expected to exist, the data should be divided into two or more subsets on the basis of possible "qualifiers." Each subset should then be analyzed separately.

4. Whenever possible, the results of causal analysis should be replicated by experimentally manipulating the cause factor in an empirical situation.

5. In the present study, the correctness of the estimation of the source and direction of causal influence is determined on the basis of the statistical significance of the corresponding causal estimates. The actual difference between these estimates is not considered. Such differences may be highly useful in the determination of the source of causal influence. An investigation of the significance (statistical or otherwise) of these differences may be highly valuable.

6. Inasmuch as the "agreement pattern" method is not entirely "foolproof," more causal data types should be investigated in the future to reduce the probability of erroneous identification of data types.

APPENDICES

Appendix A

Skewness and Kurtosis of Variables in Model A1 Data Sets

	Skewness				Kurtosis			
Data Set	X_1	X_2	Y_1	Y_2	X_1	X_2	Y_1	Y_2
1	-.15	-.13	.06	-.04	-.48	-.29	.74	.58
2	-.15	.02	.06	-.04	-.48	.37	.74	.58
3	-.15	-.16	.06	-.04	-.48	-.50	.74	.58
4	-.15	-.01	.06	-.04	-.48	-.15	.74	.58
5	-.15	-.15	.06	-.04	-.48	-.59	.74	.58
6	-.15	-.05	.06	-.04	-.48	-.43	.74	.58
7	-.18	-.12	.16	-.07	-.34	-.28	.87	.52
8	-.18	.02	.16	-.07	-.34	.31	.87	.52
9	-.18	-.15	.16	-.07	-.34	-.49	.87	.52
10	-.18	-.02	.16	-.07	-.34	-.15	.87	.52
11	-.18	-.14	.16	-.07	-.34	-.56	.87	.52
12	-.18	-.06	.16	-.07	-.34	-.40	.87	.52
13	-.19	-.11	.28	-.10	-.15	-.25	.98	.44
14	-.19	.01	.28	-.10	-.15	.25	.98	.44
15	-.19	-.14	.28	-.10	-.15	-.42	.98	.44
16	-.19	-.03	.28	-.10	-.15	-.10	.98	.44
17	-.19	-.14	.28	-.10	-.15	-.44	.98	.44
18	-.19	-.07	.28	-.10	-.15	-.28	.98	.44
19	-.14	-.11	.41	-.15	.03	-.21	1.01	.36
20	-.14	-.00	.41	-.15	.03	.23	1.01	.36
21	-.14	-.15	.41	-.15	.03	-.29	1.01	.36
22	-.14	-.06	.41	-.15	.03	.02	1.01	.36
23	-.14	-.16	.41	-.15	.03	-.25	1.01	.36
24	-.14	-.10	.41	-.15	.03	-.08	1.01	.36
25	-.03	-.12	.50	-.21	.12	-.18	.93	.29
26	-.03	-.03	.50	-.21	.12	.24	.93	.29
27	-.03	-.17	.50	-.21	.12	-.17	.93	.29
28	-.03	-.10	.50	-.21	.12	.17	.93	.29
29	-.03	-.18	.50	-.21	.12	-.08	.93	.29
30	-.03	-.14	.50	-.21	.12	.12	.93	.29

Appendix B

Skewness and Kurtosis of Variables in Model A2 Data Sets

	Skewness				Kurtosis			
Data Set	X_1	X_2	Y_1	Y_2	X_1	X_2	Y_1	Y_2
1	-.16	.04	.21	.23	-.62	-.25	-.23	-.19
2	-.06	.04	.17	.21	-.61	-.15	-.26	-.25
3	.06	.06	.15	.18	-.54	-.03	-.26	-.32
4	.15	.12	.16	.13	-.44	.06	-.30	-.40
5	.14	.22	.18	.05	-.44	.04	-.44	-.47
6	-.16	-.55	.21	.23	-.62	.31	-.22	-.19
7	-.06	-.55	.17	.21	-.61	.33	-.26	-.25
8	.06	-.50	.15	.18	-.54	.33	-.26	-.32
9	.15	-.36	.16	.13	-.44	.29	-.30	-.40
10	.14	-.13	.18	.05	-.44	.24	-.44	-.47

