Primary submission: 31.10.2017 | Final acceptance: 07.05.2018

Implementation-Neutral Causation in Structural Models

Stephen F. LeRoy[1]

ABSTRACT Analysts associated with the Cowles Commission attached great importance to the distinction between structural and reduced-form models: in their view structural models, but not reduced-form models, allow the analysis of causal relations. They did not present clear justification for this view. Here we show that this insight is correct, and make the demonstration of it precise. Causal relations are shown to depend on parameter restrictions that are explicit in the structural form, but not in the reduced form when the coefficients are interpreted as unrestricted constants. The requisite parameter restrictions are those associated with implementation-neutral causation. A graphical procedure is outlined that identifies causal orderings and also the ordering based on implementation-neutral causation. The same procedure applied to reduced form models produces the implementation-neutral causal ordering only if the parameter restrictions are explicitly incorporated in the reduced form. The analysis is applied in investigating the validity of the causal Markov condition.

KEY WORDS: causation, correlation, regression, Cowles, implementation neutrality, external variables, internal variables

JEL Classification: B16, B41, C18, C40

[1] University of California, Santa Barbara , USA

Introduction

In the early days of macroeconomics much was made of the distinction between structural and reduced-form models, particularly by economists associated with the Cowles Commission at the University of Chicago and Yale University. Most simply, a linear structural model can be written as

$$Ay = Bx, \tag{1}$$

where y denotes the internal variables of the model (those determined by the model) and x denotes its external variables (those taken as given). Both x and y are vectors. $A = \{\alpha_{ij}\}$ and $B = \{\beta_{ik}\}$ are matrices of constants.[1] A is square and nonsingular, and is normalized by setting the elements of the main diagonal equal to one. Attention is restricted to linear models in this paper.[2]

Correspondence concerning this article should be addressed to: **Stephen F. LeRoy**, University of California, Santa Barbara, Santa Barbara, CA 93106, USA. Phone: (805) 689-2344. E-mail: leroy@ucsb.edu

1. Structural Forms and Reduced Forms

The Cowles economists distinguished the structural form of a model from its solution form,

$$y = A^{-1}Bx \equiv Gx, \tag{2}$$

where $G = \{\gamma_{ik}\}$. Eq. (2) is usually called the reduced form.[3] The structural form was viewed as conveying valuable information not contained in the reduced

form. It is difficult to extract from the Cowles economists' discussions exactly how this information is connected with causation (it was usually associated with identification), and why it disappears in going from the structural form to the reduced form (again, see the discussion in (Hurwicz, 1962)), although the argument there is not easy to follow). In this connection a recurrent theme has been that the structural form coefficients can be used to analyze interventions, and therefore locate causal orderings, whereas the reduced-form coefficients cannot be used in this way (see Section 3). There remains the question of why this is so.

This paper shows that a version of the Cowles argument is correct. It discusses a definition of causation that is based on the reduced form, but is most easily implemented using the structural form. We characterize the precise nature of the information that is lost in passing from the structural to the reduced form.

2. Equality and Causation

Many contemporary applications of structural models, particularly those directed toward graphical analysis of causation, use an alternative specification of structural models, written as

$$y = Ay + Bx. \tag{3}$$

Here A has zeros on the main diagonal. In (3) the symbol = denotes causation, with the right-hand side variables of each equation interpreted as directly causing the left-hand side variable. Thus = is an assignment operator, as in computer languages. This definition appears to allow each of two internal variables to cause the other. Some analysts have accepted this implication (Heckman, 2005), but others take the view that simultaneously-determined variables should be distinguished from causally ordered variables.

There is a problem with interpreting = as an assignment operator in models incorporating simultaneity. Application of the assignment operator to the right-hand side of the equation for y_i may require y_j as an input, and also vice-versa. Therefore these equations cannot be solved by application of the assignment operator. It follows that = is interpretable as the assignment operator only in fully recursive models. One would prefer to have a treatment of models in which there may be recursive blocks, but equations within such blocks are simultaneous.

Under the alternative interpretation each equation in (3) has a distinct identity: the variables that are direct causes of y_i are all located on the right-hand side of the i-th equation. In the philosophy literature this property is known as "modularity". In the formulation (1), in contrast, the characterization of = as a reflexive, symmetric and transitive operator implies that it is arbitrary which variable appears on the left-hand side of an equation. Thus we do not have modularity: the equations are best thought of as defining a single map from an m-dimensional space of external variables to an n-dimensional space of internal variables. With = interpreted as a reflexive, symmetric and transitive operator, there is no direct connection with causation.

The alteration in the meaning of = from its mathematical definition to its interpretation as representing causation has led some writers to express the view that graphical depictions of causal models, which incorporate the altered meaning of =, are fundamentally different from their algebraic counterparts (Elwert, 2013). Below we will conclude that, contrary to this, there is no reason to avoid using = with its usual mathematical meaning in analyzing causation, and this is so in both equation-system-based and graph-based discussions. This is a major attraction: economic models are derived from primitives by using mathematical calculations in which = is interpreted as a reflexive, symmetric and transitive operator, as opposed to an assignment operator. Proposing to change the interpretation of = upon termination of such derivations creates more problems than it solves. With = preserving its mathematical interpretion in the analysis of causation these problems do not arise. Thus structural models may or may not contain simultaneous blocks, consistent with their having a well-defined causal structure.

3. Causation Based on *"Ceteris Paribus"*

Angrist and Pischke (2015) is one of the few recent sources in the economics literature that discusses causation explicitly and clearly (but not, in our view, correctly). Their account outlines a treatment of causation that is widespread, if not universal, in contemporary economics. If y_j appears on the right-hand side of the structural equation determining y_i, then

 DOI: 10.5709/ce.1897-9254.276

y_j is defined to cause y_i "ceteris paribus". Here "ceteris paribus" means that other variables in the equation determining y_i, which may include both internal and external variables, are held constant. The i, j element of A is interpreted as giving a quantitative measure of the causal dependence of y_i on y_j, ceteris paribus. The intervention on y_j is not connected with the external variables that, according to the model, determine y_j.

The ceteris paribus definition of causation relies on the problematic characterization of equality as an asymmetric relation, as discussed in the preceding section. Interpreting the equality symbol instead as having its usual mathematical meaning, as recommended here, implies that a definition of causation based on the "ceteris paribus" condition is inadmissible inasmuch as it treats the left-hand side variable differently from the right-hand side internal variables.

