STRUCTURAL EQUATION MODELING, 9(4), 459–474

A New Identification Condition for Recursive Models With Correlated Errors

Carlos Brito and Judea Pearl
Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles

This article establishes a new criterion for the identification of recursive linear models in which some errors are correlated. We show that identification is ensured as long as error correlation does not exist between a cause and its *direct* effect; no restrictions are imposed on errors associated with indirect causes.

Before structural equation models (SEM) can be estimated and evaluated against data, a researcher must make sure that the parameters of the estimated model are *identified*, namely, that they can be determined uniquely from the population covariance matrix. The importance of testing identification prior to data analysis is summarized succinctly by Rigdon (1995):

> To avoid devoting research resources toward a hopeless cause (and to avoid ignoring productive research avenues out of an unfounded fear of underidentification), researchers need a way to quickly evaluate a model's identification status before data are collected. Furthermore, because models are often altered in the course of research (Jöreskog, 1993), researchers need a technique that helps them understand the impact of potential structural changes on the identification status of the model. (p. 359)

It is well known that, in recursive path models with correlated errors, the identification problem is unsolved. In other words, we are not in possession of a necessary and sufficient criterion for deciding whether the parameters in such a model can be computed from the population covariance matrix of the observed variables. Certain restricted classes of models are nevertheless known to be identifiable, and

Requests for reprints should be sent to Judea Pearl, Cognitive Systems Laboratory, Computer Science Department, University of California, Los Angeles, CA 90024. E-mail: judea@cs.ucla.edu

these are often assumed by researchers as a matter of convenience or convention (Duncan, 1975; Kang & Seneta, 1980; Wright, 1960).[1]

McDonald (1997) characterizes a hierarchy of three such classes: (a) uncorrelated errors, (b) correlated errors restricted to exogenous variables, and (c) correlated errors restricted to pairs of causally unordered variables (i.e., variables that are not connected by uni-directed paths). The structural equations in all three classes are regressional (i.e., the error term in each equation is uncorrelated with the explanatory variables of that same equation), hence the parameters can be estimated uniquely using ordinary least squares (OLS) techniques.[2]

Although more powerful algebraic and graphical methods have been developed for testing whether a specific parameter or a specific equation in a model is identifiable (see Bekker, Merckens, & Wansbeek, 1994; Fisher, 1966; Geraci, 1983; Pearl, 1998, 2000a; Spirtes, Richardson, Meek, Scheines, & Glymour, 1998), such methods are often too complicated for investigators to apply in the pre-analytic phase of model construction. Additionally, those specialized methods are limited in scope. The rank and order criteria (Fisher, 1966), for example, which are typical of the algebraic variety, do not exploit restrictions on the error covariances (if such are available). The rank criterion further requires a precise estimate of the covariance matrix before identifiability can be decided. Identification methods based on block recursive models (Fisher, 1966; Rigdon, 1995), for another example, insist on uncorrelated errors between any pair of ordered blocks. The "back-door" and "single-door" criteria, which are typical of the graphical variety (Pearl, 2000a, pp. 150–152), are applicable only in sparse models, that is, models rich in conditional independencies (e.g., zero partial correlations). The same holds for criteria based on instrumental variables (IV; Bowden & Turkington, 1984), since these require search for variables (called *instruments*) that are uncorrelated with the error terms in specific equations (see Relations to Instrumental Variables).

This article adds a new criterion to our repertoire of sufficient identification conditions, and thus uncovers a new class of identifiable models that, unlike the regressional hierarchy of McDonald (1997), permits errors associated with causally ordered variables (or blocks of variables) to be correlated and, unlike the IV method, does not rely on finding instruments. We show that identification is ensured as long as correlated errors are restricted to pairs of variables that are not *directly* linked (i.e., do not stand in *direct* causal relationship); no restrictions need be imposed on errors associated with indirectly linked variables. We shall call this class of models "bow-free," since their associated path diagrams are free of any

[1]Rigdon (1995) reports that, out of 108 models scanned in the applied SEM literature, 81 (75%) assumed that errors were uncorrelated.

[2]Bollen (1989) lumped the first two conditions under the rubric "Recursive Rule" (p. 104). We use the term *recursive* to denote acyclic systems of equations (i.e., without reciprocal causation or feedback loops) with unrestricted error covariances.

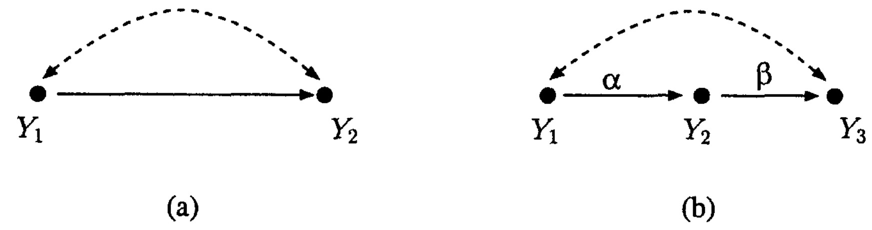


FIGURE 1 (a) A "bow-pattern"; (b) a bow-free model.

