CAUSAL RELATIONS AND STRUCTURAL MODELS*

Richard Juster, Survey Research Center, University of California, Berkeley

1. Introduction

Following early work by Simon (1954), a method of testing the adequacy of certain a priori assumptions of non-experimental causal models has gained considerable acceptance among social investigators.[1] This approach to empirical theory construction has been developed further by Blalock (1962, 1964), Alker (1965, 1966), and Boudon (1968), among others, and is generally referred to as "Simon-Blalock causal modeling."

Simon and others who have contributed to this approach to theory building have limited their interest almost exclusively to recursive-form linear structural models and have been concerned with implications of such models in so far as they are representations of "one-way" or "non-reciprocal" causal relations among sets of variables.[2] In his classic paper, Simon (1954) derives necessary and sufficient conditions (which can be empirically approximated) for specified structural coefficients in the three-equation recursive-form linear model to be zero. Assuming that the dependent variable in each structural equation is subject to change according to the causal laws postulated by the equation, he interprets his statistical results to be a test of causal relations represented in a recursive-form linear model.

The purpose of the present paper is two-fold. First, after a review of Simon's work, we shall elucidate some of the inherent weaknesses of the recursive-form linear structural model for the representation of causal processes and shall argue against its use, even when it seems an appropriate choice for such a representation. A case will be made for the employment of linear structural systems which are less restrictive than the recursive-form in two important respects. Specifically, we are interested in linear structural systems which (1) have the ability to represent reciprocal causal relations and (2) require weaker restrictions on the covariances of the disturbance terms than is required for the recursive-form system. Second, we shall derive correlational conditions which are both necessary and sufficient for specified structural coefficients to be zero in a type of linear structural model which is more general than the recursive-form in the two respects just mentioned. As does Simon for the recursive-form system, we shall take our statistical results to be a test of causal relations in this more general structural model.

Greatly simplifying Simon's paper (1954), we shall try to capture the most important implications of his efforts. In an earlier piece (1953), Simon formally develops the concepts of causality and the causal relation.[3] Offering an implicit definition of the causal relation as an asymmetrical relation between two variables, Simon points out that the temporal sequence of the variables is not the basis of the asymmetry which defines the causal relation. In fact, his implicit definition admits relations where no temporal sequence even appears. He argues that such a definition corresponds more closely to the consensual scientific usage of the concept than does a definition which employs time sequence as the basis of the asymmetry between two variables. At the heart of Simon's notion of the causal relation is his concern with a "production" or "influence" operation. A mechanistic relation exists between two variables whereby impulses from one variable influence behavior in the other. It is this influence operation which forms the basis of the asymmetry that defines the causal relation. For example: "a thrown rock produces a broken window" is a relationship which exhibits an operational relation providing the basis of an asymmetry between two variables. While a temporal sequence is certainly involved in such a relationship, it is not the sequence which forms the basis of the asymmetry.

Yet, Hume argues (and Simon agrees) that the only relationship we can observe between two variables is a "constant conjunction" in the past (an association). Since an association is all we are able to observe, it is impossible for us to establish a necessary connection or ontological relationship between a cause and an effect. Only in a tentative sense are we able to establish a causal relationship. Hence, it makes little sense to affirm that some variable is the "true" cause of some other variable. In our example, for instance, we could never demonstrate that a thrown rock really causes a broken window. We can never dismiss the possibility that other variables exist which would "explain away" our causal relationship. Nor can we avoid the possibility that the analytic level chosen by another investigator to explain a given empirical process is different than the level we have chosen -- we may choose a "microscopic" level of analysis, while someone else may choose a "macroscopic" level. Even a controlled experiment cannot demonstrate conclusively a "real" causal relationship between two variables.

However, it is possible to view the causal relation in a manner which does not violate Hume's argument. We may take a "subjective" view of the causal relation and argue that whether or not an empirical relationship is causal depends on the context in which we make our description. While it makes little sense to argue that some process is unconditionally causal, it is nevertheless consistent with the Humean view that a relationship may be called causal in a conditional sense. When we argue that a relationship is causal, we are exhibiting a perception of (or hypothesis about) the empirical world where we have made, either explicitly or implicitly, certain ceteris paribus assertions about other empirical relations. We are making causal statements about an abstraction of the real world, not about the real world itself. For example, it is necessary to consider the empirical process we are interested in as practically isolated from the rest of the world. However, a truly isolated process is almost never encountered in the real world nor are many of the other ceteris paribus conditions which yield the context in which we call a relationship causal. But, if we are willing to assume that these ceteris paribus conditions are satisfied and confine our causal descriptions to abstractions of the empirical world, then Hume's critique becomes irrelevant.

It is for the above reasons that Simon restricts his formal definition of the causal relation to refer only to models of empirical processes rather than to the empirical processes themselves. Inspired by his implicit notion of the causal relation as an asymmetrical production operation, he formally defines the causal relation within the context of a non-stochastic "segmentable" system of linear non-homogeneous equations -- segmentable in the sense that the current predetermined variables are determined by equations which do not contain any of the current endogenous variables.[4] A paraphrase of Simon's formal definition of the direct causal relation, X causes Y, is: X directly causes Y if X appears as a predetermined variable in the equation for Y, when Y is currently endogenous in a segmentable system of linear non-homogeneous equations.[5] This formal definition provides the basis for Simon's later paper (1954), to which we now turn our attention. However, before proceeding to our discussion, we shall digress to introduce an important preliminary to the remainder of this paper.

The recursive-form linear structural model is that model in which the structural disturbances are independent of each other, and the matrix of coefficients of the endogenous variables has only zeros to one side of the main diagonal such that:

$$
\begin{array}{lrcl}
 & Y_1 & = & \sum_{k=1}^{K} \gamma_{1k}X_k + u_1 \\
 & \beta_{21}Y_1 + \quad Y_2 & = & \sum_{k=1}^{K} \gamma_{2k}X_k + u_2 \\
(1.1) & \vdots & & \vdots \qquad \vdots \\
 & \beta_{G1}Y_1 + \beta_{G2}Y_2 + \cdots \; Y_G & = & \sum_{k=1}^{K} \gamma_{Gk}X_k + u_G
\end{array}
$$

where $Y_1,\ldots,Y_G$ are variables endogenous to the whole system, $X_1,\ldots,X_K$ are predetermined variables, and $u_1,\ldots,u_G$ are random disturbances with

$$E(u_g) = 0, \quad E(u_g u_{g'}) = \begin{cases} \phi_{gg}, & g' = g \\ 0, & g' \neq g, \end{cases}$$

and u_g is independent of $Y_1,\ldots,Y_{g-1},X_1,\ldots,X_K$, for all g, $1 \leq g \leq G$, so that in the gth equation there is only one endogenous variable, Y_g, all the variables $Y_1,\ldots,Y_{g-1}$ being predetermined by equations preceding in the hierarchy, and $X_1,\ldots,X_K$ being predetermined for the entire model.[6] Hence, an obvious property of such a system is that influences move only in one direction: that is, from equations earlier in the hierarchy to those appearing later. Moreover, it can be shown that every equation in such a system is identified.[7] More will be said about the recursive-form structural system after we complete our discussion of Simon's contribution to theory building.

While Simon has made an important contribution to the methodology of scientific inquiry simply on the basis of his formal definition of the causal relation, by no means is this the extent of his contribution. We shall now consider Simon's concern with the derivation of necessity and sufficiency conditions for the adequacy of specific three-variable causal models.

In order to carry out his derivation, Simon employs a three-equation recursive-form linear structural model to represent the causal relations among a set of three variables. From a close examination of the properties of system (1.1), it should be obvious that the recursive-form is segmentable and an appropriate choice for the representation of causal relations as defined by Simon in his earlier piece (1953). The recursive-form linear structural model is a stochastic analogue of the non-stochastic segmentable system employed by Simon in his development of the formal definition of the causal relation. Moreover, the direct causal relation, Y_i causes Y_j, is represented in the recursive-form linear model by the non-zero structural coefficient β_{ji} (i.e. $\beta_{ji} \neq 0$).

Assuming a system of three endogenous variables, each variable being determined by a linear mechanism so that the model has three equations, Simon represents this situation as:

$$
\begin{aligned}
Y_1 + \beta_{12}Y_2 + \beta_{13}Y_3 &= u_1 \\
\beta_{21}Y_1 + Y_2 + \beta_{23}Y_3 &= u_2 \qquad (1.2) \\
\beta_{31}Y_1 + \beta_{32}Y_2 + Y_3 &= u_3
\end{aligned}
$$

where the normalization rule $\beta_{ii} = 1$ is used and where the u's are random disturbances with $E(u_1) = E(u_2) = E(u_3) = 0$. However, (1.2) is not segmentable (or recursive) unless additional a priori assumptions are made. As in (1.1), the matrix of coefficients of the endogenous variables (in this case the variables Y_1, Y_2, Y_3) must be triangular and the covariances of the disturbances must be zero for the system to be recursive-form (this insures the identifiability of each equation in lieu of other information)[8]. Simon assumes that Y_1 is not causally dependent on Y_2 or Y_3, and Y_2 is not causally dependent on Y_3. This assertion amounts to an a priori substantive ordering of the three variables. Additionally, restricting the covariances of the disturbances to be zero so that $E(u_1u_2) = E(u_1u_3) = E(u_2u_3) = 0$, yields the following recursive-form system:

$$
\begin{aligned}
Y_1 \qquad\qquad &= u_1 \\
\beta_{21}Y_1 + Y_2 \qquad &= u_2 \qquad (1.3) \\
\beta_{31}Y_1 + \beta_{32}Y_2 + Y_3 &= u_3
\end{aligned}
$$

Given the assumed substantive ordering of the three variables, the assumed independence of the disturbances, and assuming that the variables are measured from their respective means, Simon proves the following: if certain restrictions are placed on the set $\{\beta_{21},\beta_{31},\beta_{32}\}$ (that is, certain combinations of the elements of this set are zero), then certain conditions in terms of zero-order correlations between various pairs of variables in the model hold, and conversely. Simon then deduces that if none of these correlational conditions holds, then all three structural coefficients β_{21}, β_{31}, β_{32} are non-zero.

Alternatively, we may look at the matter from a different perspective. It can be shown that the two restrictions placed on (1.3) -- that the matrix of coefficients is triangular and that the covariances of the disturbances are zero -- are jointly necessary and sufficient to identify every equation in the model without having further information on the model. Moreover, each equation of (1.3) is just identified. Now, with additional restrictions on the set $\{\beta_{21},\beta_{31},\beta_{32}\}$ (that some elements of this set are zero), either the second or the third equation (or both of these equations) becomes overidentified. Thus, upon observation of the population data, Simon's correlational conditions allow us to determine which one of the possible situations of overidentification (there are $C_1^3 + C_2^3 + C_3^3 = 7$ possible situations) is consistent with the a priori just identifying restrictions on (1.3). If none of the overidentified conditions holds, then we deduce that every equation is just identified.

Perhaps the most important interpretation we can give to Simon's statistical results is that he provides a test of the existence of direct causal relations as formally defined in his earlier paper (1953). Since his correlational conditions tell us which specific structural coefficients are non-zero or zero, and since a non-zero structural coefficient defines a causal relation, we can determine which specific causal relations hold among the variables in the model (assuming a priori, of course, that some of them do not hold). With reference to (1.3), we can summarize Simon's results as follows, where ρ_{ij} is the zero-order correlation coefficient between variables Y_i and Y_j:

Given the a priori restrictions on (1.3), then

1) $\beta_{21} = 0$, $\beta_{31} \neq 0$, $\beta_{32} \neq 0$ if and only if
$\rho_{21} = 0$, $\rho_{31} \neq 0$, $\rho_{32} \neq 0$.