Another problem is that analyzing causation using the ceteris paribus condition amounts to respecifying the model, so that the causal analysis is conducted using a model different from that actually proposed. An example will make this clear. Consider the recursive model

$$y_1 = \beta_{11}x_1 + \beta_{12}x_2 \tag{4}$$

$$y_2 = \alpha_{21}y_1 + \beta_{23}x_3 \tag{5}$$

$$y_3 = \alpha_{31}y_1 + \alpha_{32}y_2 + \beta_{34}x_4. \tag{6}$$

On the received account of causation, this model implies that y_2 causes y_3, with causal constant α_{32}, ceteris paribus. Here ceteris paribus means that y_1 and x_4 are held constant.

The operation of substituting a constant for the internal variable y_1 is inconsistent with the presence of eq. (4) in the model, since that equation states that y_1 instead depends on the external variables x_1 and x_2. We are led to accommodate the ceteris paribus condition by deleting eq. (4) from the model. Further, the exercise involves treating y_2 as an external variable, so we also delete from the model eq. (5), which characterized y_2 as an internal variable. We are left with a model consisting of the single equation

$$y_3 = \alpha_{32}\hat{y}_2 + \beta_{34}x_4, \tag{7}$$

where $\hat{y}_2$ denotes the variable y_2 now redefined to be an external variable.

In the model (7) there is no doubt that $\hat{y}_2$ causes y_3, and that the constant associated with this causation is α_{32}, as asserted in the received account of causation. However, the model from which this conclusion is drawn is completely different from the original model—eq. (4)-(6): the model as altered has different internal variables, different external variables and different equations. Rather than determining causation in the model as originally specified, applying the ceteris paribus condition in this way amounts to altering the model so as to create a setting in which causation has a clear meaning and is unambiguously associated with a constant in the model. Doing so does not constitute an analysis of causation in the original model.

There is another way to make essentially the same point. An intervention on y_2 is reducible to an intervention on x_1, x_2 or x_3. By assumption these are external, and therefore are not linked by functional equations. But holding constant y_1 effectively converts it into an external variable, thereby necessarily inducing a functional relation between x_1 and x_2. One of these (it is not clear which) becomes an internal variable. Again, the conclusion is that holding constant an internal variable constitutes an alteration of the model, and therefore is inadmissible in defining causal relations.

Properly viewed, the statement that one internal variable causes another "ceteris paribus" consists of the assertion that external variables that are not determinants of the cause variable, but not internal variables or external variables that are determinants of the cause variable, are held constant. In the remainder of this paper the term "causation" is always taken to mean causation that is ceteris paribus in this sense, so the "ceteris paribus" proviso can be omitted.

4. Interventions

We discuss our preferred treatment of causation in the remainder of this paper.

In the Cowles usage an intervention consists of a modification of the structural equations intended to allow the analyst to determine what would happen under a given hypothetical change in the environment (Haavelmo, 1943; Heckman & Pinto, 2013). Using a model in this way to analyze causation involves alter-

ing the setting, with the alteration depending on the causal question that is being asked. This practice was criticized in the preceding section.

The insistence of the Cowles economists on representing interventions as modifications of structural equations led them away from an alternative much simpler formalization of interventions using elements of the model that are already available: external variables. Representing interventions as hypothetical alterations of external variables means that no change in the model is involved in analyzing interventions. There is no loss of generality in requiring that interventions be modeled as alterations of external variables since any conceivable intervention can be accommodated by inclusion of external "shift variables" in the model.

Let us then initially set the external variables to preassigned values. The solution to the model under these values is termed the baseline. Then generate an intervention by changing the assumed value of one or more of the external variables and recompute the solution. One then determines the effect of the intervention by comparing the values taken on by the internal variables under the intervention with those under the baseline specification.

By designating a coefficient as an external variable rather than a constant the analyst is allowing for interventions on that variable. If the coefficients are variables the model is bilinear, not linear. These specifications are different. Treating constants as if they were variables, or converting constants to variables, is methodologically questionable at best: it makes no sense to alter a model in using it to analyze the effect of an intervention that is inherently ambiguous in the original formulation of the model. In an equation characterized as linear the coefficients are interpreted as constants. Labeling the coefficient a constant implies that interventions on that constant are ruled out: we do not ask mathematicians what would happen if π were equal to a number other than 3.1416, and economists should not be asking the analogous question about the constants of their models.[4]

The requirement that analysts explicitly distinguish constants from external variables and treat each consistently, even in analyzing interventions, enforces clarity about which contemplated interventions the analyst views as admissible and which are excluded from consideration. Here we part company from the Cowles economists who, as noted, were often unclear about this distinction.[5]

The same analysis applies to the frequently-encountered practice of analyzing interventions involving internal variables by relabeling cause variables as external variables and deleting from the model the equations that determine them (here we leave aside the fact that the equations to be deleted can be uniquely ascertained only in fully recursive models). Such a procedure constitutes a substantive modification of the model even in fully recursive models, as in (4)-(6). The treatment to be specified, in contrast, will provide a way to analyze causal relations among variables in a model without at the same time changing the model.

In forecasting exercises the general practice is to specify probability distributions for external variables and then derive the distributions of internal variables by applying the reduced-form equations. Analyzing interventions on such models, in contrast, involves specifying particular realizations of the external variables, as noted above. Contrary to some discussions, there is no contradiction between assigning probability distributions to external variables in using a model to generate forecasts and setting the realizations of these variables to determine effects of interventions. In modeling the price of some crop an analyst could specify that the harvest depends on the weather, and then produce a forecast by assuming a probability distribution for weather-related external variables. Equally, one could analyze what the crop would be if the weather were good. The former exercise is a forecast, while the latter constitutes analysis of an intervention. The same model can be used in either application.

5. Causation

Causal relations can be modeled in terms of interventions. One can determine for each internal variable the set of external variables each of which affects the value of that internal variable. The elements of these *external sets* are causes of the internal variable: if $\mathcal{E}(y_i)$ is the external set for y_i, then $x_j \in \mathcal{E}(y_i)$ causes y_i, written $x_j \rightarrow y_i$. The effect on internal variables of interventions on external variables can be calculated by applying the reduced form to the difference between the baseline values of external variables and their values under intervention: $\gamma_{ij}\Delta x_j$.