"bow pattern" (Pearl, 2000a)—a curved arc embracing a single arrow, as in Figure 1a—which typically represent unmeasured common cause of a dependent and an explanatory variable in the same equation.[3]

A simple example of a bow-free model is shown in Figure 1b. It represents the equations

$$\begin{aligned} Y_1 &= e_1 \\ Y_2 &= \alpha Y_1 + e_2 \\ Y_3 &= \beta Y_2 + e_3 \end{aligned} \tag{1}$$

with $cov(e_1, e_2) = cov(e_3, e_2) = 0$, and $cov(e_1, e_3) \neq 0$. The correlation between e_1 and e_3 renders the third equation nonregressional (since e_3 is correlated with Y_2), yet β is nevertheless identifiable, as can be verified by simple algebraic or graphical methods (Pearl, 2000a, pp. 155–156). For example, we can regard Y_2 as (conditional) instrumental variable, since, conditioned on Y_1,Y_2 is uncorrelated with e_3 (see Relations to Instrumental Variables).

A more elaborate bow-free model is shown in Figure 2.

Here, all errors are assumed to be correlated, except those corresponding to the five direct arrows, that is, $Y_1 \to Y_3$, $Y_1 \to Y_4$, $Y_2 \to Y_4$, $Y_3 \to Y_5$, $Y_4 \to Y_5$. Again, the equation for variable Y_5 is nonregressional, but, as we shall see, all parameters in this model are identifiable. Remarkably, the identifiability of this model is not recognized by available graphical methods (see Pearl, 1998, 2000a, pp. 149–154; Spirtes et al., 1998).

The bow-free condition thus extends existing identification criteria to a large class of models in which some omitted factors affect two or more variables, as long as the affected variables are not directly linked. Such models are quite com-

[3]McDonald called the bow-free condition "Direct Precedence Assumption" (R. P. McDonald, personal communication, November 15, 1993), as distinct of the "Precedence Assumption" discussed in McDonald (1997). Kenny, Kashy, and Bolger (1998) have in fact conjectured that such condition would ensure identification in both recursive and nonrecursive models. They stated, "Although there is no known proof of this condition, there is no known exception."

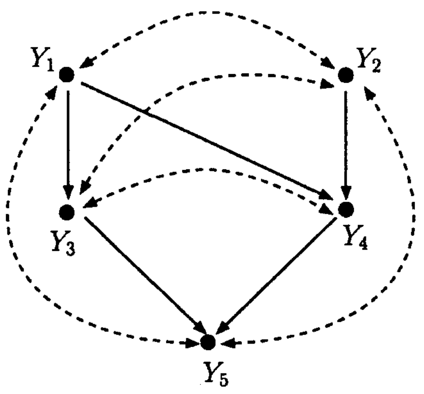


FIGURE 2 Bow-free model with five variables.

mon,[4] and can be recognized by simple inspection of the diagram (or the equations) without invoking algebraic operations or searches for conditional instrumental variables. The new condition can therefore serve as an effective, pre-analytic means of ascertaining the identifiability of candidate structural models under construction.

LINEAR RECURSIVE MODELS

A linear model for a set of random variables $Y = \{Y_1, \ldots, Y_n\}$ is defined by a set of equations of the form

$$Y_j = \sum_i c_{ji} Y_i + e_j, \quad j = 1, \ldots, n$$

and a matrix Ψ of variance/covariance for error-terms e_j.[5]

The equations and the pairs of error–terms (e_i, e_j) with nonzero correlation define the "structure" of the model. The model structure can be represented by a directed graph, called path diagram, in which the set of nodes is defined by the variables $Y_1, \ldots, Y_n$, and there is a directed edge from Y_i to Y_j if the coefficient of Y_i in

[4]All recursive models illustrated in Kenny (1979) and Bollen (1989) are bow-free.

[5]The set of Y variables for which all cs are zero are commonly called *exogenous* (or *predetermined*) and are often distinguished notationally (e.g., in Bollen, 1989, p. 81, exogenous variables are denoted $X_1, \ldots, X_q$); such distinction is not needed for our analysis. Additionally, unlike some of the econometric and SEM literature, we will not require that the exogenous variables be uncorrelated with the equation disturbances of all endogenous variables. Such requirements are rarely justified in nonexperimental studies, even when the exogenous variables represent policy decisions (see Pearl, 2000a, p. 136).

the equation for Y_j is distinct from zero. Additionally, if error–terms e_i and e_j have nonzero correlation, we add a (dashed) bidirected edge between Y_i and Y_j.