2) $\beta_{31} = 0$, $\beta_{21} \neq 0$, $\beta_{32} \neq 0$ if and only if
$\rho_{21} \neq 0$, $\rho_{31} \neq 0$, $\rho_{32} \neq 0$, and $\rho_{31} = \rho_{21}\rho_{32}$
(or, equivalently, $\rho_{31.2} = 0$).

3) $\beta_{32} = 0$, $\beta_{21} \neq 0$, $\beta_{31} \neq 0$ if and only if
$\rho_{21} \neq 0$, $\rho_{31} \neq 0$, $\rho_{32} \neq 0$, and $\rho_{32} = \rho_{21}\rho_{31}$
(or, equivalently, $\rho_{32.1} = 0$).

4) $\beta_{21} = \beta_{31} = 0$, $\beta_{32} \neq 0$ if and only if
$\rho_{21} = \rho_{31} = 0$, $\rho_{32} \neq 0$.

5) $\beta_{21} = \beta_{32} = 0$, $\beta_{31} \neq 0$ if and only if
$\rho_{21} = \rho_{32} = 0$, $\rho_{31} \neq 0$.

6) $\beta_{31} = \beta_{32} = 0$, $\beta_{21} \neq 0$ if and only if
$\rho_{31} = \rho_{32} = 0$, $\rho_{21} \neq 0$.

7) $\beta_{21} = \beta_{31} = \beta_{32} = 0$ if and only if
$\rho_{21} = \rho_{31} = \rho_{32} = 0$.

8) $\beta_{21} \neq 0$, $\beta_{31} \neq 0$, $\beta_{32} \neq 0$ if and only if none of the correlational conditions (1)-(7) holds.[9]

While Simon's conditions refer to population correlation coefficients, we can nevertheless approximate his conditions by employing sample correlation coefficients. However, it should be kept in mind that his results have utility only with respect to the a priori simplyfying assumptions of the model and the initial just identifying restrictions on its equations. Consequently, any practical application of Simon's results to sample data must be done with a sensitivity to the possibility that some of the a priori assertions may not actually hold (of course, this warning could be given to the use of any statistical model). However, even though Simon's results refer only to formal statistical models of the real world, the application of his procedures can lead to tentative real world inferences.

On heuristic grounds, Simon's work has appeal since it may be viewed as an attempt to explicate the assumptions and logical processes that are usually involved in making causal inferences from correlational data. When we observe a non-zero correlation coefficient between two variables and we wish to make some causal inference from this association, we ordinarily introduce a third variable if we have doubts whether the observed correlation is "genuine."[10] This third variable may account for the observed zero-order correlation and render it "spurious." In order to investigate this possibility, we compute the partial correlation coefficient between the original two variables, holding the third variable constant as a control. Comparing the computed partial correlation coefficient with the initial zero-order correlation coefficient, we attempt to make causal inferences. If the computed partial correlation coefficient is approximately zero, then we are apt to conclude either (1) that the third variable is an intervening variable, implying that the causal relationship between the original two variables is mediated by the third variable; or (2) that the third variable causes both of the original variables and thus accounts for the initially observed zero-order correlation. However, this procedure cannot tell us the causal direction in (1), let alone distinguish between (1) and (2).

We have seen above how, by moving to a formal representation of causal processes within the confines of the recursive-form, Simon makes clear what assumptions are necessary to distinguish between (1) and (2) and to determine the causal direction in (1). The correlational data alone do not allow us to decide among these possible conclusions. Not only does he point out the necessary assumptions, but his analysis also provides us with a set of correlational predictions that preserve the flavor of less rigorous causal investigations. The heuristic appeal of Simon's work, then, rests not so much on his ability to provide a test for causal theories embedded in a formal statistical model, but rather on his ability to use the algebraic relationships of the formal model to deduce a correlational test which maintains highly intuitive properties.

II. The Present Problem

We have seen that Simon's attempts to define formally the causal relation takes place exclusively within the confines of a recursive-form linear structural system of equations.[11] While we find his approach to be useful and important, it has several significant limitations. In this section of the paper, we shall be concerned with the nature of these limitations.

A. Reciprocal Causal Relations

Since the basis of Simon's implicit definition of the causal relation is the notion of an asymmetrical production relation between two variables, and not the temporal sequence of the pair, his formal definition of the concept is not entirely adequate. The choice of the recursive-form linear structural model to play a role in the formal definition of the causal relation prevents, in general, our consideration of reciprocal asymmetrical production relations. If we employ Simon's formal definition of the causal relation, it is clear that the recursive-form structural model does not permit a variable Y_i to both cause and be caused by some other variable Y_j, without the explicit introduction of temporal sequence into the system by "lagging" one or the other of the two variables.[12] This can easily be seen by noting that in the recursive-form linear structural model the matrix of coefficients of the endogenous variables must be triangular. (See (1.1)) Hence, if a structural coefficient $\beta_{ij} \neq 0$, then β_{ji} must be zero (or conversely).

Although we would argue that temporal sequences are present and ought to be explicitly included in the formal representations of such processes as symbiosis in biological systems, interpersonal attitude influence in the husband-wife dyad, or the influence relationship between constituency and legislative representative in the formation of public policy, it is not these sequences which lead us to consider the processes as causal.[13] If we are inclined to consider a process like symbiosis as a reciprocal causal process, in which we believe temporal sequence to be no more the basis for either of the two asymmetries than it is for the single asymmetry of our rock-window example, then Simon's approach to formalizing the causal relation is inadequate for our purposes. When we think of a symbiotic relationship between two biological organisms, we are most concerned with the fact that each organism asymmetrically provides products necessary for the existence and development of the other. The gist of such a relation is one of reciprocal production or influence. The temporal sequences involved do not form the bases of the mutual asymmetries we are interested in; it is the existence of a situation in which benefit is reciprocally given and received by both organisms which performs this function.

Of course, it can be argued that every causal relation, reciprocal or not, involves temporal sequences to some extent, or otherwise no change could ever take place as a consequence of the relation. Thus, it might be inferred from this argument that temporal sequences should be explicitly included in the formal representation of any process which we regard as causal. Yet, no matter how closely the concept of the causal relation may hang together in empirical situations with the concept of temporal sequence, these two principles are logically independent.[14] Simon does not argue that time is not involved in causal relations, but rather that the basis of the asymmetry of a causal relation is provided by the mechanistic or operational notion of production. Knowledge of the temporal sequence of two variables does not imply that a mechanistic relation holds between the pair. Hence, the concept of time should play no role whatsoever in either the development or the statement of a satisfactory formal definition of the causal relation.[15]

However, as already mentioned, if we admit the possibility of reciprocal causation, it is impossible to define formally the causal relation within the context of the recursive-form structural model without the explicit representation of temporal sequence. But, by the above argument, such an explicit representation would violate the logical independence of the causal relation and temporal sequence. Therefore, Simon's formal definition actually precludes the notion of reciprocal causation, even though this notion is perfectly consistent with his implicit definition of the causal relation.

In other words, we regard a reciprocal causal process as a process exhibiting two asymmetries, each of which satisfies Simon's conceptual notion of the causal relation. There is nothing contained in his implicit definition of the causal relation which excludes the possibility that two asymmetrical relations can hold between a pair of variables. Indeed, the concept of reciprocal causation does not in any way change or distort Simon's implicit definition of the causal relation. Furthermore, we feel quite comfortable with his implicit definition. The notion of reciprocal causation merely implies that two causal relations, rather than only one, are present between a pair of variables. We intend no additional meaning, either implicit or explicit, to be attached to this notion.

Like Simon, we shall restrict our formal definition of the causal relation to refer only to a representational model. However, as we have seen in the preceding discussion, the recursive-form linear structural system is inappropriate for our purposes (as is any segmentable system) if we desire to consider reciprocal causation. Since we desire to confine our interest in causal relations to the domain of a structural model, it is clear that we must have at our disposal a structural model which has the ability to represent reciprocal causal relations. After a discussion of some of the additional problems of the recursive-form, we shall consider a type of linear structural model which meets our requirements.

B. Restrictions on the Covariances of Disturbance Terms

In addition to its inability to represent reciprocal causal relations, the recursive-form linear structural system has another disadvantage: its restriction that the covariances of the disturbance terms are zero is a rather stringent a priori condition to impose upon models of many empirical processes. In practice, this restriction is usually not a consequence of the theory which gives rise to the particular structural model being considered. It is more likely to be either a consequence of the investigator's implicitly held ideas about some causal process being contained in the disturbance terms or a consequence of the desire for computational convenience. We generally regard the disturbances to be the net effects of many independently operating tiny influences that have been excluded from the structural parts of our equations. Therefore, when we assume that the disturbances of two equations are independent, we are implying that the omitted variables influencing the two equations have few elements in common. Whether or not such an assumption is plausible may depend largely on the theoretical proximity of the mechanisms described by the two equations. For example, even if we were to consider attitude influence in the husband-wife dyad to be non-reciprocal, with the husband influencing his wife, but not conversely, it would probably be rather unreasonable for us to expect that the variables omitted from a stochastic two-equation formal description of this process have no large number of elements in common. On the other hand, with respect to a stochastic two-equation economic model, where each equation describes widely separated, yet operationally linked, segments of some market, there is probably a greater likelihood that the excluded variables affecting the two equations have few common elements.[16] However, even in the latter example, we would probably be at a loss to give a theoretically meaningful defense of the assertion that the disturbances are independent.

Unfortunately, the restriction that the covariances of the disturbances are zero is often made simply because it offers a gain in the ease of estimation of the structural parameters of an equation system with a triangular matrix of coefficients. This restriction is used to aid in the identification of individual equations in such a system, and helps to guarantee the consistency of ordinary least squares estimates.[17] However, as indicated above, there is a good deal of risk associated with this restriction. Invoking this restriction becomes even more problematical if the only reason for doing so is the desire for computational convenience. This is not only due to the reasons already mentioned, but also due to the critical relationship between this restriction and the identification of equations in systems with triangular coefficient matrices.

The concept of an identified equation is generally defined in the econometric literature in terms equivalent to the following: a particular structure equation of a model is identified if that equation is the only equation, among the entire set of structure equations compatible with the data, which is also compatible with the a priori restrictions imposed by the model on that particular equation. (Essentially, an equation is identified if we have enough information to distinguish that equation from other equations in a simultaneous system.) Using this definition, it can be shown that in a linear model only those structure equations which are linear combinations of the true structure equations are the equations which satisfy the data. Moreover, an equation in such a model is identified if and only if no more than one of this set of linear combinations is consistent with the restrictions placed by the model on that particular equation.[18] When the only a priori information employed is that which specifies that certain variables are excluded from certain equations (i.e. that certain structural coefficients are zero), a necessary (but not sufficient) condition (order condition) for the identifiability of a particular equation in a linear model of G equations is that at least $G-1$ of the variables that appear in the entire model be excluded a priori from that equation. Using the same type of a priori exclusion information, a necessary and sufficient condition (rank condition) for the identifiability of a particular equation in a model of G linear equations is that it be possible to form at least one nonzero determinant of order $G-1$ from the matrix of coefficients constructed as follows: from the matrix of coefficients of the entire system, delete every column which corresponds to a variable not excluded a priori from that equation, and delete the row of coefficients of that equation.[19] (It is presumed, however, that the reader is already familiar with the concept of identification and with the order and rank conditions for the identifiability of equations in linear models, since the discussion we have just offered is certainly not sufficient to prepare the reader for all of what follows in this paper. We intend the above discussion to be more of a reminder than anything else. If the reader is not familiar with the identification problem, he should consult one of the standard works: for example, Christ (1966), Johnston (1963), or Fisher (1966).)