Appendix C

Skewness and Kurtosis of Variables in Model A3 Data Sets

	Skewness				Kurtosis			
Data Set	X_1	X_2	Y_1	Y_2	X_1	X_2	Y_1	Y_2
1	-.15	-.05	.06	-.04	-.48	-.82	.74	.58
2	-.15	-.08	.06	-.04	-.48	-.51	.74	.58
3	-.15	-.10	.06	-.04	-.48	-.46	.74	.58
4	-.18	-.06	.16	-.07	-.34	-.81	.87	.52
5	-.18	-.08	.16	-.07	-.34	-.49	.87	.52
6	-.18	-.10	.16	-.07	-.34	-.43	.87	.52
7	-.18	-.06	.28	-.10	-.15	-.78	.98	.44
8	-.18	-.09	.28	-.10	-.15	-.44	.98	.44
9	-.18	-.11	.28	-.10	-.15	-.34	.98	.44
10	-.14	-.07	.41	-.15	.03	-.72	1.01	.36
11	-.14	-.11	.41	-.15	.03	-.34	1.01	.36
12	-.14	-.13	.41	-.15	.03	-.18	1.01	.36
13	-.03	-.09	.50	-.21	.12	-.62	.93	.29
14	-.03	-.13	.50	-.21	.12	-.20	.93	.29
15	-.03	-.16	.50	-.21	.12	-.01	.93	.29

Appendix D

Skewness and Kurtosis of Variables in Model B1 Data Sets

	Skewness				Kurtosis			
Data Set	X_1	X_2	Y_1	Y_2	X_1	X_2	Y_1	Y_2
1	-.03	.20	.35	.43	-.06	.23	.29	.43
2	-.03	-.04	.35	.43	-.06	-.36	.29	.43
3	-.03	.06	.35	.43	-.06	.28	.29	.43
4	-.03	-.05	.35	.43	-.06	-.30	.29	.43
5	-.03	.00	.35	.43	-.06	.21	.29	.43
6	-.03	-.05	.35	.43	-.06	-.15	.29	.43
7	-.06	.19	.24	.42	-.20	.32	.15	.43
8	-.06	-.00	.24	.42	-.20	-.45	.15	.43
9	-.06	.06	.24	.42	-.20	.29	.15	.43
10	-.06	-.02	.24	.42	-.20	-.42	.15	.43
11	-.06	.00	.24	.42	-.20	.15	.15	.43
12	-.06	-.03	.24	.42	-.20	-.27	.15	.43
13	-.11	.20	.13	.36	-.22	.44	-.00	.33
14	-.11	.00	.13	.36	-.22	-.51	-.00	.33
15	-.11	.08	.13	.36	-.22	.29	-.00	.33
16	-.11	-.01	.13	.36	-.22	-.48	-.00	.33
17	-.11	.03	.13	.36	-.22	.10	-.00	.33
18	-.11	-.01	.13	.36	-.22	-.33	-.00	.33
19	-.17	.21	.02	.28	-.15	.55	-.10	.16
20	-.17	-.02	.02	.28	-.15	-.47	-.10	.16
21	-.17	.10	.02	.28	-.15	.32	-.10	.16
22	-.17	-.02	.02	.28	-.15	-.39	-.10	.16
23	-.17	.05	.02	.28	-.15	.14	-.10	.16
24	-.17	-.01	.02	.28	-.15	-.25	-.10	.16
25	-.20	.21	-.04	.19	-.10	.65	-.11	-.03
26	-.20	-.06	-.04	.19	-.10	-.30	-.11	-.03
27	-.20	.11	-.04	.19	-.10	.41	-.11	-.03
28	-.20	-.03	-.04	.19	-.10	-.16	-.11	-.03
29	-.20	.06	-.04	.19	-.10	.26	-.11	-.03
30	-.20	-.01	-.04	.19	-.10	-.04	-.11	-.03

Appendix E

Skewness and Kurtosis of Variables in Model B2 Data Sets

	Skewness				Kurtosis			
Data Set	X_1	X_2	Y_1	Y_2	X_1	X_2	Y_1	Y_2
1	-.06	.34	-.01	-.01	-.59	.09	-.25	-.23
2	-.15	.21	-.06	-.00	-.45	.09	-.17	-.23
3	-.05	.01	-.14	.05	-.08	.10	-.02	-.24
4	.17	-.19	-.21	.10	.12	.11	.09	-.29
5	.33	-.28	-.24	.11	.04	.08	.06	-.34
6	-.06	-.22	-.01	-.01	-.59	-.24	-.25	-.23
7	-.15	-.26	-.06	-.00	-.45	-.19	-.17	-.23
8	-.05	-.29	-.14	.05	-.08	-.07	-.02	-.24
9	.18	-.30	-.20	.10	.12	.09	.09	-.29
10	.33	-.24	-.24	.11	.04	.20	.06	-.34