The models considered in this work are assumed to be recursive, that is, $c_{ij} = 0$ for $j \geq i$. The structural parameters of the model, denoted by θ, are the coefficients c_{ij}, and the values of the nonzero entries of the error covariance matrix $\Psi_{ij} = E(e_i e_j)$. We will also assume that all the Y variables are measured without errors, that the Y variables are standardized, and that all error–terms are unmeasured.

Fixing the model structure and assigning values to the parameters θ, the model determines a unique covariance matrix Σ over the observed variables $\{Y_1, ..., Y_n\}$, given by (see Bollen, 1989, p. 85)

$$\Sigma(\theta) = (I - C)^{-1} \Psi (I - C)^{-1} \tag{2}$$

where C is the matrix of structural coefficients, $[C]_{ij} = c_{ij}$.

Conversely, given the structure of a model, one may attempt to solve for C in terms of the observed covariance Σ. This is not always possible. In some cases, no parameterization of the model could be compatible with a given Σ. In other cases, as in Figure 1a, the structure of the model may permit several distinct solutions for the parameters. In these cases, the model is called *nonidentifiable*. Our task is to find conditions on the model structure that would guarantee a unique solution for the parameters if such a solution exists. Additionally, since the conditions we seek involve the structure of the model alone, not a particular numerical value of the parameters of θ, we will insist on having a unique solution for *almost all* values of θ, allowing for some pathological exceptions as formulated below.

A convenient way of relating parameter identification to the structure of the model is to write Equation 2 for each term σ_{ij} of Σ using Wright's (1960) method of path coefficients. Wright's method consists of equating the (standardized) covariance σ_{ij} with the sum of products of path coefficients and error covariances along the unblocked paths connecting Y_i and Y_j (examples are given in Example). Whenever the resulting equations give a unique solution to some path coefficient c_{ij}, independently of the (unobserved) residual correlations, that coefficient is said to be identifiable.

In this article we establish a sufficient condition for the identification of all structural parameters.

IDENTIFICATION OF BOW-FREE MODELS

We claim, in almost every case, given the structure of a recursive linear bow-free model, and a covariance matrix Σ that is compatible with the model, all structural parameters can be determined uniquely. This claim is formalized in the statement of Theorem 1.

Theorem 1

Let M be a bow-free model, and let θ be the set of parameters of M, that is, $\theta = \{\Psi, C\}$ such that $\Psi_{ij} = 0$ whenever $[C]_{ij} \neq 0$. Then, for almost all θ, we have

$$\Sigma(\theta) = \Sigma(\theta') \text{ implies } \theta = \theta'$$

In other words, if two sets of parameters, θ and θ', give rise to the same covariance Σ, then θ and θ' must be identical, except when θ resides in a set of zero Lebesgue measure.

Before proving the theorem, some definitions and preliminary results are required.

Definition 1

A *path* in a graph is a sequence of edges (directed or bidirected) such that each edge starts in the node ending the preceding edge. A *directed path* is a path composed only by directed edges, all oriented in the same direction. We say that node X is an *ancestor* of node Y if there is a directed path from X to Y. A path is said to be *blocked* if there is a node Z and a pair of consecutive edges in the path such that both edges are oriented toward Z (e.g., $\ldots \rightarrow Z \leftarrow \ldots$). In this case, Z is called a *collider.*

For example, the path $Y_1 \rightarrow Y_4 \leftrightarrow Y_3 \rightarrow Y_5$ in Figure 2 is blocked, because Y_4 is a collider. The path $Y_1 \rightarrow Y_4 \rightarrow Y_5$, on the other hand, is a directed path from Y_1 to Y_5, and it is unblocked.

Definition 2

The *depth* of a node in a directed acyclic graph (DAG) is the length (in number of edges) of the longest directed path between the node and any of its ancestors. Nodes with no ancestors have a depth of zero.

For example, in Figure 2 nodes Y_1 and Y_2 have a depth of zero, Y_3 and Y_4 have a depth of one, and Y_5 has a depth of two.

Lemma 1

Let X,Y be nodes in the path diagram of a recursive model, such that $depth(X) \geq depth(Y)$. Then, every path between X and Y that includes a node Z with $depth(Z) \geq depth(X)$ must be blocked by a collider.

Proof

Consider a path p between X and Y and node Z satisfying the conditions above. We observe that Z cannot be an ancestor of either X or Y, otherwise we would have $depth(Z) < depth(X)$ or $depth(Z) < depth(Y)$.