Even though it can be shown that every equation in a recursive-form linear structural system is identified, it is not uncommon for investigators to hold the belief that it is the triangularity of the coefficient matrix that is sufficient to identify all equations of the system. However, without additional *a priori* information, the triangularity of the coefficient matrix is *insufficient* to identify *any* equation of the system except the first. To see this, consider the following two-equation model, where the normalization rule $\beta_{ii} = -1$ has been used:

$$(2.1)\qquad \begin{aligned} -Y_1 \qquad &= u_1 \\ \beta_{21}Y_1 - Y_2 &= u_2 \end{aligned}$$

where we assume $\beta_{12} = 0$, and all that we assume about the disturbances, u_1, u_2, is that their variance-covariance matrix, call it Σ, is positive definite. While the coefficient matrix of (2.1), call it B, is clearly triangular (since $\beta_{12} = 0$), nevertheless the second equation is not identified, since it does not satisfy the order condition for identifiability (a necessary condition). However, the first equation is identified (trivially), since it satisfies the rank condition (a necessary and sufficient condition).

Let us examine the two equations of (2.1) more closely. The matrix of coefficients of the entire system is

$$(2.2)\qquad \underset{\sim}{B} = \begin{pmatrix} -1 & 0 \\ \beta_{21} & -1 \end{pmatrix}$$

For the first equation of (2.1), the matrix of coefficients relevant to the rank condition is found to be

$$(2.3)\qquad (-1)$$

which has been formed by deleting from B: (a) every column not containing an *a priori* assumed zero in the first equation, and (b) the row of coefficients of the first equation. Now, for the first equation to satisfy the rank condition for identifiability, which is both necessary and sufficient, we must be able to form at least one non-zero determinant of order $G-1$, where G is the number of equations in the system, from matrix (2.3). Since the determinant of (2.3) is non-zero and of order $G-1 = 1$, the first equation of (2.1) is identified.

However, by the order condition for identifiability, which is necessary, the second equation of (2.1) must *a priori* exclude at least $G-1$ of the variables that appear in the system. For system (2.1), $G-1 = 1$. But, the second equation excludes zero of the variables appearing in the system. Hence, by the order condition, the second equation of (2.1) is not identified.

It should now be clear that the triangularity of the coefficient matrix is insufficient to identify every equation of a linear model. Moreover, in order to identify the entire system, having only the knowledge that the coefficient matrix is triangular, it is both necessary and sufficient to make the *additional* restriction that the covariances of the disturbances are zero. (See Fisher (1966) for a proof of this.)

Nor does the triangularity of the coefficient matrix give license to readily assume zero covariances of the disturbances, even though the hierarchical appearance of the system for some reason seems to tempt many investigators to make this assumption. The triangularity of the coefficient matrix only *guarantees* that perturbations in a disturbance *directly* influence only one of the endogeneous variables which appear in the same equation as that disturbance. The triangularity does *not* imply that movements in one disturbance are not associated with movements in other disturbances. Indeed, if the triangularity did imply such independence among the disturbances, then the triangularity alone would suffice to identify every equation in the model. But, we already have seen that the triangularity alone is not sufficient to identify the entire system. Therefore, it is clear that the triangularity does not imply that the disturbances are independent of each other.

Now, let us refer to our earlier example of attitude influence in the husband-wife dyad. If we regard influence in this dyad as non-reciprocal, then we could certainly represent the process in a structural system with a triangular matrix of coefficients. Yet, we already know that this triangularity does not alone suffice to identify the entire model. Let us suppose that we do not impose any *a priori* restrictions on the model other than those exclusion restrictions which led to the triangularity of the coefficient matrix. Since we have no other information, it is clear that we must additionally assume that the covariances of the disturbances are zero, if the entire system is to be identified. However, as indicated above, we cannot rely on the triangularity alone to permit us to assume that the disturbances are independent.

Nor, are we really convinced that it is even desirable to make such an assertion about the disturbances. We have a strong feeling that such an assertion would have rather implausible implications. If we were to make this assertion, we would be implying that the influences omitted from the structural parts of our equations, yet whose net effects are represented by the disturbances, are essentially different for each equation in the model. But, we are inclined to think that such an implication is not very plausible for the husband-wife dyad. The only alternative left to us is to make the assertion about the disturbances out of a desire for computational convenience. However, given that we already believe that the assertion is probably inappropriate, it certainly would be foolish for us to choose this last alternative.

Let us now examine how a situation of associated disturbances may affect the application of Simon's correlational testing procedure. We shall use a set of computer simulated data to illustrate such a situation. We are interested in assessing whether Simon's technique is sensitive to the "true" structural model, given a situation of associated disturbances. This illustration may shed some light on the problems associated with making the assumption that the disturbances in a triangular system are independent when, in fact, they are not.

The following structural system was used to generate a sample of 1000 cases of computer simulated data.[20]

$$(2.4)\qquad \begin{aligned} -Y_1 \qquad\qquad &+ u_1 = 0 \\ 5Y_1 - Y_2 \qquad &+ u_2 = 0 \\ 6Y_1 \qquad - Y_3 &+ u_3 = 0 \end{aligned}$$

where u_1 is a normally distributed variable with $E(u_1) = 0$ and $\sigma u_1 = 1$; and where $u_2 = u_1^2$, $u_3 = u_1^3$. While the coefficient matrix of (2.4) is triangular, we have a situation where the disturbances of this system are not independent. If we were to assume the disturbances to be independent (although, in fact, they are not) and since $\beta_{32} = 0$, Simon's procedure would predict that (2.4) is correct if and only if $\rho_{21} \neq 0$, $\rho_{31} \neq 0$, $\rho_{32} \neq 0$, and $\rho_{32.1} = 0$.

Upon observation of the correlation coefficients computed from the simulated data generated by (2.4), we find that $\rho_{32.1} \neq 0$ but rather is equal to 0.20. Therefore, Simon's method fails to detect the "true" structural model (2.4) which generated the data. This analysis suggests that to naively assume independence of structural disturbances may lead to serious consequences. The application of Simon's procedure to such a misspecified model may produce quite misleading conclusions about the configuration of causal relations within the model.

It should now be clear that we should be very cautious in choosing the recursive-form linear model to represent our theories, even when a triangular matrix of structural coefficients seems appropriate (i.e. when a complete *a priori* substantive ordering of the variables is theoretically *justified*). For unless we are willing to make the *additional* restriction (having no other identifiability information at hand) that the covariances of the disturbances are zero -- a restriction that is most likely not a consequence of our theory, but more likely a consequence of implicit notions of isolability (*ceteris paribus*) or of a desire for computational ease -- the recursive-form is not suitable. Indeed, even when we do think it suitable, its advantages usually just do not seem to outweigh its disadvantages.

If, as a consequence of our abstraction of some empirical situation, we feel that our theory can be best represented in a structural system with a triangular matrix of coefficients, it might be preferable to employ some structural system which requires weaker restrictions on the disturbance covariances, and avoid the recursive-form altogether, along with its high risks. More often than not, the theory which we intend to represent has implications only for the structural parts of our equations, and does not give rise to restrictions about the behavior of the disturbances. Fortunately, there do exist alternatives to the recursive-form for the representation of theories which inspire a triangular matrix of structural coefficients. One of these alternatives will be discussed in the following pages. Although the structural system which we shall discuss has some drawbacks, it will be seen to be superior to the recursive-form in a number of important respects.

III. An Alternative Formal Definition of the Causal Relation

A. Introductory Remarks

In the preceding sections of this paper, we have focused on a number of issues associated with the concept of causality. A large part of our discussion was devoted to a critical examination of Simon's approach to defining the causal relation. We saw that his approach is both useful and important, but that it is not entirely satisfactory. While we found his implicit notion of the causal relation (an asymmetrical influence or production relationship between two variables) to have desirable properties, we found his formal definition of this concept (confined to the recursive-form) to be inadequate for our purposes. In particular, we saw that if we admit the possibility of reciprocal causation, it is impossible to define formally the causal relation within the context of the recursive-form linear structural model without the explicit representation of temporal sequence. We argued, however, that such an explicit representation would violate the logical independence of the causal relation and temporal sequence.

In addition, we found the recursive-form to be unsatisfactory due to its rather stringent restriction on the covariances of the disturbances. We saw that the recursive-form requires the assumption that the disturbance covariances are zero. We examined this assumption in some detail and pointed out its critical role in the identification of systems with triangular coefficient matrices. While it was seen that this assumption leads to a computationally convenient situation, it was also seen that high risks are involved when it is invoked. We saw that even when a system has a triangular coefficient matrix, to assume that the disturbances are independent often has implausible implications and can lead to misleading causal conclusions (employing Simon's procedure).

Therefore, since the recursive-form is unsatisfactory in a number of important respects, we must offer some alternative structural system in which we may adequately define the causal relation and in which we may adequately represent causal theories. Such an alternative structural system must (1) have the ability to represent reciprocal causal relations, without requiring the explicit representation of temporal sequence and (2) allow the possibility that disturbance terms are associated. We shall now proceed to the description of one such structural system which meets our requirements. After this description, we shall then be ready to offer our formal definition of the causal relation.

B. The Simultaneous Linear Structural Model

Before we begin our description of the structural model in which we shall eventually imbed our notion of the causal relation, let us make a distinction between two types of variables. Let us call a variable which is determined in a particular model by other variables in that model an endogeneous variable. A variable which is not determined in that particular model will be called a predetermined variable.[21] While we have occasionally used this distinction in the preceding pages of this paper, we were not very precise about its meaning. However, for the model we are about to describe, this distinction plays a very crucial role. This role will become apparent in the course of our description.

Let us assume that our model refers to characteristics of a population, and not to characteristics of some sample of that population. Let us further assume that we have knowledge of the values of the variables in that population, but not of the structual coefficients. We shall treat these structural coefficients as unknowns. We shall also assume that the equations of our model are linear in the variables and unknowns. Thus, we consider the following model of G equations:[22]

$$(3.1)\quad \begin{array}{l} y_1\beta_{11}+\cdots+y_G\beta_{1G}+z_1\gamma_{11}+\cdots+z_k\gamma_{1k}+u_1 \;\cdots \\ y_1\beta_{G1}+\cdots+y_G\beta_{GG}+z_1\gamma_{G1}+\cdots+z_k\gamma_{Gk}+u_G = \underset{\sim}{0} \end{array}$$

which may be written in matrix form as:

$$(3.2)\quad \underset{\sim}{Y}\underset{\sim}{B} + \underset{\sim}{Z}\underset{\sim}{\Gamma} + \underset{\sim}{U} = \underset{\sim}{0}$$

where

(3.3) $\underset{\sim}{Y} = (y_1,\ldots,y_G)$ is the $1\times G$ row vector of endogenous variables in the model;

(3.4) $\underset{\sim}{B} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{G1} \\ \vdots & & \vdots \\ \beta_{1G} & \cdots & \beta_{GG} \end{pmatrix}$ is the $G\times G$ matrix of coefficients (constant but unknown population parameters) of the endogenous variables;

(3.5) $\underset{\sim}{Z} = (z_1,\ldots,z_K)$ is the $1\times K$ row vector of predetermined variables in the model;

(3.6) $\underset{\sim}{\Gamma} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{G1} \\ \vdots & & \vdots \\ \gamma_{1K} & \cdots & \gamma_{GK} \end{pmatrix}$ is the $K\times G$ matrix of coefficients (constant but unknown population parameters) of the predetermined variables;

and

(3.7) $\underset{\sim}{U} = (u_1,\ldots,u_G)$ is the $1\times G$ row vector of the random structural disturbances.