 DOI: 10.5709/ce.1897-9254.276

Causal relations may also be defined in which the cause variable as well as the effect variable is internal. Then the intervention consists of any of the possible changes in the values of the variables in the external set of the cause variable that lead to a given change in the value of the cause variable. If $\mathcal{E}(y_j)$ is a proper subset of $\mathcal{E}(y_i)$ we will say that y_j *causes* y_i, written $y_j \rightarrow y_i$ (here we follow the lead of Simon, 1953). The condition means that an intervention on any element of $\mathcal{E}(y_j)$ induces an alteration in y_i, and also that there exists some element of $\mathcal{E}(y_i)$ that is not in $\mathcal{E}(y_j)$, so as to assure the asymmetry of causation. The requirement of a proper subset rather than a subset allows us to distinguish between variables that are causally ordered and those that are simultaneously determined (in which case the variables have the same external sets). The contribution here is to provide a graphical counterpart to Simon's analysis and, starting in Section 8, to extend the analysis to a more restricted characterization of causation.

Note that analyzing causation so defined does not involve alteration of the model.

The proper subset condition implies that no internal variable has an external set consisting of a singleton. It would be counterintuitive at best to have a scale multiple of an external variable labeled an internal variable: doing so involves treating differently two variables that are not substantively different in terms of the structure of the model. Finally, allowing singleton external sets would seem to imply that external variables always cause themselves. This specification, besides conflicting with the definition of an external variable, would be odd at best.

Ruling out singleton external sets is more a notational convention than a substantive restriction; to ensure exclusion of internal variables with singleton external sets one has only to solve them out of the model by replacing the internal variable with the external variable and adjusting equation coefficients appropriately to allow for the scale factor.

6. Causal Graphs

As we have seen, causal orderings based on comparisons of external sets are derived by comparing all the pairs $\{x_j, y_i\}$ and $\{y_j, y_i\}$. Pairs $\{x_i, x_j\}$ are not connected because, being external, they are not causally related. If $x_j \in \mathcal{E}(y_i)$ we have $x_j \rightarrow y_i$. If $\mathcal{E}(y_j)$ is a proper subset of $\mathcal{E}(y_i)$ we have $y_j \rightarrow y_i$.

Given availability of numerical values for the elements of A and B one determines the causal ordering by solving for the reduced form, and then determining the existence or nonexistence of causation between any two variables by checking whether the external set of one is a proper subset of that of the other. This is directly inferred from the location of zeros in the reduced form. If one knows the location of zeros in A and B but does not have numerical values for parameters, as is the case in theoretical analysis, this method requires using symbolic matrix inversion software to determine the location of zeros in the reduced form. This may not always be convenient. An alternative procedure would be to assign arbitrary values to all structural coefficients, and then compute the reduced-form coefficients. Barring an unlucky choice of numerical values, this procedure would produce zeros in the same locations as the symbolic calculation would. The easiest approach, at least with simple models, involves working directly with the structural model instead of the reduced form, and using graphical methods.

The problem consists of deriving from the equations of the model a graph such that we have $x_j \rightarrow y_i$ and $y_j \rightarrow y_i$ if and only if there exists a path from x_j to y_i or y_j to y_i involving an ordered n-tuple of internal variables such that each member has an external set that is strictly greater than that of its predecessor. In this section we outline the construction of such a graph.

In advocating the use of graphical methods in analyzing causation we follow the mainstream in causal analysis, notably Pearl (2000). However, our use of graphical methods differs from that found in the mainstream tradition. In the received analysis the causal graph is taken directly from the structural model: the variables on the right-hand side of each equation are identified as direct causes of the left-hand side variable. We took issue with this specification in Section 2.

A preliminary—and trivial—first step consists of construction of the *structural graph*. This consists of a graph in which each variable is connected by an undirected edge to each other variable that appears in the same structural equation. The use of undirected edges reflects the interpretation of = as a reflexive, symmetric and transitive operator, as discussed in Section 2. Each edge can be labeled with the corresponding coefficient

in the mathematical form of the model, although these labels may be omitted depending on the application. Hereafter we will use the structural graphs of a model in place of its mathematical form, the two being interchangeable. Doing so facilitates comparison of structural and causal graphs.

A *causal graph* consists of a graph that connects with arrows variables that are causally linked, where the causal arrow here has the same meaning as in Section 5 (except that, as noted below, in causal graphs it is convenient to omit causal arrows when they are implied by the transitivity of causation). This involves connecting pairs of variables such that the effect variable has an external set that is strictly larger than that of the cause variable, except when there exists an intermediate variable that is causally related to both. In the presence of such a variable the two original variables are directly connected by an arrow only if there also exists an undirected edge connecting them in the structural graph. Existence of such an edge reflects the fact that the two variables are causally linked both directly and indirectly. We want the graph to show both direct and indirect causation when both are present; this is possible only if we delete all causal arrows that are implied by the transitivity of causation.

Some models have simultaneous blocks, meaning that nonsingleton sets of internal variables have the same external sets. Such variables are not connected by arrows. Instead, each member of a simultaneous block is connected with an arrow to each variable outside the block that is connected by a directed edge to any member of the block. In this case causal arrows may represent the connection between non-adjacent variables. Example 2 below illustrates this construction.

The procedure just outlined generates a graph in which each internal variable is connected to its ancestors by paths of incoming arrows and to its descendants by paths of outgoing arrows. Parents and children are special cases of ancestors and descendants where the connection is achieved via a single arrow, so that causation is direct. Causation is indirect when more than one arrow is involved in the path from cause variable to effect variable.

The causal graph allows an easy representation of the reduced-form coefficients. With each causal path is associated a path coefficient consisting of the product of the edge coefficients associated with the arrows that generate the path. Each internal variable is connected to each of the variables in its external set by one (or more) causal path(s). If there exists only one path connecting the two variables the reduced-form coefficient of that internal variable with respect to each external variable in its external set equals the path coefficient for that path. If there exist more than one path connecting the two variables, the reduced-form coefficient equals the sum of the path coefficients (see Examples 3 and 4).

Note that this characterization of causal coefficients applies without qualification only when the cause variable is external, as with the reduced form. The corresponding characterization when the cause variable is internal is found in Sections 8, 9 and 10.