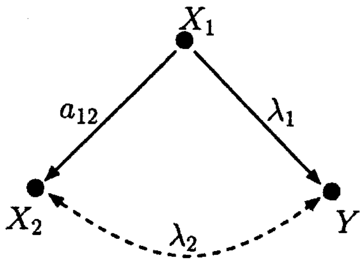


FIGURE 3 Simple linear model.

Now, consider the subpath of p between Z and Y. If this subpath has the form $Z \rightarrow \ldots Y$, then it must contain a collider, since it cannot be a directed path from Z to Y. Similarly, if the subpath of p between X and Z has the form $X \ldots \leftarrow Z$, then it must contain a collider.

In all the remaining cases Z is a collider blocking the path.

Definition 3

For each node Y, the set of incoming edges to Y, denoted $I(Y)$, is defined as the union of (a) the set of all directed edges oriented toward Y, and (b) the set of all bidirected edges between Y and a node X, such that $depth(X) < depth(Y)$.

Lemma 2

Let Y be a variable at depth k of a bow-free model. Assume that the parameter of every edge connecting variables at depth smaller than k is identifiable. Then, the parameter of each edge in $I(Y)$ is identifiable almost everywhere.

Proof

Let $\mathbf{X} = \{X_1, \ldots, X_m\}$ be the set of variables at depth smaller than k that are connected to Y by some directed or undirected edge. The bow-free property of the model implies a one-to-one correspondence between the variables in $\mathbf{X}$ and the elements of $I(Y)$. Thus, we can write $I(Y) = \{(X_1,Y), \ldots, (X_m, Y)\}$. For example, in the model illustrated in Figure 3, we have $\mathbf{X} = \{X_1, X_2\}$ and $I(Y) = \{(X_1 \rightarrow Y), (X_2 \leftrightarrow Y)\}$.

Applying Wright's (1960) method to each pair (X_i, Y), $X_i \in \mathbf{X}$, we obtain the following equations:

$$\sigma_{X_i,Y} = \sum_{\text{paths } p_l} T(p_l), \; i = 1, \ldots, m \tag{3}$$

where the term $T(p_l)$ is the product of the parameters of the edges along the path p_l, and the summation is over all unblocked paths between X_i and Y.

For $i = 1, \ldots, m$, we denote the parameter of edge (X_i, Y) by λ_i. Then, we can rewrite the set of equations as

$$\sigma_{X_1,Y} = \sum_{j=1}^{m} a_{1j} \cdot \lambda_j$$

$$\ldots$$

$$\sigma_{X_mY} = \sum_{j=1}^{m} a_{mj} \cdot \lambda_j$$

where the terms in a_{ij} correspond to unblocked paths between X_i and Y including edge (X_j, Y).

In the model of Figure 3, the only unblocked path between X_1 and Y consists of the edge $X_1 \rightarrow Y$, since the path $X_1 \rightarrow X_2 \leftrightarrow Y$ is blocked by the collider X_2. Also, we have two unblocked paths between X_2 and Y: $X_2 \leftarrow X_1 \rightarrow Y$, and $X_2 \leftrightarrow Y$. Thus, equations (3) for this model read

$$\sigma_{X_1,Y} = \lambda_1$$
$$\sigma_{X_2,Y} = a_{21}\lambda_1 + \lambda_2$$

Nonlinear terms (i.e., $\lambda_j\lambda_i$) are excluded from these equations because each unblocked path from X_i to Y contains exactly one element of $I(Y)$. Moreover, it follows from Lemma 1 and the assumptions of the lemma, that every factor in a_{ij} is an identifiable parameter. This yields a linear system of m equations (one for each X_i) and m unknown, $\lambda_1, \ldots, \lambda_m$, that we can write in matrix notation as $\Sigma = A \cdot \Lambda$.

Writing in matrix form the system of equations for the model in Figure 3, we obtain

$$\begin{bmatrix} \sigma_{X_1,Y} \\ \sigma_{X_2,Y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

The parameters λ_is are identifiable if the system of linear equations has unique solution, and this is the case if and only if $\text{Det}(A) \neq 0$.

We now observe some special properties of matrix A. First, note that, in bow-free models, the only unblocked path between X_i and Y including edge (X_i, Y) consists precisely of such edge. Hence, $a_{ii} = 1, i = 1, \ldots, m$. On the other hand, for $i \neq j$, entry a_{ij} is obtained by considering unblocked paths between X_i and Y which include edge (X_j, Y). Since $X_i \neq X_j$, any such path must have at least two edges.

Hence, a_{ij} is either zero (if no such path exists) or a polynomial in the parameters of the model with no constant term.