Clearly, (3.2) contains as many equations as there are endogenous variables in the model.

Assumption 3.1: $\underset{\sim}{B}$ is nonsingular.

Given that $\underset{\sim}{B}$ is non-singular, postmultiplication of (3.2) by B^{-1} yields a solution for the values of the G endogenous variables of the model in terms of the predetermined variables and the structural disturbances. Thus, we have:

$$(3.8)\quad \underset{\sim}{Y} = \underset{\sim}{Z}\underset{\sim}{\Pi} + \underset{\sim}{V}$$

which is the reduced form of (3.2) where

(3.9) $\underset{\sim}{\Pi} = -\underset{\sim}{\Gamma}\underset{\sim}{B}^{-1}$ is the $K\times G$ matrix of reduced form coefficients

and

(3.10) $\underset{\sim}{V} = -\underset{\sim}{U}\underset{\sim}{B}^{-1}$ is the $1\times G$ matrix of reduced form disturbances.

It is obvious that each equation of the reduced form (3.8) has only one endogenous variable.

Assumption 3.2: $E(\underset{\sim}{U}) = \underset{\sim}{0}$

That is, each disturbance has zero expectation.

Now, denote the variance-covariance matrix of the disturbances by $\underset{\sim}{\Sigma}$, so that: $E(\underset{\sim}{U}'\underset{\sim}{U}) = \underset{\sim}{\Sigma}$, where the prime mark stands for transposition.

Assumption 3.3: $\underset{\sim}{\Sigma}$ is positive definite.

Assumption 3.4: The predetermined variables are linearly independent.

Assumption 3.5: $E(\underset{\sim}{Z}'\underset{\sim}{U}) = \underset{\sim}{0}$

In other words, the predetermined variables are assumed to be independent of the disturbances.

Let us now discuss some properties of the reduced form (3.8). Since the reduced form disturbances are linear combinations of the disturbances in (3.2), it can easily be shown that:

$$(3.11)\quad E(\underset{\sim}{V}) = \underset{\sim}{0};$$

(3.12) $E(\underset{\sim}{V}'\underset{\sim}{V})$, call it $\underset{\sim}{\Omega}$, is positive definite;

and

$$(3.13)\quad E(\underset{\sim}{Z}'\underset{\sim}{V}) = \underset{\sim}{0}$$

Let us assume that the only kind of identifying restrictions on model (3.2) are exclusion restrictions. Suppose that extraneous a priori information indicates to us that certain of the y's, H in number, are permitted to appear in the gth equation of (3.2), and the remainder of them, $G-H$ in number, are excluded. Suppose also that certain of the z's, J in number, are permitted to appear, and the remainder of them, $K-J$, are excluded. It is always possible for the y's to be numbered in such a way that $y_1,\ldots,y_H$ are the ones that appear and $y_{H+1},\ldots,y_G$ are excluded, and for the z's to be so numbered that $z_1,\ldots,z_J$ appear and $z_{J+1},\ldots,z_K$ are excluded. Of course, this a priori information amounts to exact linear restrictions on the coefficients of (3.2) such that $\beta_{g,H+1} = \cdots = \beta_{gG} = \gamma_{g,J+1} = \cdots = \gamma_{g,K} = 0$. Let us further assume that each equation in (3.2) satisfies the rank condition for identifiability, which implies that the order condition must also be satisfied, i.e. $(G-H) + (K-J) \geq G-1$ or $K-J \geq H-1$ for all g, $1 \leq g \leq G$. Moreover, let us employ the normalization rule $\beta_{gg} = -1$ for all g, $1 \leq g \leq G$. Hence, the g^{th} equation of (3.2) may be written, after solving for y_g, as:

$$(3.14)\quad y_g = \underset{\sim}{y}_1\underset{\sim}{\beta} + \underset{\sim}{z}_1\underset{\sim}{\gamma} + u_g$$

where $\underset{\sim}{y}_1$ is a $1\times H-1$ row vector of the $H-1$ endogenous variables other than y_g included in the equation, so that:

$$(3.15)\quad \underset{\sim}{y}_1 = (y_1,\ldots,y_H);$$

$\underset{\sim}{\beta}$ is an $H-1\times 1$ column vector of the coefficients of $\underset{\sim}{y}_1$, so that:

$$(3.16)\quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_{g1} \\ \vdots \\ \beta_{gH} \end{pmatrix};$$

$\underset{\sim}{z}_1$ is a $1\times J$ row vector of the J predetermined variables included in the equation:

$$(3.17)\quad \underset{\sim}{z}_1 = (z_1,\ldots,z_J);$$

$\underset{\sim}{\gamma}$ is a $J\times 1$ column vector of the coefficients of $\underset{\sim}{z}_1$:

$$(3.18)\quad \underset{\sim}{\gamma} = \begin{pmatrix} \gamma_{g1} \\ \vdots \\ \gamma_{gJ} \end{pmatrix};$$

and u_g is the structural disturbance for the equation.

C. Defining the Causal Relation

Wold argues that a causal interpretation should not be given to the behavior relations represented by the structural coefficients of model (3.2), i.e., the elements of $\underset{\sim}{B}$ and $\underset{\sim}{\Gamma}$ (See Wold and Jureen (1953), Wold (1959, 1960), and Strotz and Wold (1960).). His argument is that such coefficients do not lend themselves to direct operative use in the sense of permitting a stimulus-response interpretation of the relations which they represent. In other words, Wold uses the concept of causality to correspond with the usual laboratory meaning of the term. The basis for this use of the concept rests largely on the notion of control. Therefore, control in a system like (3.2) is a result of the operative significance of the behavior relations of the reduced form of (3.2), i.e., (3.8). This is so because an analogy to the direct control of the stimulus in the laboratory setting can only be gained by the assumed manipulability of the predetermined variables in (3.2), whose behavioral relations are represented in the reduced form. Wold argues that the structural relations do not lend themselves to this interpretation since they contain interdependencies among the various endogenous variables in the model -- variables which are not subject to direct operational control.

However, Wold's notion of causality is much narrower than ours. While we agree that direct operative significance in the experimental sense cannot be attributed to the structural relations of (3.2), our view of causality does give causal meaning to these structural relations. Our concept of causality as an asymmetrical influence process includes more than just the notion of operational control. This should be apparent from earlier sections of this paper. We are also interested in giving causal significance to the asymmetric relations among the endogenous variables themselves. Relying only on the reduced form, as Wold would have it, tells us nothing about such structural relationships. When a structural model is constructed, it is not done so arbitrarily. It is formulated in an attempt to represent meaningful relations derived from theory and astute perception. The structural model is not just a set of arbitrary linear combinations of variables (See Goldberger (1964)). In fact, the structural model, not the reduced form, is the formulation which is inspired by the underlying theory of the model. Therefore, we consider giving causal significance to the structural relations of (3.2) entirely appropriate.

We are now ready to offer a formal definition of the causal relation. We shall do this in terms of the g^{th} equation (3.14) of model (3.2).

Definition 3.1: A variable is a direct cause of y_g if that variable is an element of either $\underset{\sim}{y}_1$ or $\underset{\sim}{z}_1$.

Therefore, a variable which is a direct cause of y_g has a (nonzero) structural coefficient which is an element of either $\underset{\sim}{\beta}$ or $\underset{\sim}{\gamma}$.

Thus, we have defined the direct causal relation within a model which admits the possibility of reciprocal causal relations, without the explicit introduction of temporal sequence. However, it should be apparent that not all variables may be involved in reciprocal causal relationships; this clearly applies to the predetermined variables in the model. While a predetermined variable may be the cause of some other variable(s) in the model, it cannot itself be caused by any other variable in the model. That certain variables are restricted from engaging in reciprocal causal relationships in a particular model should not, however, detract from the usefulness of Definition 3.1. It is certainly reasonable to suppose that, as a consequence of the theory which gives rise to a particular structural model, certain variables in that model are given as predetermined. Our critique of Simon's formal definition of the causal relation should not be misconstrued. We did not argue that all variables in a given model should enjoy the possibility of taking part in reciprocal relationships. We merely pointed out that Simon's formal definition is inadequate because it does not permit the possibility of any variable taking part in such a relationship.

We also noted that to assume the independence of the structural disturbances in a given model, as is required by the recursive-form, is an arbitrary decision usually not a consequence of the underlying theory and often having rather implausible implications. A similar charge might be leveled at the assumption in (3.2) that the predetermined variables are independent of all disturbances in the model. Since we can never actually observe the disturbances, we cannot obtain more observational information about the relationships between the disturbances and the variables we consider predetermined than we can about the relationships among the disturbances themselves. Yet, while both assumptions are arbitrary, one may be more arbitrary than the other. It is considerably more difficult to give theoretical justification for the assertion that variables determined in the same model have independent disturbances than for the assertion that variables determined outside of the model are independent of the disturbances of variables determined within the model. Therefore, we feel somewhat more comfortable assuming that predetermined variables are independent of the model's disturbances than we feel assuming that the model's disturbances themselves are independent.

IV. A Test of Causal Theories

A. Causal Theories

Let us suppose that (3.2) represents some particular theory. In addition, let us require that all previous assumptions made about this system still hold. Suppose further that the structural relations in (3.2) are causal relations, as defined in Definition (3.1). We have already specified that certain variables are excluded from certain equations in (3.2) -- that is, that certain elements of $\underset{\sim}{B}$ and/or $\underset{\sim}{\Gamma}$ are a priori assumed to be zero. Let us assume that such exclusion restrictions on the model are consequences of our theory. In other words, our theory indicates to us which variables are, and which variables are not, direct causes of the g^{th} endogeneous variable, for all g, $1 \le g \le G$. Moreover, we have already assumed that the exclusion restrictions placed on (3.2) are sufficient to identify each equation of the model.

However, we may not be certain that our a priori exclusion restrictions are correct with respect to the population. Our causal theory, represented by (3.2), may not be appropriate. Therefore, it is desirable to have some statistical test of the exclusion restrictions on the equations of (3.2). Alternatively, we may interpret such a test to be a test of the causal relations postulated in (3.2). It should be obvious that we shall be unable to provide a separate test for each a priori assumption of (3.2). We shall have to consider at least some of the assumptions of (3.2) to be correct without their being tested. Therefore, our statistical test of the a priori exclusion restrictions of (3.2) will be useful only if we are willing to assume that certain other assumptions of (3.2) are valid. In fact, it will be seen that not even every exclusion restriction on (3.2) will be tested. At least some of these restrictions will have to be assumed to be correct. Hence, our statistical test of the causal relations of (3.2) will allow us to make only conditional statements about the adequacy of the model. But, this should not bother us. We have already argued, much earlier in this paper, that it makes little sense ever to state that a relationship is unconditionally causal.