7. Examples

The algorithm presented here is illustrated using examples. In each case the model is defined using a structural graph. The associated causal graph is presented when the structural and causal graphs differ. As will be seen, sometimes the two graphs coincide (with undirected edges replaced by arrows; hereafter this proviso is omitted), and sometimes not.

Example 1

The structural graph shown in Figure 1 depicts the simplest model in which the internal variables are causally ordered. It is discussed in Section 9. The associated causal graph is identical to the structural graph.

Example 2

The standard economist's supply-demand model, in which each of two equations includes price, quantity and one external variable, is the simplest model that includes simultaneous determination of a block of internal variables, here consisting of y_1 and y_2. The internal variables, price and quantity in the supply-demand example, are not causally ordered. Its structural graph is shown as Figure 2(a). The causal graph, shown as Figure 2(b), consists of arrows linking each of the internal variables to each of the external variables, with the internal variables not causally connected.

Example 3

In Figure 3 the variables y_1 and y_2 have external sets neither of which is a subset of the other, and y_3 has

 DOI: 10.5709/ce.1897-9254.276

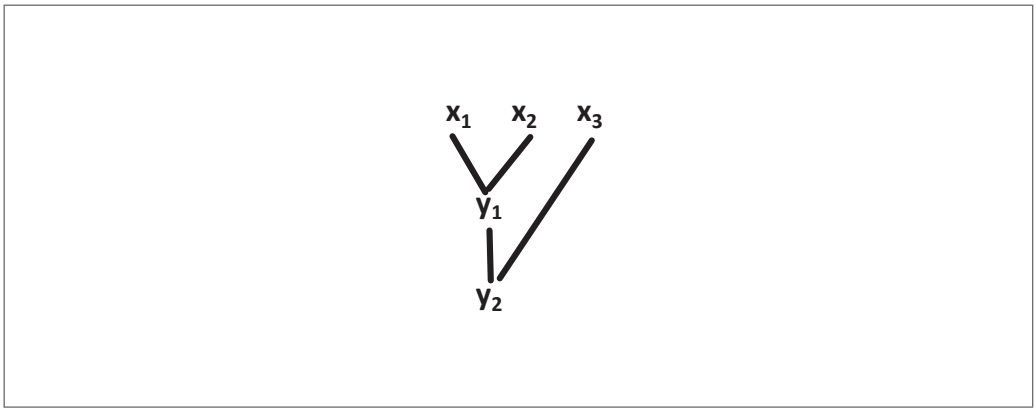

Figure 1.

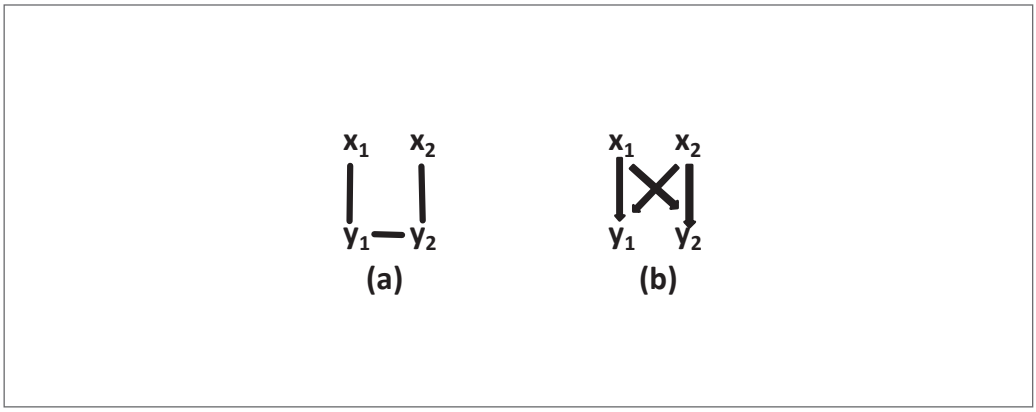

Figure 2.

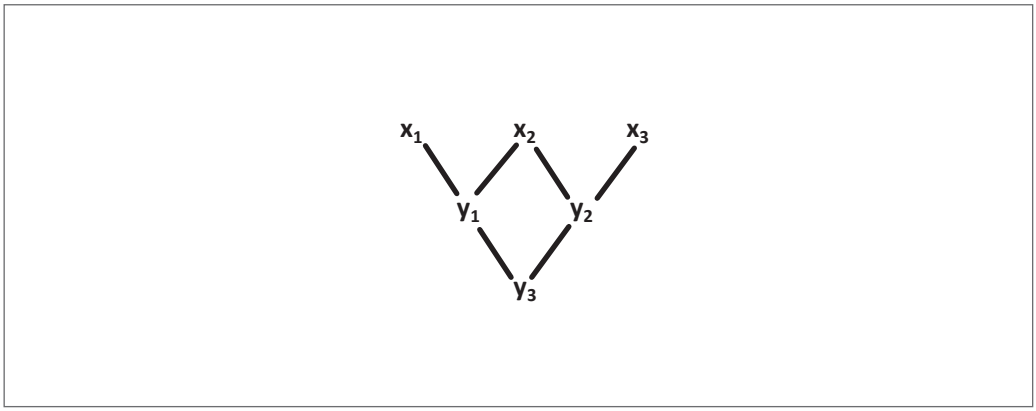

Figure 3.

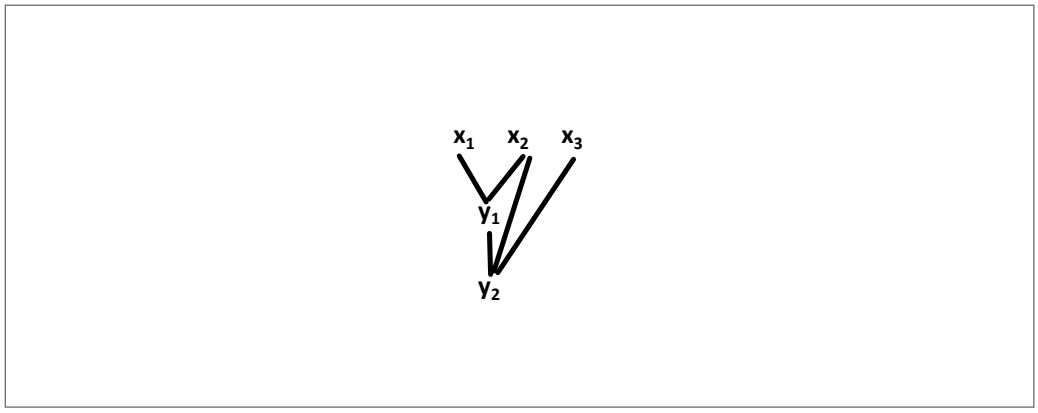

Figure 4.