Thus, the general form of matrix A is

$$A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1m} \\ a_{21} & 1 & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & 1 \end{bmatrix}$$

From the definition of determinant, we can write

$$Det(A) = 1 + T$$

where term T is either zero or a polynomial in the parameters of the model with no constant term. Thus, Det(A) only vanishes on the set of zeros of the polynomial $1 + T$. And, from Okamoto (1973), we can further conclude that this set has a Lebesgue measure of zero.

Proof of Theorem 1

We argue by induction on the depth of the variables. At step k, we show that all the parameters associated with edges connecting variables at depth smaller than or equal to k are identifiable.

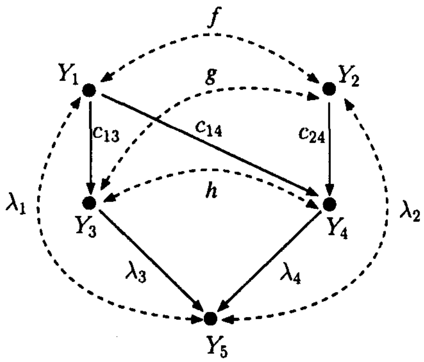
Let Y be a variable at depth k.

If $k = 0$, then Y is a root and $I(Y)$ is empty. To show that the parameter of a bidirected edge connecting two roots Y and Y' is identifiable, we note that the only unblocked path between Y and Y' is precisely given by edge (Y, Y'). In this case, according to Wright's (1960) method, the parameter of edge (Y, Y') is identifiable and given by the correlation coefficient between Y and Y'.

Now, assume that the parameter of every edge connecting two variables at depth smaller than k is identifiable. Then, by Lemma 2, the parameters of edges in $I(Y)$ are identifiable.

Once we know that the parameters of edges in $I(Y)$ are identifiable, for every variable Y at depth k, we can easily show that the parameter of a bidirected edge connecting two variables at depth k is identifiable (see Example). This follows by writing the corresponding Wright's equation and noting that every unblocked path between two such variables is either the desired bidirected edge itself, or a path labeled with identifiable parameters.

Since the model has a finite number of variables, and for each variable Y the parameters of edges in $I(Y)$ are all identifiable except in a set of Lebesgue measure zero, the statement of the theorem holds.



$$\Sigma = \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & 1 & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & 1 & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & 1 & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & 1 \end{bmatrix}$$

FIGURE 4 A bow-free model illustrating the proof of Theorem 1.

EXAMPLE

Assume we are given the model structure and the matrix Σ shown in Figure 4. Since the structure of the model is bow-free, Theorem 1 ensures that all model parameters are identifiable. In the following, we show how to find expressions for the model's parameters, starting from the roots of the graph (corresponding to the exogenous variables Y_1 and Y_2) and proceeding in increasing depths.

Wright's (1960) equation for the pair (Y_1, Y_2) gives

$$\sigma_{12} = f$$

Next, we consider variable Y_3. Wright's equations for pairs (Y_1, Y_3) and (Y_2, Y_3) read

$$\begin{cases} \sigma_{13} = c_{13} \\ \sigma_{23} = g + c_{13} f \end{cases}$$

from which we obtain two additional parameters: $c_{13} = \sigma_{13}$ and $g = \sigma_{23} - \sigma_{12} \cdot \sigma_{13}$.

Now, we consider variable Y_4. Wright's equations for pairs (Y_1, Y_4), (Y_2, Y_4), and (Y_3, Y_4) read

$$\begin{cases} \sigma_{14} = c_{14} + f c_{24} \\ \sigma_{24} = f c_{14} + c_{24} \\ \sigma_{34} = h + g c_{24} + c_{13} c_{14} + c_{13} f c_{24} \end{cases}$$

Solving for c_{24}, c_{14}, and h, we obtain

$$c_{24} = \frac{\sigma_{24} - f\sigma_{14}}{1-f^2} = \frac{\sigma_{24} - \sigma_{14}\cdot\sigma_{12}}{1-\sigma_{12}^2}; \; c_{14} = \frac{\sigma_{14} - f\sigma_{24}}{1-f^2} = \frac{\sigma_{14} - \sigma_{24}\cdot\sigma_{12}}{1-\sigma_{12}^2}$$

$$\begin{aligned} h &= \sigma_{34} - c_{24}\sigma_{23} - c\sigma_{13} \\ &= \sigma_{34} - \left[\frac{\sigma_{23}\cdot\sigma_{24} + \sigma_{13}\cdot\sigma_{14} - \sigma_{12}\cdot\sigma_{14}\cdot\sigma_{23} - \sigma_{12}\cdot\sigma_{13}\cdot\sigma_{24}}{1-\sigma_{12}^2}\right] \end{aligned}$$

Note that the edge (Y_3, Y_4) connects two variables of equal depth and the corresponding parameter, h, is not needed for the computation of c_{24}, c_{14}. h is obtained immediately from the corresponding equation for σ_{34}, once we solve for c_{24}, c_{14}.