Even though a number of tests of identifying restrictions may be found in the econometric literature (for example, tests developed by Anderson and Rubin (1950) and by Basmann (1960); also see Christ (1966), Chapter X.), none of these tests has the heuristic appeal of the correlational test developed for the recursive-form by Simon. While it is clear that Simon's testing procedure is inappropriate for non-recursive systems, it would be desirable to develop a test which has some of the intuitive characteristics associated with his procedure. Particularly, it would be desirable to develop a correlational test which in some sense reflects the logical process associated with introducing control variables into a zero-order correlational relationship and examining the partial correlations as a test for "spuriousness." We earlier argued that Simon's procedure is appealing precisely because it reflects this logical process, even though the procedure was deduced from a formal statistical model. Therefore, it is hoped that the results of the derivation that follows will also reflect this process.

B. Derivation of the Statistical Test

Without loss of generality, let us assume that all variables in (3.2) have zero means. We still assume that we are referring to the population and not to a sample of that population. Employing our normalization rule, we again consider the g^{th} equation of (3.2):

(4.1) $$y_g = \underset{\sim}{y}_1\underset{\sim}{\beta} + \underset{\sim}{z}_1\underset{\sim}{\gamma} + u_g$$

Let

(4.2) $\underset{\sim}{z}_2 = (z_{J+1},\ldots,z_K)$ be the $1 \times K\text{-}J$ row vector of the predetermined variables excluded from the g^{th} equation.

Since $\underset{\sim}{z}_1$ is the $1 \times J$ row vector of predetermined variables included in the g^{th} equation, we may write the reduced form of $\underset{\sim}{y}_1$ as:

(4.3) $$\underset{\sim}{y}_1 = \underset{\sim}{z}_1\underset{\sim}{\pi}^{11} + \underset{\sim}{z}_2\underset{\sim}{\pi}^{12} + \underset{\sim}{v}_1$$

where

(4.4) $\underset{\sim}{\pi}^{11}$ is the $J \times H\text{-}1$ matrix of reduced form coefficients of $\underset{\sim}{z}_1$,

(4.5) $\underset{\sim}{\pi}^{12}$ is the $K\text{-}J \times H\text{-}1$ matrix of reduced form coefficients of $\underset{\sim}{z}_2$,

and

(4.6) $\underset{\sim}{v}_1$ is the $1 \times H\text{-}1$ row vector of reduced form disturbances for $\underset{\sim}{y}_1$.

Now, let us define:

$$(4.7) \qquad y_1^* = y_1 - v_1 = z_1\pi^{11} + z_2\pi^{12}$$

Solving (4.7) for y_1 we obtain:

$$(4.8) \qquad y_1 = y_1^* + v_1$$

Substituting (4.8) into (4.1) for y_1 we have:

$$(4.9) \qquad y_g = y_1^*\beta + z_1\gamma + (u_g + v_1\beta)$$

Since the elements of y_1^* are linear combinations of the original predetermined variables in (3.2), we have, using Assumption (3.5) and (3.13):

$$(4.10) \qquad E(y_1^{*\prime}U) = 0$$

and

$$(4.11) \qquad E(y_1^{*\prime}V) = 0$$

Therefore, we may consider the elements of y_1^* to be predetermined variables. Let

(4.12) $y_2 = (y_{H+1}, \dots, y_G)$ be the $1 \times G\text{-}H$ row vector of the $G\text{-}H$ endogenous variables excluded from the $g^{\underline{th}}$ equation.

We may write the reduced form of y_2 as:

$$(4.13) \qquad y_2 = z_1\pi^{21} + z_2\pi^{22} + v_2$$

As with y_1^*, we define:

$$(4.14) \qquad y_2^* = y_2 - v_2 = z_1\pi^{21} + z_2\pi^{22}$$

and we obtain:

$$(4.15) \qquad E(y_2^{*\prime}U) = 0$$

and

$$(4.16) \qquad E(y_2^{*\prime}V) = 0$$

Hence, we may consider the elements of y_2^* to be predetermined variables. Therefore, we have as predetermined variables the elements of

$$(4.17) \qquad z_1,\ z_2,\ y_1^*, \text{ and } y_2^*.$$

So, let us define:

(4.18) $I = \begin{pmatrix} y_1^{*\prime} \\ z_1' \\ y_2^{*\prime} \\ z_2' \end{pmatrix}$ to be the $G\text{-}1\text{+}K \times 1$ column vector of instrumental variables formed so that the predetermined variables appearing in the right-hand side of (4.9) come first, and the predetermined variables not appearing in (4.9) come last.

Now, let us partition I so that

$$(4.19) \qquad I = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$$

where

$$(4.20) \qquad I_1 = \begin{pmatrix} y_1^{*\prime} \\ z_1' \end{pmatrix} \quad \text{and} \quad I_2 = \begin{pmatrix} y_2^{*\prime} \\ z_2' \end{pmatrix}$$

Rewriting (4.9), we obtain:

$$(4.21) \qquad y_g = (y_1^* \;\; z_1)\begin{pmatrix} \beta \\ \gamma \end{pmatrix} + (u_g + v_1\beta)$$

Premultiplying (4.21) by (4.19), we have:

$$(4.22) \qquad \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} y_g = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} (y_1^* \;\; z_1)\begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} u_g + \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} v_1\beta$$

which may be written as:

$$(4.23) \qquad \begin{pmatrix} I_1y_g \\ I_2y_g \end{pmatrix} = \begin{pmatrix} I_1y_1^* & I_1z_1 \\ I_2y_1^* & I_2z_1 \end{pmatrix}\begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} I_1u_g \\ I_2u_g \end{pmatrix} + \begin{pmatrix} I_1v_1\beta \\ I_2v_1\beta \end{pmatrix}$$

Taking the expected value of (4.23), we have:

$$(4.24) \qquad E\begin{pmatrix} I_1y_g \\ I_2y_g \end{pmatrix} = E\left[\begin{pmatrix} I_1y_1^* & I_1z_1 \\ I_2y_1^* & I_2z_1 \end{pmatrix}\begin{pmatrix} \beta \\ \gamma \end{pmatrix}\right] + E\begin{pmatrix} I_1u_g \\ I_2u_g \end{pmatrix} + E\begin{pmatrix} I_1v_1\beta \\ I_2v_1\beta \end{pmatrix}$$

and this gives us:

$$(4.25) \qquad E\begin{pmatrix} I_1y_g \\ I_2y_g \end{pmatrix} = E\left[\begin{pmatrix} I_1y_1^* & I_1z_1 \\ I_2y_1^* & I_2z_1 \end{pmatrix}\begin{pmatrix} \beta \\ \gamma \end{pmatrix}\right],$$

since $E\begin{pmatrix} I_1u_g \\ I_2u_g \end{pmatrix} = E\begin{pmatrix} I_1v_1\beta \\ I_2v_1\beta \end{pmatrix} = 0$, by the fact that the elements of I_1 and I_2 are predetermined variables. Rewriting (4.25) we get:

$$(4.26) \qquad E\begin{pmatrix} I_1y_g \\ I_2y_g \end{pmatrix} = E\begin{pmatrix} I_1y_1^* & I_1z_1 \\ I_2y_1 & I_2z_1 \end{pmatrix}\begin{pmatrix} \beta \\ \gamma \end{pmatrix},$$

which is a system of $G\text{-}1\text{+}K$ equations in $H\text{-}1\text{+}J$ unknowns (unknown structural coefficients).

Assume Equation (4.1) is correct. Thus, all $G\text{-}1\text{+}K$ equations of (4.26) are satisfied by the true population parameter vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$, since all of these equations are consequences of (4.1). However, $G\text{-}1\text{+}K > H\text{-}1\text{+}J$; that is, (4.26) has more equations than unknown structural parameters. Therefore, there is a possibility of extracting additional information if we substitute the true vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ into system (4.26). Indeed, we might find necessary conditions for (4.1) to be correct. Yet, the true parameter vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ is unknown to us, by (3.4). Hence, in order to investigate the possibility of extracting such additional information, we must find some way of obtaining the true parameter vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ from the information we have at hand.

Since the true parameter vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ satisfies the entire system (4.26), it is a solution to any subsystem of (4.26). In particular, if we choose any subsystem of $H\text{-}1\text{+}J$ equations from (4.26), we get at least one solution (the true parameter vector). For various reasons, it is possible to extract a particular subsystem of $H\text{-}1\text{+}J$ equations from (4.26) such that one of the solutions has as its expected value the true parameter vector and such that the matrix of coefficients (of this solution) is nonsingular; that is, invertible. Therefore, if we substitute this obtained solution into system (4.26), we should be able to extract the same (in an expectational sense) additional information as if we were to substitute the true parameter vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ into (4.26).

Now, let us examine this solution to (4.26) which has as its expected value the true parameter vector $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$.

Let us partition (4.26) into two subsystems so that one subsystem is

$$(4.27) \qquad E(I_1y_g) = E(I_1y_1^* \;\; I_1z_1)\begin{pmatrix} \beta \\ \gamma \end{pmatrix},$$

a system of $H\text{-}1\text{+}J$ equations in $H\text{-}1\text{+}J$ unknowns, and the other subsystem is

$$(4.28) \qquad E(I_2y_g) = E(I_2y_1^* \;\; I_2z_1)\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$$

a system of $G\text{-}H\text{+}K\text{-}J$ equations in $H\text{-}1\text{+}J$ unknowns. It is clear that this partitioning corresponds to the distinction initially made in (4.19) and (4.20) when we partitioned I. That is, system (4.27) is the system of equations derived from (4.9) by using as instruments those predetermined variables appearing in (4.9) (i.e., the elements of I_1, $I_1 = \begin{pmatrix} y_1^{*\prime} \\ z_1' \end{pmatrix}$) and system (4.28) is the system derived from (4.9) by using as instruments those predetermined variables not appearing in (4.9) (i.e., the elements of I_2, $I_2 = \begin{pmatrix} y_2^{*\prime} \\ z_2' \end{pmatrix}$).

If we solve (4.27) for $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$, we obtain a solution which has as its expected value the true parameter vector.[23] $E(I_1y_1^* \;\; I_1z_1)$ is a square matrix of order $H\text{-}1\text{+}J$. By Assumption (3.4), $E(I_1y_1^* \;\; I_1z_1)$ exists. By the assumption that (4.1) is identified, $E(I_1y_1^* \;\; I_1z_1)$ is nonsingular. Therefore, the inverse of $E(I_1y_1^* \;\; I_1z_1)$ exists. Hence, by Cramer's rule, we solve (4.27) for $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ and we get:

$$(4.29) \qquad \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = [E(I_1y_1^* \;\; I_1z_1)]^{-1}E(I_1y_g),$$

a unique solution to the subsystem (4.27).

To show that (4.29) has as its expected value the true parameter vector $\binom{\beta}{\gamma}$, let us denote (4.29) by $\binom{\beta}{\gamma}''$. We define the error of (4.29) to be the difference between (4.29) and the true parameter vector $\binom{\beta}{\gamma}$, so that:

(4.30) $$\underset{\sim}{e} = \binom{\beta}{\gamma}'' - \binom{\beta}{\gamma},$$

where $\underset{\sim}{e}$ is the error of $\binom{\beta}{\gamma}''$. Our task is to show that $E(e) = \underset{\sim}{0}$.