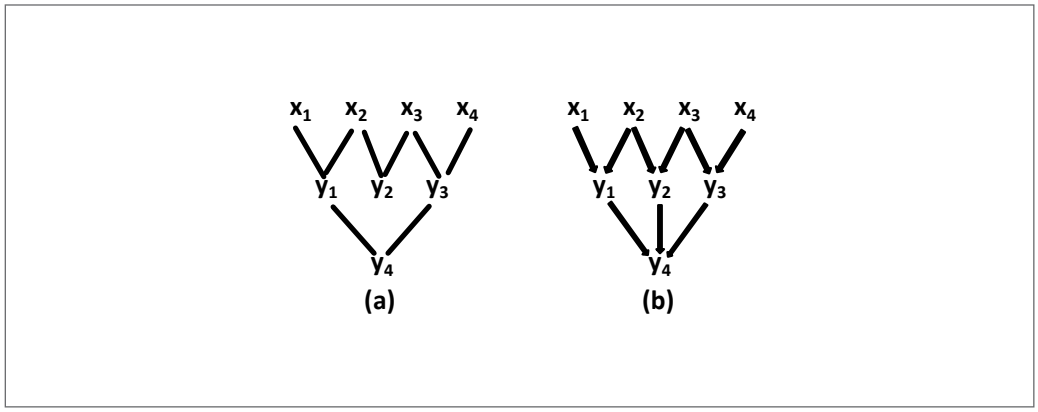

Figure 5.

an external set that properly contains the external sets of each of y_1 and y_2. Therefore the causal form of the model coincides with the structural form: y_1 and y_2 are not causally related, but each causes y_3. The external variable x_2 affects y_3 via two indirect paths, so the reduced-form coefficient of y_3 with respect to x_2 is $\alpha_{31}\beta_{12} + \alpha_{32}\beta_{22}$.

Example 4

The structural model shown in Figure 4 differs from that in Figure 1 due to inclusion of an edge connecting x_2 and y_2. Here x_2 and y_2 are connected both directly and indirectly. The causal graph in Figure 4 coincides with its structural graph.

Example 5

Here y_2 causes y_4 in Figure 5(b) despite the absence of an edge connecting them in Figure 5(a). This is so because an intervention on y_2 is necessarily attributable to an intervention on x_2 or x_3, and these cause a change in y_4. Note that, from comparison of Figures 2 and 5, presence of an edge connecting two variables in the structural graph is neither necessary nor sufficient for presence of an arrow connecting them in the causal graph.

8. Implementation-Neutral Causation

For many purposes the notion of causation just outlined is too weak. This is so because even if y_j causes

 DOI: 10.5709/ce.1897-9254.276

y_i different interventions consistent with a given Δy_j can induce different y_i. For example, suppose that we consider an intervention Δy_1 on y_1 in some model. Any intervention on an internal variable is viewed as generated by an underlying intervention on the variables in its external set. In the model of Example 4 the intervention Δy_1 could have been caused by an intervention of $\Delta y_1 / \beta_{11}$ on x_1 or $\Delta y_1 / \beta_{12}$ on x_2 (or, of course, a linear combination of these). There results $\Delta y_2 = \alpha_{21} \Delta y_1$ in the first case and $(\alpha_{21} + \beta_{22} / \beta_{12}) \Delta y_1$ in the second. The question "What is the effect of y_1 on y_2?" does not specify which is the case, leading to the conclusion that the magnitude of the causal effect of y_1 on y_2 is not well defined.

One could object against this line that in the model of Example 3 Δy_1 results unambiguously in an effect $\alpha_{31} \Delta y_1$ on y_3 if y_2 is held constant. We argued in Section 3 that holding constant an internal variable in this way constitutes an alteration of the model (by inducing a functional relation between variables specified as external; in this case x_2 and x_3). Avoiding altering the model leaves us with the conclusion that the effect of y_1 on y_2 in the model of Example 3 as specified is in fact inherently ambiguous.

In other cases this ambiguity does not occur. If in addition to $y_j \to y_i$ we have a model in which all the interventions that lead to a given value of Δy_j map onto the same value of y_i, the effect of Δy_j on y_i does not depend on how Δy_j is implemented (that is, which element(s) of $\mathcal{E}(y_j)$ is (are) intervened upon). In that case causation is *implementation neutral*. The causal relation between y_1 and y_2 in Figure 1 is implementation neutral: the effect on y_2 of an intervention of $\Delta y_1 / \beta_{11}$ on x_1 (equal to $\alpha_{21} \Delta y_1$) is equal to that of an intervention of $\Delta y_1 / \beta_{12}$ on x_2. We refer to the causal relation so defined as *IN-causation*. If y_j causes y_i and the causation is implementation neutral we will write $y_j \Rightarrow y_i$. For each internal variable y_i, each x_j in $\mathcal{E}(y_i)$ IN-causes y_i; for arbitrary internal variable y_j, y_j may or may not IN-cause y_i, even if $y_j \to y_i$, as we have just seen.

If x_j or y_j IN-causes y_i by definition there exists a constant that measures the effect of x_j or y_j on y_i. This constant is the same for all possible underlying interventions. That coefficient may or may not coincide with the parameter α_{ij} in the structural model (it does so when only one path connects y_j and y_i, but not otherwise). Note that, in the discussion in Section 3 of the ceteris paribus condition, in the recursive model (4)-(6) y_2 does not IN-cause y_3: if the intervention inducing Δy_2 is on x_3, the effect on y_3 is different from that occurring if the intervention is on x_1 or x_2. Therefore the constant α_{32} cannot be interpreted causally. In contrast, y_1 does IN-cause both y_2 and y_3, so α_{21} and α_{31} can be interpreted causally.