Appendix F

Skewness and Kurtosis of Variables in Model B3 Data Sets

	Skewness				Kurtosis			
Data Set	X_1	X_2	Y_1	Y_2	X_1	X_2	Y_1	Y_2
1	-.03	.05	.35	.43	-.06	-.61	.29	.43
2	-.03	.03	.35	.43	-.06	-.14	.29	.43
3	-.03	-.01	.35	.43	-.06	.03	.29	.43
4	-.06	.05	.24	.42	-.20	-.60	.15	.43
5	-.06	.03	.24	.42	-.20	-.17	.15	.43
6	-.06	-.00	.24	.42	-.20	-.04	.15	.43
7	-.11	.05	.13	.37	-.22	-.51	-.00	.33
8	-.11	.03	.13	.37	-.22	-.14	-.00	.33
9	-.11	.01	.13	.37	-.22	-.09	-.00	.33
10	-.17	.04	.03	.28	-.15	-.32	-.10	.16
11	-.17	.03	.03	.28	-.15	-.03	-.10	.16
12	-.17	.02	.03	.28	-.15	-.03	-.10	.16
13	-.20	.03	-.04	.19	-.10	-.05	-.11	-.03
14	-.20	.03	-.04	.19	-.10	.14	-.11	-.03
15	-.20	.02	-.04	.19	-.10	.13	-.11	-.03

Appendix G

Skewness and Kurtosis of Gates-MacGinitie Test Scores for Second and Third Graders

	Skewness				Kurtosis			
Group	Voc_1	Voc_2	$Comp_1$	$Comp_2$	Voc_1	Voc_2	$Comp_1$	$Comp_2$
Gr. 2 and Gr. 3	.63	.18	1.22	.52	-.31	-1.13	.82	-.27
Gr. 2	.62	.03	1.34	.23	-.52	-1.37	.93	-.96
Gr. 3	.63	.35	.93	.49	-.19	- .81	-.08	-.55

Appendix H

Skewness and Kurtosis of Measures of Height for Ages 9, 11, 13, and 15

	Skewness				Kurtosis			
Group	H_9	H_{11}	H_{13}	H_{15}	H_9	H_{11}	H_{13}	H_{15}
Total	-.04	.09	-.02	.44	.36	.46	-.25	.23
Male	.10	.11	.31	.27	-.21	-.44	-.79	-.44
Female	-.15	.02	-.55	-.09	.74	.93	.92	.33

Appendix I

Skewness and Kurtosis of Measures of Weight
for Ages 9, 11, 13, and 15

	Skewness				Kurtosis			
Group	W_9	W_{11}	W_{13}	W_{15}	W_9	W_{11}	W_{13}	W_{15}
Total	1.85	1.24	.83	1.16	7.83	2.79	1.45	3.38
Male	2.99	2.09	1.30	1.20	14.51	8.59	3.27	3.66
Female	.57	.69	.40	1.07	-.17	.13	.08	2.75

Appendix J

Skewness and Kurtosis of Measures of Strength
for Ages 9, 11, 13, and 15

	Skewness				Kurtosis			
Group	S_9	S_{11}	S_{13}	S_{15}	S_9	S_{11}	S_{13}	S_{15}
Total	.41	.64	.61	.46	1.14	1.33	.91	-.52
Male	.52	.54	.42	-.25	1.69	2.41	.21	-.45
Female	.45	.81	.72	.45	1.04	.71	2.20	1.57

REFERENCES

REFERENCES

Blalock, H. M. "Correlational analysis and causal inferences." American Anthropologist, 62(1970), 624-631.

Blalock, H. M. "Correlation and causality: the multivariate case." Social Forces, 39(1961a), 246-251.

Blalock, H. M. Causal Inferences in Non-experimental Research. Chapel Hill: The University of North Carolina Press, 1961b.

Blalock, H. M. "Four-variable causal models and partial correlations." American Journal of Sociology, 68(1962a), 182-194.

Blalock, H. M. "Further observations on asymmetric causal models." American Sociological Review, 27(1962b), 542-545.

Blalock, H. M. "Some implications of random measurement error for causal inferences." American Journal of Sociology, 71(1965), 37-47.

Blalock, H. M. "Causal inferences, closed populations, and measures of association." The American Political Science Review, 61(1967), 130-136.