Substitution of (4.9) into (4.29) for y_g yields:

(4.31) $$\binom{\beta}{\gamma}'' = [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E\{\underset{\sim}{I}_1[(\underset{\sim}{y}_1^* \;\; \underset{\sim}{z}_1)\binom{\beta}{\gamma} + u_g + \underset{\sim}{v}_1\underset{\sim}{\beta}]\}$$

(4.32) $$= [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1\underset{\sim}{I}_1')\binom{\beta}{\gamma} + [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1u_g) + [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1\underset{\sim}{v}_1\beta)$$

(4.33) $$= [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)\binom{\beta}{\gamma} + [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1u_g) + [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1\underset{\sim}{v}_1\underset{\sim}{\beta})$$

(4.34) $$= \binom{\beta}{\gamma} + [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1u_g) + [E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)]^{-1}E(\underset{\sim}{I}_1\underset{\sim}{v}_1\underset{\sim}{\beta})$$

(4.35) $$= \binom{\beta}{\gamma} + 0 + 0;$$

by Assumption (3.5) and Equations (3.13), (4.10), and (4.11). Hence $E(\underset{\sim}{e}) = \underset{\sim}{0}$. Therefore, $E\binom{\beta}{\gamma}'' = \binom{\beta}{\gamma}$.

Earlier we pointed out that, by substituting an appropriate solution into (4.26), we should be able to extract some additional information. Instead of substituting $\binom{\beta}{\gamma}''$ directly into (4.26), we will substitute it into another system, each of which can be derived from the other.

Let us define

(4.36) $$\underset{\sim}{\sigma} = \begin{pmatrix} \sigma_{y_1^*} & 0 & \cdots & & & & 0 \\ 0 & \ddots & & & & & \vdots \\ \vdots & & \sigma_{y_H^*} & & & & \vdots \\ \vdots & & & \sigma_{z_1} & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ \vdots & & & & & \sigma_{z_J} \;\; \sigma_{y_{H+1}^*} & 0 \\ 0 & \cdots & & & & 0 \;\; \ddots & \sigma_{z_K} \end{pmatrix}$$

to be the $G-1+K \times G-1+K$ diagonal matrix of population standard deviations of the elements of $\underset{\sim}{I}$. We see that

(4.37) $$\underset{\sim}{\sigma} = \begin{pmatrix} \underset{\sim}{\sigma}_1 & 0 \\ 0 & \underset{\sim}{\sigma}_2 \end{pmatrix},$$

where $\underset{\sim}{\sigma}_1$ is the $H-1+J \times H-1+J$ diagonal matrix of population standard deviations of the elements of $\underset{\sim}{I}_1$ and $\underset{\sim}{\sigma}_2$ is the corresponding $G-H+K-J \times G-H+K-J$ matrix for the elements of $\underset{\sim}{I}_2$. Similarly, let σ_g be the population standard deviation of y_g.

Since none of the diagonal elements of $\underset{\sim}{\sigma}$ is zero (assuming none of our variables degenerates to a constant), σ is invertible. Therefore, $\underset{\sim}{\sigma}_1$ is also invertible, as well as the scalar σ_g. Premultiplying system (4.26) by $\underset{\sim}{\sigma}^{-1}$, postmultiplying by σ_g^{-1}, and noting that $\underset{\sim}{\sigma}_1^{-1}\underset{\sim}{\sigma}_1$ is the identity matrix, we have:

(4.38) $$\underset{\sim}{\sigma}^{-1}E\begin{pmatrix} \underset{\sim}{I}_1y_g \\ \underset{\sim}{I}_2y_g \end{pmatrix}\sigma_g^{-1} = \underset{\sim}{\sigma}^{-1}E\begin{pmatrix} \underset{\sim}{I}_1\underset{\sim}{y}_1^* & \underset{\sim}{I}_1\underset{\sim}{z}_1 \\ \underset{\sim}{I}_2\underset{\sim}{y}_1^* & \underset{\sim}{I}_2\underset{\sim}{z}_1 \end{pmatrix}\underset{\sim}{\sigma}_1^{-1}\underset{\sim}{\sigma}_1\binom{\beta}{\gamma}\sigma_g^{-1}$$

Since a diagonal matrix is symmetric, and since we assume that all variables in (3.2) have zero means, and since the elements of $\underset{\sim}{y}_1^*$ and $\underset{\sim}{y}_2^*$ have zero means (since they are linear combinations of variables with zero means), (4.38) becomes:

(4.39) $$\begin{pmatrix} \underset{\sim}{P}_{I_1y_g} \\ \underset{\sim}{P}_{I_2y_g} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{P}_{I_1y_1^*} & \underset{\sim}{P}_{I_1z_1} \\ \underset{\sim}{P}_{I_2y_1^*} & \underset{\sim}{P}_{I_2z_1} \end{pmatrix}\underset{\sim}{\sigma}_1\binom{\beta}{\gamma}\sigma_g^{-1}$$

where

(4.40) $$\underset{\sim}{P}_{I_1y_g} = \begin{pmatrix} \rho_{y_1^*y_g} \\ \vdots \\ \rho_{y_H^*y_g} \\ \rho_{z_1y_g} \\ \vdots \\ \rho_{z_Jy_g} \end{pmatrix}$$

is the $H-1+J \times 1$ column vector of population correlation coefficients between the elements of $\underset{\sim}{I}_1$ and y_g;

(4.41) $$\underset{\sim}{P}_{I_2y_g} = \begin{pmatrix} \rho_{y_{H+1}^*y_g} \\ \vdots \\ \rho_{y_G^*y_g} \\ \rho_{z_{J+1}y_g} \\ \vdots \\ \rho_{z_Ky_g} \end{pmatrix}$$

is the $G-H+K-J \times 1$ column vector of population correlation coefficients between the elements of $\underset{\sim}{I}_2$ and y_g;

(4.42) $$\underset{\sim}{P}_{I_1y_1^*} = \begin{pmatrix} \rho_{y_1^*y_1^*} & \cdots & \rho_{y_1^*y_H^*} \\ \vdots & & \vdots \\ \rho_{y_H^*y_1^*} & \cdots & \rho_{y_H^*y_H^*} \\ \rho_{z_1y_1^*} & \cdots & \rho_{z_1y_H^*} \\ \vdots & & \vdots \\ \rho_{z_Jy_1^*} & \cdots & \rho_{z_Jy_H^*} \end{pmatrix}$$

is the $H-1+J \times H-1$ matrix of population correlation coefficients between the elements of $\underset{\sim}{I}_1$ and the elements of $\underset{\sim}{y}_1^*$; similarly,

(4.43) $\underset{\sim}{P}_{I_1z_1}$ is the $H-1+J \times J$ matrix of population correlation coefficients between the elements of $\underset{\sim}{I}_1$ and $\underset{\sim}{z}_1$;

(4.44) $\underset{\sim}{P}_{I_2y_1^*}$ is the $G-H+K-J \times H-1$ matrix of population correlation coefficients between the elements of $\underset{\sim}{I}_2$ and $\underset{\sim}{y}_1^*$;

and

(4.45) $\underset{\sim}{P}_{I_2z_1}$ is the $G-H+K-J$ matrix of population correlation coefficients between the elements of $\underset{\sim}{I}_2$ and $\underset{\sim}{z}_1$.

System (4.39) is the above mentioned system into which we will substitute $\binom{\beta}{\gamma}''$.

We will now examine the necessary and sufficient conditions for (4.1) to be correct. We have assumed (4.1) to be identified; it may be just identified (i.e., $K-J = H-1$) or overidentified (i.e., $K-J > H-1$). If (4.1) is just identified, certain (necessary) conditions will hold. If (4.1) is overidentified, then we will extract a suitable amount of information such that we can treat (4.1) to be just identified, on the basis of only this information. We will then find necessary and sufficient conditions for the "left-over" information to hold. It is this latter case in which we are most interested.

Case I. Assume (4.1) is just identified (i.e., $K-J = H-1$). Let us define

(4.46) $$\underset{\sim}{\beta}^2 = \begin{pmatrix} \beta_{g,H+1} \\ \vdots \\ \beta_{gG} \end{pmatrix}$$ to be the $G-H \times 1$ column vector of coefficients of the elements of $\underset{\sim}{y}_2$ (the endogeneous variables excluded from (4.1)) and

$$(4.47)\quad \underset{\sim}{\gamma}^2 = \begin{pmatrix} \gamma_{g,J+1} \\ \vdots \\ \gamma_{gK} \end{pmatrix} \text{ to be the } K\text{-}J \times 1 \text{ column vector of}$$

coefficients of the elements of $\underset{\sim}{z}_2$ (the predetermined variables excluded from (4.1)).

Furthermore, let

$$(4.48)\quad \underset{\sim}{Q} = \begin{pmatrix} \underset{\sim}{\beta}^2 \\ \underset{\sim}{\gamma}^2 \end{pmatrix}$$

Let us also define

$$(4.49)\quad \underset{\sim}{P}_{\underset{\sim}{I}_2 y_g.\underset{\sim}{I}_1} = \begin{pmatrix} \rho_{y_{H+1}y_g.y_1^*\cdots y_H^* z_1\cdots z_J} \\ \vdots \\ \rho_{y_G y_g.y_1^*\cdots y_H^* z_1\cdots z_J} \\ \rho_{z_{J+1}y_g.y_1^*\cdots y_H^* z_1\cdots z_J} \\ \vdots \\ \rho_{z_K y_g.y_1^*\cdots y_H^* z_1\cdots z_J} \end{pmatrix}$$

to be the G-H+K-J × 1 vector of partial correlation coefficients between the elements of $\underset{\sim}{I}_2$ and y_g, controlling for every element of $\underset{\sim}{I}_1$. Let us state the following theorem:

Theorem 4.1: If $\underset{\sim}{Q} = \underset{\sim}{0}$, then $\underset{\sim}{P}_{\underset{\sim}{I}_2 y_g.\underset{\sim}{I}_1} = \underset{\sim}{0}$.

Essential to the proof of Theorem 4.1 is the following lemma about partial correlation coefficients. Let us state, without proof, this easily verified lemma.

Lemma 4.1: $\rho_{12.34\ldots k} = 0$ if and only if $\rho_{12} = B_{23}\rho_{13} + B_{24}\rho_{14} + \cdots + B_{2k}\rho_{1k}$ where $B_{ij} = \beta_{ij}\frac{\sigma_j}{\sigma_i}$.

Proof of Theorem 4.1: By the assumption that $\underset{\sim}{Q} = \underset{\sim}{0}$, (4.1) is identified. Therefore, $\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$" exists. Substitution of $\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$" into (4.39) for $\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$ yields a set of relations which, by Lemma 4.1, implies that $\underset{\sim}{P}_{\underset{\sim}{I}_2 y_g.\underset{\sim}{I}_1} = \underset{\sim}{0}$. Q.E.D.

We see that when (4.1) is just identified $\underset{\sim}{Q} = \underset{\sim}{0}$, since both $\underset{\sim}{\beta}^2$ and $\underset{\sim}{\gamma}^2$ are equal to zero, by the assumption that the elements of $\underset{\sim}{y}_2$ and $\underset{\sim}{z}_2$ are excluded from (4.1). Therefore, by Theorem 4.1, for the just identified case, $\underset{\sim}{P}_{\underset{\sim}{I}_2 y_g.\underset{\sim}{I}_1} = \underset{\sim}{0}$.

However, Theorem 4.1 does not really provide us with a test (i.e., necessary and sufficient conditions) of the correctness of the identifying restrictions on equation (4.1). We have merely provided a necessary condition for these restrictions to be correct. This is not a sufficient condition for the correctness of these restrictions. Indeed, if the restrictions on (4.1) were incorrect (i.e., if some elements of $\underset{\sim}{Q}$ were nonzero), then (4.1) would be underidentified. Therefore, $\begin{pmatrix}\beta \\ \gamma\end{pmatrix}$" would not exist, since $E(\underset{\sim}{I}_1\underset{\sim}{y}_1^* \;\; \underset{\sim}{I}_1\underset{\sim}{z}_1)$ would be singular and, hence, not invertible. Thus, using our method, we cannot deduce any sufficient conditions for the identifying restrictions to be correct in the just identified case.