Finally, we consider variable Y_5. Wright's equations for pairs (Y_1, Y_5), (Y_2, Y_5), (Y_3, Y_5), and (Y_4, Y_5) read

$$\begin{cases} \sigma_{15} = \lambda_1 + c_{13}\lambda_3 + \lambda_4\sigma_{14} \\ \sigma_{25} = \lambda_3\sigma_{23} + c_{24}\lambda_4 + \lambda_2 \\ \sigma_{35} = c_{13}\lambda_1 + \lambda_3 + \lambda_4\sigma_{34} \\ \sigma_{45} = c_{14}\lambda_1 + \lambda_3\sigma_{34} + \lambda_4 + c_{24}\lambda_2 \end{cases}$$

which need to be solved for the four depth-3 parameters: λ_1, λ_2, λ_3, λ_4.

To illustrate Theorem 1, we write these equations in matrix form:

$$\begin{bmatrix} \sigma_{15} \\ \sigma_{25} \\ \sigma_{35} \\ \sigma_{45} \end{bmatrix} = \begin{bmatrix} 1 & 0 & c_{13} & \sigma_{14} \\ 0 & 1 & \sigma_{23} & c_{24} \\ c_{31} & 0 & 1 & \sigma_{34} \\ c_{41} & c_{42} & \sigma_{43} & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix}$$

and we see that matrix A indeed has only 1's on the main diagonal. Furthermore, the determinant of A is given by

$$\begin{aligned} \det[A] = 1 + c_{13}^2c_{24}^2 - c_{13}^2 - c_{24}^2 - \sigma_{34}^2 + \\ c_{13}\sigma_{34}(c_{14} + \sigma_{14}) - c_{14}\sigma_{14} + c_{24}\sigma_{23}(\sigma_{34} - c_{13}\sigma_{14}) \end{aligned} \tag{4}$$

Thus, the system has unique solution except on the zeros of the polynomial in Equation 4. Solving for the parameters λ_1, λ_2, λ_3, and λ_4, we obtain

$$\lambda_4 = \frac{(1-c_{13}^2)(\sigma_{45} - c_{14}\sigma_{15} - c_{24}\sigma_{25}) - h(\sigma_{35} - c_{13}\sigma_{15})}{(1-c_{13}^2)(1 - c\sigma_{14} - c_{24}^2) - h(\sigma_{34} - c_{13}\sigma_{14})}$$

$$\lambda_3 = \frac{\sigma_{35} - c_{13}\sigma_{15}}{1-c_{13}^2} - \lambda_4 \cdot \frac{\sigma_{34} - c_{13}\sigma_{14}}{1-c_{13}^2}; \; \lambda_1 = \frac{\sigma_{15} - c_{13}\sigma_{35}}{1-c_{13}^2} - \lambda_4 \cdot \frac{\sigma_{14} - c_{13}\sigma_{34}}{1-c_{13}^2}$$

$$\begin{aligned} \lambda_2 &= \sigma_{25} - \frac{\sigma_{23}}{1-c_{13}^2}\cdot(\sigma_{35} - c_{13}\sigma_{15}) - \lambda_4 \cdot \left(c_{24} - \frac{\sigma_{23}}{1-c_{13}^2}\cdot(\sigma_{34} - c_{13}\sigma_{14})\right) \\ &= \sigma_{25} - \lambda_3\sigma_{23} - \lambda_4 c_{24} \end{aligned}$$

RELATIONS TO INSTRUMENTAL VARIABLES (IV)

Current conditions of identifiability, both algebraic and graphical, are based primarily on IVs (Bowden & Turkington, 1984), the graphical representation of which are described in (Pearl, 2000a, pp. 150–154). These conditions are applicable to models containing bow-patterns, as is illustrated in the following simple example:

$$W = e_1$$
$$X = aW + e_2$$
$$Y = cX + e_3$$

with $cov(e_1,e_2) = cov(e_1,e_3) = 0$. The potential correlation between e_2 and e_3 induces a bow-pattern in the corresponding diagram, as shown in Figure 5a. Nevertheless, the fact that $cov(W, e_3) = 0$ permits us to identify parameter c from Wright's equations:

$$\sigma_{WX} = a$$
$$\sigma_{WY} = ac$$

yielding

$$c = \sigma_{WY}/\sigma_{WX}$$

More generally, whenever a variable W can be found which is correlated with X but is uncorrelated with the error (e) in the equation

$$Y = cX + e$$

that variable is called an *instrument* for c, for it permits the identification of c, using Equation 5. If the equation for Y contains several explanatory variables, then W need be uncorrelated with e and with each of those explanatory variables (excluding X).