For a related analysis under the rubric of "spurious correlation" (see Simon, 1954). By "spurious correlation" Simon meant correlation where there is no causation. Here Simon can be interpreted as anticipating the idea of implementation neutrality, although his analysis differs from that found here.[6]

The IN-causal ordering consists of all the pairs x_j, y_i and y_j, y_i such that $x_j \Rightarrow y_i$ and $y_j \Rightarrow y_i$. IN-causation will be our primary notion of causation: if $y_j \to y_i$ but not $y_j \Rightarrow y_i$ we do not have enough information about the intervention to characterize its effect on y_i quantitatively. IN-causation is discussed more fully in LeRoy (2016), and is applied to treatment evaluation in LeRoy (2018).

9. IN-Causation in Reduced-Form Models

We review the criteria for IN-causation in the simplest nontrivial reduced-form model discussed above in which two internal variables are causally ordered:

$$y_1 = \gamma_{11} x_1 + \gamma_{12} x_2 \tag{8}$$

$$y_2 = \gamma_{21} x_1 + \gamma_{22} x_2 + \gamma_{23} x_3. \tag{9}$$

In (8)-(9) the term in the 1,3 position in G is zero, with other terms nonzero. Hereafter we will refer to equation systems like (8)-(9), where the nonzero values of G are interpreted as unrestricted constants, as "generic reduced forms".[7] In (8)-(9) we have $y_1 \to y_2$, but not necessarily $y_1 \Rightarrow y_2$.

As discussed in LeRoy (2016), under the parameter restriction

$$\frac{\gamma_{21}}{\gamma_{11}} = \frac{\gamma_{22}}{\gamma_{12}} \tag{10}$$

we can define the constant α_{21} by

$$\frac{\gamma_{21}}{\gamma_{11}} = \frac{\gamma_{22}}{\gamma_{12}} \equiv \alpha_{21} \tag{11}$$

and write the model in structural form as

$$y_1 = \beta_{11}x_1 + \beta_{12}x_2 \tag{12}$$

$$y_2 = \alpha_{21}y_1 + \beta_{23}x_3. \tag{13}$$

In (12)-(13) we have $y_1 \Rightarrow y_2$. The same is true in the reduced-form version of the model, eq. (8)-(9), under the parameter restriction (10), but not otherwise.

Inspection of the generic reduced form allows determination of whether two variables are causally related, in the sense of $\rightarrow$ as defined above. This is so because the generic reduced form contains the information about whether the external set of y_1 is a proper subset of that of y_2, in which case we have $y_1 \rightarrow y_2$, or not. Whether the parameter restrictions required for IN-causation, (10) in our example, are satisfied cannot be determined from the generic reduced form (8)-(9) without restrictions on the parameters.

If the generic reduced form can be rewritten as

$$y_1 = \gamma_{11}x_1 + \gamma_{12}x_2 \tag{14}$$

$$y_2 = \alpha_{21}\gamma_{11}x_1 + \alpha_{21}\gamma_{12}x_2 + \gamma_{23}x_3 \tag{15}$$

for some α_{21}, it satisfies the parameter restrictions (11) by construction, implying that the derivation of the structural form (12)-(13) is immediate. We will use the term "restricted reduced form" to refer to the version of the reduced form that incorporates the reduced-form restrictions implied by some structural model, as in (14)-(15). Thus structural models, or equivalently restricted reduced forms, contain information about both causation and IN-causation, as distinguished from unrestricted reduced forms, which do not contain information about IN-causation.[8]

The possibility of encoding structural information in reduced forms has relevance for the ongoing debate between statisticians, economists and members of other disciplines about the meaning of structural equations. Statisticians and econometricians (see Haavelmo, 1943; Pearl, 2015; Wermuth, 1992) have taken the view that the coefficients of structural models have no clear meaning because they are not connected to the probability distribution of internal variables. This statement is correct when the probability distribution of internal variables is viewed as generated by applying the generic reduced form to the external variables, the probability distribution of which is assumed. However, it is incorrect as applied to the restricted reduced form: as the above example shows, structural parameters like α_{21} in fact appear in restricted reduced forms, and therefore can be viewed as figuring in the link between assumed distributions of external variables and the implied distributions of internal variables.

10. IN-Causation in Structural Models

As observed in Section 8, $y_j \rightarrow y_i$ does not imply $y_j \Rightarrow y_i$. In graphical terms this is so because there may exist paths communicating causation that connect variables in the external set of y_j with y_i but do not pass through y_j. Existence of such paths implies that the effect on y_i of an intervention resulting in Δy_j differs under different such interventions. When there are no such paths we have $y_j \Rightarrow y_i$.[9]

A path connecting two variables is a *directed path* when all the arrows along the path point in the same direction. If that condition fails causation along that path is blocked, a result that has figured prominently in earlier causation discussions. This observation is relevant in considering the converse of the proposition stated above (that if all paths from elements of $\mathcal{E}(y_j)$ to y_i pass through y_j then we have $y_j \Rightarrow y_i$). The converse would be that if $y_j \Rightarrow y_i$ then all paths from $x_k \in \mathcal{E}(y_j)$ to y_i pass through y_j. This, however, is not true. Consistent with $y_j \Rightarrow y_i$ we may have paths connecting elements of $\mathcal{E}(y_j)$ and y_i that are defined in the causal form but are not defined in the structural form. Also, there may exist paths from $x_k \in \mathcal{E}(y_j)$ to y_i that are defined in the structural form but are not directed.

To see that such paths can coexist with IN-causation, consult Figure 6. We have that $y_1 \Rightarrow y_3$ despite the existence of two paths in the causal graph that connect x_2 to y_3 but do not pass through y_1: $\{x_2, y_2, {}_{3,} y_3\}$ and $\{x_2, y_2, y_3\}$. However, the first of these is not a directed path and the second is not the counterpart of a path in the structural graph, since there is no edge connecting y_2 and y_3.