Case II. Assume (4.1) is overidentified (i.e., K-J > H-1).

We have seen that our method cannot provide a test of the correctness of the a priori exclusion restrictions on (4.1) if (4.1) is just identified. However, if (4.1) is overidentified, then our method does provide a test for some of the a priori exclusion restrictions on (4.1) to be correct. If (4.1) is overidentified, then there are more exclusion restrictions on (4.1) than are needed to insure its identifiability. Therefore, if we choose a subset of these a priori exclusion restrictions -- just enough to guarantee the identifiability (4.1) -- and consider the restrictions contained in this subset to be untestable, then we may consider the remainder of the exclusion restrictions on (4.1) to be testable. That is, by our method, we are able to find necessary and sufficient conditions for the correctness of these overidentifying (extra) restrictions, given that our chosen subset is assumed to be sufficient to identify (4.1).[24]

Let us assume that, for theoretical reasons, we feel that the restrictions contained in our chosen subset have a stronger basis than the remainder of the restrictions on (4.1). Let us additionally assume that even if our test rejects the validity of all of the overidentifying restrictions on (4.1), the identifiability of all other equations in (3.2) is unaffected. Renumbering if necessary, we can order the variables excluded from (4.1) (i.e., the elements of $\underset{\sim}{y}_2$ and $\underset{\sim}{z}_2$) so that:

$$(4.57)\quad \underset{\sim}{y}_2 = (y_{H+1},\ldots,y_{H'},y_{H'+1},\ldots,y_G)$$

and

$$(4.58)\quad \underset{\sim}{z}_2 = (z_{J+1},\ldots,z_{J'},z_{J'+1},\ldots,z_K),$$

where the exclusion from (4.1) of the set of variables $\{y_{H'+1},\ldots,y_G,z_{J'+1},\ldots,z_K\}$ suffices to identify (4.1). The number of elements in this set is G-H'+K-J' = G-1 (since an equation in a linear system of G equations must exclude at least G-1 of the variables that appear in the entire system in order to be identified and exactly G-1 to be just identified). Let us assume that the exclusion of this set of variables is taken as unquestioned (for theoretical reasons). Let the exclusion of the set of variables $\{y_{H+1},\ldots,y_{H'},z_{J+1},\ldots,z_{J'}\}$ from (4.1) be the overidentifying restrictions on (4.1). The number of elements in this set is H'-H+J'-J. Since the exclusion of the variables contained in the set $\{y_{H'+1},\ldots,y_G,z_{J'+1},\ldots,z_K\}$ suffices to identify that equation, the exclusion of the variables of the set $\{y_{H+1},\ldots,y_{H'},z_{J+1},\ldots,z_{J'}\}$ is a testable assumption.

Now, partition $\underset{\sim}{y}_2$ and $\underset{\sim}{z}_2$ so that

$$(4.59)\quad \underset{\sim}{y}_2 = (\underset{\sim}{y}_2^1 \;\; \underset{\sim}{y}_2^2)$$

$$\text{where } \underset{\sim}{y}_2^1 = (y_{H+1},\ldots,y_{H'}), \quad \underset{\sim}{y}_2^2 = (y_{H'+1},\ldots,y_G)$$

and

$$(4.60)\quad \underset{\sim}{z}_2 = (\underset{\sim}{z}_2^1 \;\; \underset{\sim}{z}_2^2)$$

$$\text{where } \underset{\sim}{z}_2^1 = (z_{J+1},\ldots,z_{J'}), \quad \underset{\sim}{z}_2^2 = (z_{J'+1},\ldots,z_K)$$

Corresponding to the distinction just made and recalling that $\underset{\sim}{I} = \begin{pmatrix}\underset{\sim}{I}_1 \\ \underset{\sim}{I}_2\end{pmatrix}$, let us partition $\underset{\sim}{I}_2$ so that

$$(4.61)\quad \underset{\sim}{I}_2 = \begin{pmatrix} \underset{\sim}{I}_2^1 \\ \underset{\sim}{I}_2^3 \\ \underset{\sim}{I}_2^2 \\ \underset{\sim}{I}_2^4 \end{pmatrix}.$$

where $\underset{\sim}{I}_2^1 = (\underset{\sim}{y}_2^{1*})$, $\underset{\sim}{I}_2^3 = (\underset{\sim}{y}_2^{2*})$, $\underset{\sim}{I}_2^2 = (\underset{\sim}{z}_2^1)$, and $\underset{\sim}{I}_2^4 = (\underset{\sim}{z}_2^2)$.

Likewise, we partition (4.39), so that:

$$(4.62)\quad \begin{pmatrix} \underset{\sim}{P}_{\underset{\sim}{I}_1 y_g} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^1 y_g} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^3 y_g} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^2 y_g} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^4 y_g} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{P}_{\underset{\sim}{I}_1\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_1\underset{\sim}{z}_1} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^1\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_2^1\underset{\sim}{z}_1} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^3\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_2^3\underset{\sim}{z}_1} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^2\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_2^2\underset{\sim}{z}_1} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^4\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_2^4\underset{\sim}{z}_1} \end{pmatrix} \underset{\sim}{\sigma}_1 \begin{pmatrix}\beta \\ \gamma\end{pmatrix} \sigma_g^{-1}$$

We can rearrange (4.62) into two subsystems so that one subsystem is:

$$(4.63)\quad \begin{pmatrix} \underset{\sim}{P}_{\underset{\sim}{I}_1 y_g} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^1 y_g} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^2 y_g} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{P}_{\underset{\sim}{I}_1\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_1\underset{\sim}{z}_1} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^1\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_2^1\underset{\sim}{z}_1} \\ \underset{\sim}{P}_{\underset{\sim}{I}_2^2\underset{\sim}{y}_1^*} & \underset{\sim}{P}_{\underset{\sim}{I}_2^2\underset{\sim}{z}_1} \end{pmatrix} \underset{\sim}{\sigma}_1 \begin{pmatrix}\beta \\ \gamma\end{pmatrix} \sigma_g^{-1},$$

a system of H'-1+J' equations in H-1+J unknowns and the other subsystem is

$$(4.64)\qquad \begin{pmatrix} P_{I_2^3 y_g} \\ P_{I_2^4 y_g} \end{pmatrix} = \begin{pmatrix} P_{I_2^3 y_1^*} & P_{I_2^3 z_1} \\ P_{I_2^4 y_1^*} & P_{I_2^4 z_1} \end{pmatrix} \sigma_1 \begin{pmatrix} \beta \\ \gamma \end{pmatrix} \sigma_g^{-1},$$

a system of $G-H'+K-J'$ equations in $H-1+J$ unknowns, where (4.63) is the subsystem of (4.62) which corresponds to those equations derived from (4.9) by using as instruments the predetermined variables appearing in (4.9) (i.e., the elements of I_1) and the predetermined variables which are elements of I_2^1 and I_2^2, and where (4.64) is the subsystem of (4.62) which corresponds to those equations derived from (4.9) by using as instruments the predetermined variables of I_2^3 and I_2^4. Since the elements of I_2^3 and I_2^4 correspond to the set of variables $\{y_{H'+1},\dots,y_G, z_{J'+1},\dots,z_K\}$ whose exclusion from (4.1) is taken as unquestioned, we shall ignore (4.64) and concern ourselves only with (4.63).

Define

$$(4.65)\qquad Q = \begin{pmatrix} \beta_2^1 \\ \gamma_2^1 \end{pmatrix}$$

to be a $H'-H+J'-J \times 1$ column vector, where

$$(4.66)\qquad \beta_2^1 = \begin{pmatrix} \beta_{g,H+1} \\ \vdots \\ \beta_{gH'} \end{pmatrix}$$

is the $H'-H \times 1$ column vector of coefficients of the elements of y_2^1 and

$$(4.67)\qquad \gamma_2^1 = \begin{pmatrix} \gamma_{g,J+1} \\ \vdots \\ \gamma_{gJ'} \end{pmatrix},$$

is the $J'-J \times 1$ column vector of coefficients of the elements of z_2^1. Let

$$(4.68)\qquad \begin{pmatrix} P_{I_2^1 y_g . I_1} \\ P_{I_2^2 y_g . I_1} \end{pmatrix} = \begin{pmatrix} \rho_{y^*_{H+1} y_g . y_1^* \dots y_H^* z_1 \dots z_J} \\ \vdots \\ \rho_{y^*_{H'} y_g . y_1^* \dots y_H^* z_1 \dots z_J} \\ \rho_{z_{J+1} y_g . y_1^* \dots y_H^* z_1 \dots z_J} \\ \vdots \\ \rho_{z_{J'} y_g . y_1^* \dots y_H^* z_1 \dots z_J} \end{pmatrix}$$

be the $H'-H+J'-J \times 1$ vector of partial correlation coefficients between the elements of $\begin{pmatrix} I_2^1 \\ I_2^2 \end{pmatrix}$ and y_g, controlling for every element of I_1.

Consider the following theorem:

<u>Theorem 4.2</u>: $Q = 0$ if and only if

$$\begin{pmatrix} P_{I_2^1 y_g . I_1} \\ P_{I_2^2 y_g . I_1} \end{pmatrix} = 0.$$

<u>Proof:</u>

<u>Part I</u>: If $Q = 0$ then $\begin{pmatrix} P_{I_2^1 y_g . I_1} \\ P_{I_2^2 y_g . I_1} \end{pmatrix} = 0.$

<u>Proof of Part I</u>: By assumption, $Q = 0$. Therefore, $\beta_2^1 = 0$ and $\gamma_2^1 = 0$, which implies that the elements of y_2^1 and z_2^1 do not appear in equation (4.1). Since (4.1) is identified, $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}^{*}$ exists. Therefore, substituting $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}^{*}$ into (4.63), we obtain, by Lemma 4.1, our result:

$$\begin{pmatrix} P_{I_2^1 y_g . I_1} \\ P_{I_2^2 y_g . I_1} \end{pmatrix} = 0.$$

<u>Part II</u>: $\begin{pmatrix} P_{I_2^1 y_g . I_1} \\ P_{I_2^2 y_g . I_1} \end{pmatrix} = 0$, then $Q = 0$.

<u>Proof of Part II</u>: Assume $Q \neq 0$; then either $\beta_2^1 \neq 0$ or $\gamma_2^1 \neq 0$ (or both). Therefore there exists some j such that $H+1 \leq j \leq H'$ (or $J+1 \leq j \leq J'$) and $\beta_{gj} \neq 0$ (or $\gamma_{gj} \neq 0$). We now have k non-zero parameters in equation (4.1), where $k \geq H+J$. Therefore $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ has k rows. Since, by assumption, the elements of y_2^2 and z_2^2 are excluded from (4.1), then (4.1) is identified and $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}^{*}$ exists. Therefore substitution of $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}^{*}$ into (4.63) for $\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ yields an expression for $\rho_{y_j^* y_g}$ (or $\rho_{z_j y_g}$) which, by Lemma 4.1, implies that $\rho_{y_j^* y_g . y_1^* \dots y_H^* z_1 \dots z_J} \neq 0$ (or $\rho_{z_j y_g . y_1^* \dots y_H^* z_1 \dots z_J} \neq 0$). Therefore,

$$\begin{pmatrix} P_{I_2^1 y_g . I_1} \\ P_{I_2^2 y_g . I_1} \end{pmatrix} \neq 0.$$

This is a contradiction to our hypothesis; therefore we must have $Q = 0$. Q.E.D.