The graphical criterion for recognizing an instrument is based on replacing the condition $cov(W, e) = 0$ with its graphical analogue; namely, all paths between W and Y should be blocked in the subgraph G_c formed by deleting the link labeled by

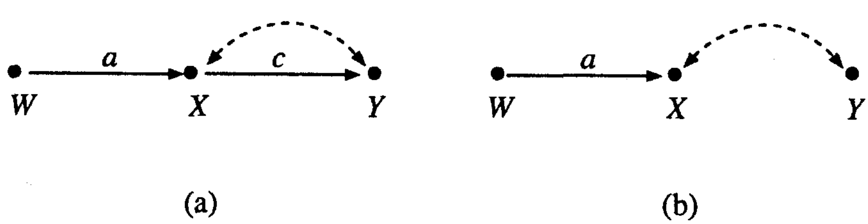


FIGURE 5 (a) A model identified by using W as instrumental variable for c. (b) The subgraph G_c used for classifying W as an instrumental variable: All paths between W and Y are blocked in G_c.

c (Pearl, 2000a, p. 150). For example, the subgraph G_β corresponding to Figure 5a is shown in Figure 5b; in this subgraph, all paths between W and Y are indeed blocked, since X is a collider.

An important generalization of the IV method is offered through the use of *conditional IVs*. A conditional IV is a variable W that may be correlated with the equation error-term but the two become uncorrelated when we condition on some set Z of observed variables (also called *covariates*). When such a set is found, parameter c is identified and (paralleling Equation 5) is given by

$$c = \frac{\sigma_{WY \cdot Z}}{\sigma_{WX \cdot Z}} \tag{6}$$

For example, there exists no variable that is uncorrelated with e_3 in Equation 1 (Figure 1b). However, variable Y_2 acts as a conditional instrumental variable (relative to β), since, conditional on Y_1, Y_2 is uncorrelated with e_3. In this case, substituting $c = \beta$, $W = Y_2$, $X = Y_2$, $Y = Y_3$, $Z = Y_1$ in Equation 6, we obtain $\beta = \sigma_{Y_2Y_3 \cdot Y_1} = \sigma_{23 \cdot 1}$ because $\sigma_{Y_2Y_2 \cdot Y_1} = \sigma_{22 \cdot 1} = 1$.

Clearly, searching a system of equations for a pair (W, Z) such that W is a conditional IV (conditioned on Z) may be a tedious task, because zero conditional covariances (or zero partial correlations) are not easily recognizable in a system of equations. Fortunately, a graphical criterion for recognizing such zero correlations is available, which reduces the search task substantially, and renders the discovery of conditional IVs manageable by simple inspection of the graph. The criterion is called *d-separation* (Pearl, 1986), and is defined in the following.

Definition 4 (*d*-Separation)

A set Z of nodes is said to *d-separate* node X from node Y in a graph if Z blocks every path between X and Y. A set Z of nodes is said to block a path p if either (a) p contains at least one arrow-emitting node that is in Z, or (b) p contains at least one collider that is outside S and has no descendant in Z.[6]

For example, the set $Z = \{Z_1\}$ blocks the path $X \leftarrow W \leftarrow Z_1 \rightarrow Z_2 \rightarrow Y$, because the arrow-emitting node Z_1 is in Z. However, the set $Z = \{Z_3\}$ does not block the path $X \leftarrow W \leftarrow Z_1 \rightarrow Z_3 \leftarrow Z_2 \rightarrow Y$, because none of the arrow-emitting nodes, W, Z_1, and Z_2, are in Z, and the collider Z_3 is not outside Z.

Because any *d*-separation condition in the graph implies $\sigma_{XY \cdot Z} = 0$ (Verma & Pearl, 1988; Pearl 2000a, p. 142), we can define *conditional instrumental variables* (Pearl, 2000b).

[6]The terms *arrow-emitting node* and *collider* are to be interpreted literally as illustrated by the examples given.

Lemma 3

Let c stand for the path coefficient assigned to the arrow $X \rightarrow Y$ in a causal graph G, and let Z be a set of variables. A variable W is an instrument relative to c, conditional on Z, if the following conditions hold:

1. Z consists of nondescendants of Y.
2. Z d-separates W from Y in the subgraph G_c formed by removing $X \rightarrow Y$ from G.
3. Z does not d-separate W from Y in G_c.

Moreover, if the preceding conditions hold, c is identified and is given by Equation 6.

To illustrate, consider the model of Figure 6a. Although there exists no d-separation condition in this graph, such condition does exist in the subgraph G_β formed by deleting the arrow from Y_2 to Y_3 (Figure 6b).

In this subgraph, Y_2 and Y_3 are d-separated by the set $Z = \{Y_1\}$. Therefore, Y_2 is an instrument for β, conditioned on $Z = \{Y_1\}$. By Lemma 3, this implies that β is identified and is given by $\beta = \sigma_{23.1}$.