Blalock, H. M. Theory Construction. Englewood Cliffs: Prentice-Hall, 1969.

Buros, O. K. (Ed.) The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1972.

Campbell, D. T. "From description to experimentation: interpreting trends as quasi-experiments." In C. W. Harris (Ed.) Problems in Measuring Change. Madison: Univ. of Wisconsin Press, 1963, 212-242.

Campbell, D. T. and K. N. Clayton. "Avoiding regression effects in panel studies of communication impact." In E. L. Uliassi (Ed.) Studies in Public Communication. Chicago: The University of Chicago, 1961, 99-118.

Campbell, D. T. and J. C. Stanley. "Experimental and quasi-experimental designs for research on teaching." In N. L. Gage (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally, 1963, 171-246.

Crano, W. D., D. A. Kenny, and D. T. Campbell. "Does intelligence cause achievement?: a cross-lagged panel analysis." Journal of Educational Psychology, 63 (1972), 258-275.

Darlington, R. B. "Multiple regression in psychological research and practice." Psychological Bulletin, 69 (1968), 161-182.

Eron, L. D., M. M. Lefkowitz, L. R. Huesmann, and L. O. Walder. "Does television violence cause aggression?" American Psychologist, 27 (1972), 253-263.

Feshback, S. and R. D. Singer. Television and Aggression. San Francisco: Jossey-Bass, 1971.

Gulliksen, H. Theory of Mental Tests. New York: Wiley and Sons, 1950.

Hume, D. Enquiries Concerning the Human Understanding (2nd Ed.). Oxford: Clarendon Press, 1902.

IBM. Random Number Generation and Testing. New York: IBM, 1959.

IBM. System/360 Scientific Subroutine Package. New York: IBM, 1968.

Lazarsfeld, P. F. Mutual effects of statistical variables. New York: Columbia University, Bureau of Applied Social Research. 1947.

Lipset, S. M., P. F. Lazarsfeld, A. H. Barton, and J. Linz. "The psychology of voting: an analysis of political behavior." In G. Lindzey (Ed.) Handbook of Social Psychology, Vol. 2. Cambridge, Mass: Addison-Wesley, 1954, 1124-1175.

Linn, R. L. and C. E. Werts. "Assumptions in making causal inferences from part correlations, partial correlations, and partial regression coefficients." Psychological Bulletin, 72 (1969), 307-310.

McNemar, Q. Psychological Statistics. New York: Wiley and Sons, 1969.

Pels, D. C. and F. M. Andrews. "Detecting causal priorities in panel study data." American Sociological Review, 29 (1964), 836-848.

Polk, K. "A note on asymmetric causal models." American Sociological Review, 27 (1962), 539-542.

Portwood, C. S. Foundations of Mathematical Epistemology: a Derivation of Causal Statistics. Unpublished doctoral dissertation, University of California, Berkeley, 1972.

Richards, J. M. "A simple analytic model for college effects." The School Review, 74 (1966), 380-392.

Rozelle, R. M. and D. T. Campbell. "More plausible rival hypotheses in the cross-lagged panel correlation technique." Psychological Bulletin, 71 (1969), 74-80.

Simon, H. A. "Spurious correlation: a causal interpretation." Journal of the American Statistical Association, 49 (1954), 467-479.

Tuddenham, R. D. and M. M. Synder. Physical Growth of California Boys and Girls from Birth to Eighteen Years. Berkeley: University of California Press, 1954.

Tukey, J. W. "Causation, regression, and path analysis." In O. Kempthorne, et al. (Eds.) Statistics and Mathematics in Biology. Ames: Iowa State College Press, 1954, 35-66.

Werts, C. E. and D. J. Watley. "Analyzing college effects: correlation vs. regression." _American Educational Research Journal_, 5 (1968), 585-598.

Werts, C. E. and R. L. Linn. "Analyzing school effects: how to use the same data to support different hypotheses." _American Educational Research Journal_, 6 (1969), 439-447.

Wold, H. "Causality and econometrics." _Econometrica_, 22 (1954), 162-177.

Wold, H. "Causal inference from observational data: a review of ends and means." _Journal of the Royal Statistical Society_, A 119 (1956), 28-61.

Wonnacott, R. J. and T. H. Wonnacott. _Econometrics_. New York: Wiley and Sons, 1970.

Yee, A. H. _Factors involved in determining the relationship between teachers' and pupils' attitudes_. U. S. Department of Health, Education, and Welfare, Office of Education Research Project No. 5-8346, Austin: The University of Texas, 1966.