11. The Causal Markov Condition

Up to now the analysis has been theoretical: we have separated the task of defining causation from that of

 DOI: 10.5709/ce.1897-9254.276

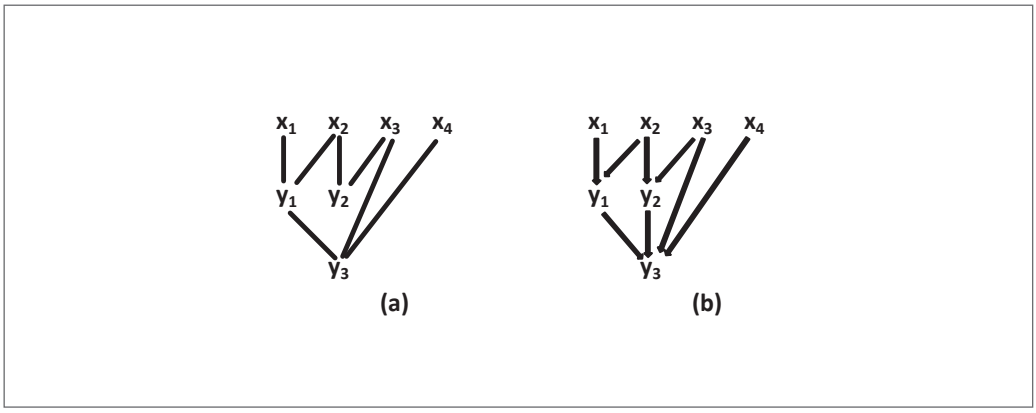

Figure 6.

testing causal models and estimating causal coefficients. Focusing on the former, as we have done, meant that there was no need to discuss probability distributions, observability of variables by the analyst or identification. Ultimately we do want to test causal models. An important tool that has been applied to this end is the causal Markov condition which, it is asserted, makes possible empirical testing of causal orderings and empirical estimation of causal parameters.

The causal Markov condition, as formulated by Spirtes, Glymour and Scheines (1993) for example, states that every variable of a model is probabilistically independent of all variables other than its descendants and parents, given its parents. The proposition that lack of correlation implies causation thus reverses the usual statement.

The status of the causal Markov condition is ambiguous. In places it is treated as an axiom separate from other assumptions specifying the structure of the model. In other places it is regarded as part of the definition of Bayesian networks, which usually involves sidestepping the question of whether a causal graph is a Bayesian network. Finally, it is sometimes treated as a substantive proposition that can be evaluated on philosophical grounds (see Hausman & Woodward, 1999).

The most obvious problem here is that, from elementary probability theory, two random variables are always independent conditional on one of them. It follows that we can certainly delete "and parents" from the definition of the causal Markov condition (this point was noted by Hausman and Woodward (1999)). A slightly less obvious point is that, because any variable can be written as a deterministic function of its parents (shocks, being random variables, are included in the set of parents of the variables they cause), any variable is independent of all variables, including its descendants, conditional on its parents. It follows that the causal Markov condition as just stated is valid, but trivially so.

These points, of course, depend on the definition adopted in this paper of parents as the set of all variables that directly cause the variable in question. In treatments of causation one often sees discussions that presume that error terms are not causal parents. However, no guidance is given as to the basis for distinguishing variables that are causal parents from those that cause a variable but are not counted among its parents. Variables characterized as errors are, of course, unobserved, but there is no apparent justification for denying their status as causal parents for this reason: the definition of causal orderings does not depend on which variables are observable. Hausman and Woodward (1999) explicitly posit existence of causal variables that are not included in the model under consideration and therefore do not qualify as parents. Presumably these appear as variables in some unspecified meta-model. It is not explained what purpose it serves to make this distinction, at least without connecting the discussion to the topic of model misspecification.

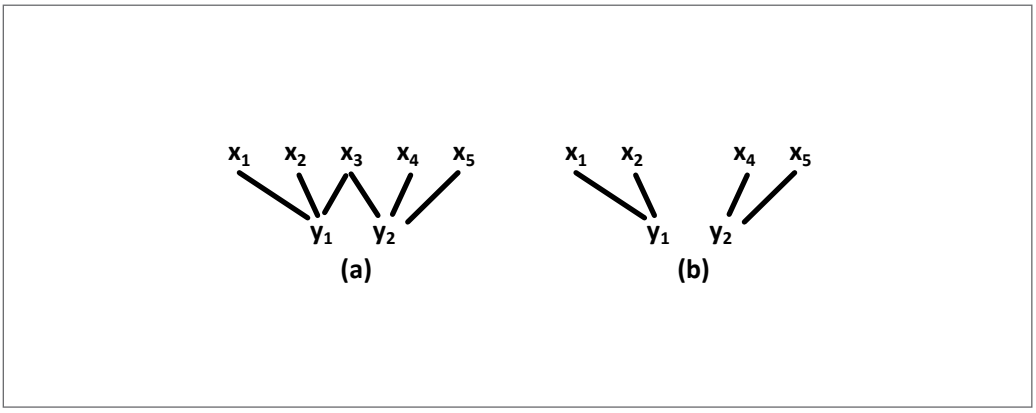

Figure 7.

Despite the arguments just made, there are several propositions similar to the causal Markov condition that are correct and nontrivial, and are easily derived in the framework set out here. These derivations all depend on the assumption that the external variables are probabilistically independent. If that condition fails the interpretation is that some correlations between the variables of a model are generated by the model's causal structure, while others are buried as unmodeled correlations among external variables. There is no empirical evidence that can determine which is which. Thus one cannot distinguish empirically between correlations generated as a consequence of the model's causal structure and those resulting from correlations among external variables.

We set forth two such propositions; no doubt there are others. The first proposition is that, assuming that the external variables are independent, if two variables have a single common ancestor they are correlated unconditionally, but are independent conditional on that ancestor. The reason is that when external variables are assumed to be independent the correlation between two internal variables results from overlap between their external sets. By assumption that overlap consists of a single external variable (if it were internal its parents would also be ancestors). Conditioning on the common ancestor effectively implies that the external sets of the two variables are disjoint, and therefore the variables are independent. If there are more than one common ancestor it is necessary to condition on all of them; conditioning on fewer is consistent with the two variables being conditionally correlated, as is easy to confirm by example.

In the model shown as Figure 7(a) the variables y_1 and y_2, having the common ancestor x_3, are correlated. Conditioning on x_3 effectively removes that variable from the model, and therefore removes also the edges connecting it with its children. The graph Figure 7(b) results. With the external sets of y_1 and y_2 now disjoint, those variables are independent.

The second proposition, unlike the first, involves implementation-neutral causation. It states that $y_1 \Rightarrow y_2$ implies that y_2 is independent of any ancestor of y_1 conditional on y_1. This conclusion is illustrated in Figure 1. Conditional on y_1, y_2 depends only on x_3. But the definition of IN-causation implies that x_3 is not an element of the external set of any ancestor of y_1. Therefore any variables that are ancestors of y_1, in this case x_1 and x_2, are independent of x_3, and therefore also of y_2 conditional on y_1. Note that this result, like that of the preceding paragraph, depends critically on the assumption that the external variables are independently distributed. Here if x_3 is correlated with x_1 or the proposition under discussion fails.