C. Interpretation of Results

By Definition (3.1), we can see why the results of Theorem 4.2 can be interpreted to be a statistical test of the adequacy of causal theories represented in the general linear structural model. If we are willing to make certain untestable <u>a priori</u> assumptions, then, by our statistical results, we are able to determine whether or not certain other restrictions on the equations of our model are correct. Since those restrictions which we test are exclusion restrictions, we have by Definition (3.1), that they are, additionally, restrictions about causal relations.

We have also accomplished our heuristic objective in deriving a statistical test which reflects the logical process associated with introducing control variables into a zero-order correlational relationship and examining the partial correlations as a test for "spuriousness." Our results give us precisely which partial correlations (if any) ought to be equal to zero in any given causal situation (as we have defined such a situation).

Furthermore, it should be easily seen that, as a consequence of Theorem (4.2), we can determine which one of the $C_0^{(K-1)} + C_1^{(K-1)} + \dots + C_{(K-1)}^{(K-1)} = 2^{(K-1)}$ possible causal mechanisms represented by (4.1) is the correct one, given the <u>a priori</u> just identifying restrictions on (4.1). We can do this by observing which set of partial correlational conditions holds, since it is obvious that only one set of these conditions holds for each one of the $2^{(K-1)}$ possibilities. Also, since (3.2) is a system of G equations, we can determine which of the $[2^{(K-1)}]^G$ possible causal theories represented by (3.2) is correct, given the just identifying restrictions on all g, $g = 1,\dots,G$.

While our test was developed only in an expectational sense, employing population parameters and expected values, we conjecture that our test can be approximated with information on samples of data. In fact, there is a close relationship between much of the derivation of our test and the parameter estimating method of two-stage least squares (See Footnote 23). A good part of our thinking was inspired by this method. Combining such an estimating procedure with the use of sample statistics rather than the population parameters used in this piece, one should be able, at least in principle, to approximate our test on sample data. However, as mentioned earlier when we discussed Simon's test for the recursive-form, the possibility of violation of untestable <u>a priori</u> assumptions is a serious problem in empirical work and, therefore, any application of our test to sample data should be done with a good deal of caution.

FOOTNOTES

[1]For example, see the following works: Alker (1966); Cnudde and McCrone (1966); Forbes and Tufte (1968); and Goldberg (1966).

[2]See Wold (1953), Wold (1960), and Strotz and Wold (1960), for discussion of the causal implications of the recursive-form.

[3]Throughout this paper, we shall use "causal relation" interchangeably with "direct causal relation."

[4]A good discussion of segmentable systems may be found in Christ (1966), pp. 61-62.

[5]See Simon (1953), p. 18, for the exact statement of his definition.

[6]The basic distinction between an endogeneous variable and a predetermined variable is the following: a predetermined variable is a variable which is independent of all disturbances in the model at time <u>t</u>. All other variables at time <u>t</u> are endogeneous variables. The distinction between these two types of variables will be made more explicit later in this paper.

[7]The notion of identifiability is not essential for an understanding of this section. We will discuss this notion in a later section.

[8]We shall consider the identification problems associated with triangular systems in a later section.

[9]See Simon (1954), p. 47.

[10]See Lazarsfeld (1959).

[11]We will use "recursive-form" interchangeably with "segmentable" in the following pages.

[12]See Bentzel and Hansen (1954-1955), pp. 153-168, for a good discussion of the problems associated with the use of the recursive-form for dealing with extended periods of time. Also see Samuelson (1965), pp. 139-40.

[13]Miller and Stokes (1963) initially desired to consider some reciprocal links in their model of constituency influence in Congress, but they never really carried out such an analysis. See Alker (1969) for a discussion of alternatives to the Miller and Stokes formulation.

[14]For a somewhat different approach to the relationship between time and causality see Fisher (1970). In this piece Fisher takes a different epistomological position than do we and admits a <u>logical</u> relationship between time and causality. He then examines the implications of the stance that simultaneous equation models are limiting approximations to nonsimultaneous ones as time lags go to zero. See also Granger (1969).

[15]See Bunge (1963), pp. 188-190 and J. Simon (1969), pp. 458-460.

[16]This paragraph is essentially a paraphrase of Fisher (1966), p. 92.

[17]Basically, an estimator is consistent if it converges in the probability limit to the "true" parameter being estimated.

[18]See Christ (1966), p. 317, for a proof of this.

[19]See Fisher (1966), pp. 36-39, for a proof.

[20]The data was generated by a program called DATSIM on the Berkeley Computer Center CDC 6400.

[21]We will include in this definition both variables which are independent of the disturbances at all times <u>t</u> (exogenous) and variables which are independent of the disturbances at a single time <u>t</u> (predetermined).

[22]See any standard econometric text for discussion of the general linear structural model. For example, Christ (1966) or Goldberger (1964).

[23]This is essentially the two-stage least squares solution. See Christ (1966), pp. 432-46; Goldberger (1964), pp. 329-36, or Johnston (1963), pp. 258-60 for discussions of the method of two-stage least squares.

[24]See Christ (1966), pp. 539-40.

*I wish to thank J. Merrill Shanks of the University of California, Berkeley and Hayward R. Alker, Jr. of the Massachusetts Institute of Technology for inspiration and helpful criticisms of an earlier draft of this paper. I also wish to thank Ruth Suzuki for her many hours on the typewriter.

REFERENCES

Alker, Hayward R., Jr. <u>Mathematics and Politics</u>. New York: Macmillan Co., 1965.

Alker, Hayward R., Jr. "Causal Inference and Political Analysis." in <u>Mathematical Applications in Political Science, II</u>. Edited by Joseph A. Bernd. Dallas: Southern Methodist University Press, 1966, pp. 3-43.

Alker, Hayward R., Jr. "Statistics and Politics: The Need for Causal Data Analysis." in <u>Politics and the Social Sciences</u>. Edited by Seymour Martin Lipset. New York: Oxford University Press, 1969, pp. 244-313.

Anderson, Theodore Wilbur, and Herman Rubin. "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations, <u>Annals of Mathematical Statistics</u>, 21 (December, 1950), pp. 570-582.

Basmann, Robert L. "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics, <u>Journal of the American Statistical Association</u>, 55 (December, 1960), pp. 650-659.

Bentzel, R. and Hansen, B. "On Recursiveness and Interdependency in Economic Models." <u>Review of Economic Studies</u>, XXII (1954-1955), pp. 153-68.

Blalock, Hubert M., Jr. "Four-Variable Causal Models and Partial Correlations." <u>American Journal of Sociology</u>, LXVIII (September, 1962), pp. 182-94.

Blalock, Hubert M., Jr. <u>Causal Inferences in Nonexperimental Research</u>. Chapel Hill: University of North Carolina Press, 1964.

Boudon, Raymond. "A New Look at Correlation Analysis." in <u>Methodology in Social Research</u>. Edited by Hubert M. Blalock, Jr. and Ann B. Blalock. New York: McGraw-Hill, Inc., 1968, pp. 199-235.

Bunge, Mario. <u>The Myth of Simplicity</u>. Englewood Cliffs: Prentice-Hall, Inc., 1963.

Christ, Carl F. <u>Econometric Models and Methods</u>. New York: John Wiley and Sons, Inc., 1966.

Cnudde, Charles F., and McCrone, Donald J. "Constituency Attitudes and Congressional Voting: A Causal Model." <u>American Political Science Review</u>, LX (March, 1966), pp. 66-72.

Fisher, Franklin M. <u>The Identification Problem in Econometrics</u>. New York: McGraw-Hill, 1966.

Fisher, Franklin M. "A Correspondence Principle for Simultaneous Equation Models," <u>Econometrica</u>, 38 (January, 1970), pp. 73-92.

Forbes, Donald, and Tufte, Edward. "A Note of Caution in Causal Modelling." <u>American Political Science Review</u>, LXII (December, 1968), pp. 1258-64.

Granger, C.W.J. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," <u>Econometrica</u>, 37 (July, 1969), pp. 424-38.

Goldberg, Arthur S. "Discerning a Causal Pattern among Data on Voting Behavior." <u>American Political Science Review</u>, LX (December, 1966), pp. 913-22.

Goldberger, Arthur S. <u>Econometric Theory</u>. New York: John Wiley and Sons, Inc., 1964.

Johnston, J. <u>Econometric Methods</u>. New York: John Wiley and Sons, Inc., 1963.

Kadane, Joseph B. "Testing a Subset of Overidentifying Restrictions," Unpublished Manuscript, 1970.

Kaplan, Abraham. <u>The Conduct of Inquiry</u>. San Francisco: Chandler Publishing Co., 1964.

Koopmans, Tjalling C. "Identification Problems in Economic Model Construction," in <u>Studies in Econometric Method</u>. Edited by William C. Hood and Tjalling C. Koopmans. Cowles Commission Monograph 14, New York: John Wiley and Sons, Inc., 1953, pp. 27-48.

Lazarsfeld, Paul F. "The Algebra of Dichotomous Systems," in <u>Item Analysis and Prediction</u>. Edited by H. Solomon. Stanford: Stanford University Press, 1959.

Liu, Ta-Chung and William J. Breen. "The Covariance Matrix of the Limited Information Estimator and the Identification Test," <u>Econometrica</u>, 37 (April, 1969), pp. 222-227.

Miller, Warren E., and Stokes, Donald E. "Constituency Influence in Congress." <u>American Political Science Review</u>, LVII (March, 1963), pp. 45-56.

Samuelson, Paul A. "Some Notions on Causality and Teleology in Economics," in <u>Cause and Effect</u>. Edited by Daniel Lerner. New York: The Free Press, 1965, pp. 99-143.

Simon, Herbert A. "Causal Ordering and Identifiability." in <u>Studies in Econometric Method</u>. Edited by William C. Hood and Tjalling C. Koopmans. Cowles Commission Monograph 14, New York: John Wiley and Sons, Inc., 1953, pp. 49-73. Reprinted in <u>Models of Man</u>. By Herbert A. Simon. New York: John Wiley and Sons, Inc., 1957, pp. 10-36.

Simon, Herbert A. "Spurious Correlation: A Causal Interpretation." <u>Journal of the American Statistical Association</u>, XLIX (September, 1954). Reprinted in <u>Models of Man</u>. By Herbert A. Simon. New York: John Wiley and Sons, Inc., 1957, pp. 37-49.

Simon, Julian L. <u>Basic Research Methods in Social Science: The Art of Empirical Investigation</u>. New York: Random House, 1969.

Strotz, Robert H. "Interdependence as a Specification Error," <u>Econometrica</u>, 28 (April, 1960), pp. 428-442.

Strotz, Robert H. and H.O.A. Wold. "Recursive vs. Nonrecursive Systems: An Attempt at Synthesis," <u>Econometrica</u>, 28 (April, 1960), pp. 417-427.

Wold, H.O.A., in association with Lars Jureen. <u>Demand Analysis: A Study in Econometrics</u>. New York: John Wiley and Sons, Inc., 1953.

Wold, H.O.A. "Ends and Means in Econometric Model Building. Basic Considerations Reviewed," in <u>Probability and Statistics</u>, Edited by U. Grenander. Stockholm: Almqvist & Wiksell, 1959.

Wold, H.O.A. "A Generalization of Causal Chain Models," <u>Econometrica</u>, 28 (April, 1960), pp. 443-463.