However, the method of conditional instrumental variables has its limitations. Figure 4, for example, represents an identifiable model that has no instrumental variable for the parameters λ_3 and λ_4. This can be verified using Lemma 3 by noting in the subgraph $G_{\lambda 3}$ there exists no pair (W, Z) that d-separates Y_3 from Y_4. Thus, the method of conditional instrumental variables is not able to identify the bow-free model of Figure 4. Such status is easily ascertained by the criterion of Theorem 1, since this model is bow-free and any bow-free model can immediately be classified as identifiable. Note also that criteria based on instrumental variables ensure the identifiability of one parameter at a time, while the criterion established in this paper ensures the identifiability of the model as a whole.

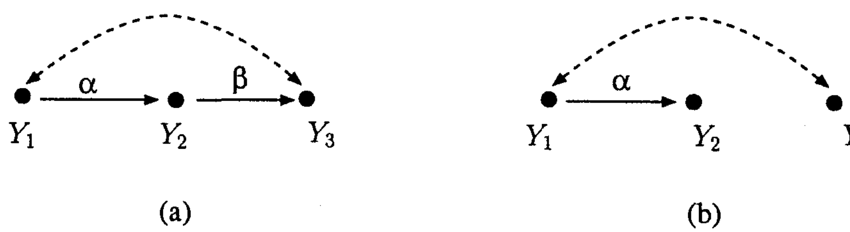


FIGURE 6 (a) A model illustrating the use of conditional instrumental variables: Y_2 is an instrument for β, conditional on Y_1. (b) The subgraph G_β used for classifying Y_2 as conditional instrument, per Lemma 3.

CONCLUSION AND DISCUSSION

We have shown that any linear, recursive bow-free model permits the identification of all its parameters. In other words, a linear recursive model is identifiable as long as the error associated with each variable is uncorrelated with the errors associated with the direct effects of that variable; other errors may be correlated. The proof of Theorem 1 further provides a systematic, recursive method of computing each structural parameter as a function of the observed covariance matrix. This result supplements our arsenal of identification conditions with a new criterion that is widely applicable and easily discernible by unaided investigators.

ACKNOWLEDGMENT

This research was supported in parts by grants from the National Science Foundation, the Office of Naval Research (Multidisciplinary University Research Initiative), the Air Force Office of Scientific Research, and the Conselho Nacional de Desenvolvimento Cientifico e Tecnologico fellowship proc. 200201/99–9.

REFERENCES

Bekker, P. A., Merckens, A., & Wansbeek, T. J. (1994). *Identification, equivalent models, and computer algebra*. Boston: Academic.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bowden, R. J., & Turkington, D. A. (1984). *Instrumental variables*. Cambridge, England: Cambridge University Press.

Duncan, O. D. (1975). *Introduction to structural equation models*. New York: Academic.

Fisher, F. M. (1966). *The identification problem in econometrics*. New York: McGraw-Hill.

Geraci, V. (1983). Errors in variables and the individual structure model. *International Economic Review, 24*(1), 217.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.

Kang, K. M., & Seneta, E. (1980). Path analysis: An exposition. In P. R. Krishnaiah (Ed.), *Developments in statistics* (Vol. 3, pp. 217–246). New York: Academic.

Kenny, D. A., Kashy, D. A., & Bolger. N. (1998). Data analysis in social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). Boston: McGraw-Hill.

Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.

McDonald, R. P. (1997). Haldane's lungs: A case study in path analysis. *Multivariate Behavioral Research, 32,* 1–38.

Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Annals of Statistics, 1,* 763–765.

Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence, 29,* 241–288.

Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods and Research, 27,* 226–284.

Pearl, J. (2000a). *Causality: Models, reasoning, and inference.* New York: Cambridge University Press.

Pearl, J. (2000b, August). *Parameter identification: A new perspective* (Tech. Rep. No. R-276). Los Angeles: University of California, Los Angeles.

Rigdon, E. E. (1995). A necessary and sufficient identification rule for structural models estimated in practice. *Multivariate Behavioral Research, 30,* 359–383.

Spirtes, P., Richardson, T., Meek, C., Scheines, R., & Glymour, C. (1998). Using path diagrams as a structural equation modelling tool. *Sociological Methods and Research, 27,* 182–225.

Verma, T., & Pearl, J. (1988). Causal networks: Semantics and expressiveness. In R. D. Shachter, T. S. Levitt, L. N. Kanal, & J. F. Lemmer (Eds.), *Proceedings of the 4th workshop on uncertainty in artificial intelligence* (pp. 352–359). Mountain View, CA: Elsevier.

Wright, S. (1960). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics, 16,* 189–202.