The proposition just stated has a partial converse: if $y_k \to y_j \to y_i$ and y_k is independent of y_i conditional on y_j, then we have $y_j \Rightarrow y_i$. The fact that we have $y_k \to y_j \to y_i$ implies that there exist paths connecting y_k and y_i. The fact that y_k and y_i are independent conditional on y_j means that all directed paths con-

 DOI: 10.5709/ce.1897-9254.276

necting y_k and y_i pass through y_j. This is the definition of IN-causation.

Existence of this theoretical result implies that, subject to maintained assumptions, IN-causation is testable. The availability of a partial converse suggests that in some settings the test may have high power.

12. Conclusion

In the introduction it was noted that the Cowles economists did not provide a clear statement of why structural models are better suited for causal analysis than reduced forms. We now have such a statement: generic reduced-form models incorporate information about the subset relations implying the causal ordering we have denoted by $\rightarrow$, but not about the parameter restrictions necessary for implementation neutral causation, denoted by $\Rightarrow$. Structural models and restricted reduced-form models contain both sources of information. Comparing results for structural and generic reduced-form models, we see that solving for the generic reduced form and using it to diagnose causation without incorporating the parameter restrictions involves a loss of information, just as the Cowles economists asserted.

It is worth noting that our justification for the Cowles analysis involves only causation. In contrast, most discussions connect the Cowles analysis primarily with identification. While there is no question of the importance of parameter restrictions for identification in structural models, here we have not introduced a distinction between observed and unobserved variables. This omission reflects the fact that causal orderings do not depend on which variables are observed. Questions relating to identification and estimation of causal parameters do require specification of which variables are observed, of course.

References

Angrist, J. D., & Pischke, J. S. (2015). *Mastering 'Metrics: The Path from Cause to Effect.* Priceton, NJ: Princeton University Press.

Elwert, F. (2013). Graphical causal models. In Morgan, S. L. (Ed.), Handbook of causal analysis for social research (pp. 245–273). Dordrecht: Springer.

Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica*, *11*(1), 1–12.

Hausman, D. M., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *The British Journal of the Philosophy of Science*, *50*(4), 521–583.

Heckman, J. (2005). The scientific model of causality. *Sociological Methodology*, *35*, 1–97.

Heckman, J., & Pinto, R. (2013, September 25). Causal analysis after Haavelmo (Working Paper No. 19453). Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w19453

Hurwicz, L. (1962). On the structural form of interdependent systems. In E. Nagel, P. Suppes, & A. Tarski (Eds.), Logic, methodology and the philosophy of science (pp. 232-239). Standford, CA: Stanford University Press.

LeRoy, S. F. (2016). Implementation neutrality and causation. *Economics and Philosophy*, *32,* 121–142.

LeRoy, S. F. (2018). Implementation neutrality and treatment evaluation. *Economics and Philosophy*, *34*(1), 45-52.

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge, UK: Cambridge University Press.

Pearl, J. (2015). Trygve Haavelmo and the emergence of causal calculus. *Econometric Theory*, *31,* 152–179.

Simon, H. A. (1953). Causal ordering and identifiability. In W. C. Hood, & T. C. Koopmans (Eds.), *Studies in Econometric Method, Cowles Commission for Research in Economics Monograph No. 14* (pp. 49-74). New York, NY: John Wiley and Sons, Inc.

Simon, H. A. (1954). Spurious correlation: A causal interpretation. *Journal of the American Statistical Association*, *49,* 467–479

Spirtes P., Glymour, C., & Scheines, R. (1993). *Causation, prediction and search (2nd ed.).* New York, NY: Springer-Verlag.

Wermuth, N. (1992). On block-recursive regression equations. *Brazilian Journal of Probability and Statistics*, 6, 1–56.

Wold, H. O. (1954). Causality and econometrics. *Econometrica*, 22, 162–177.

Woodward, J. (2007). Causation with a human face. In H. Price, R. Corry (Eds.), *Causation, physics and the constitution of reality (pp. 66-105).* New York, NY: Oxford University Press.

13. Endnotes

1. Hurwicz (1962) was one of the few Cowles economists who proposed a precise definition of "structural". Rather than simply characterizing structural models as those in which y is preceded by A, as here, his discussion brought in the idea of identification. In Hurwicz's usage, causal analysis consisted of determining the effects of alterations in structure. In that analysis the elements of A and B were characterized as constants but treated as external variables. Here, in contrast, we maintain the status of A and B as matrices of constants.

2. Hurwicz (1962) wrote the prototype structural model as

$$Ay = b, \qquad (16)$$

where b is a vector of constants and A is not necessarily square. This formulation allows the possibility that the model has fewer equations than variables, so that some of the variables are external. Hurwicz's notation did not distinguish internal from external variables. Failing to do so left it unclear which variables the model is intended to explain.
Other Cowles analysts, such as Wold (1954), did distinguish between external and internal variables. In some applications of causation analysis the assumption of linearity is unsuitable. For example, in analyzing treatment effects the treatment variable, which is binary, is specified to be a function of real-valued external variables. Such a function is necessarily nonlinear. See LeRoy (2018) for analysis of such models.

3. Here x can include both observed and unobserved variables. The coefficients with respect to unobserved external variables are well defined only subject to an arbitrary scaling of the latter. The scaling usually adopted is to set either α_{ij} or γ_{ij}, depending on whether one is working with the structural form or the reduced form, equal to 1 when x_j is unobserved. We follow this convention.

4. Thus analyses of interventions differ from comparative statics or comparative dynamics exercises, in which changes in constants are acceptable. This is so because the purpose of the latter exercises is to compare different models, not to determine the effects of changing the value of an external variable in a given model.

5. In the Cowles treatment of causation, and also in many recent discussions in the philosophy literature, analysts insisted that causal interpretation of a model requires a property of invariance. The meaning of invariance in the context of implementing alterations of a model's structure was never made clear despite much discussion. However, with interventions characterized as consisting of hypothetical changes in the values of external variables rather than as general structural changes, failure of invariance can only mean that terms specified as constants should instead be modeled as variables. In well-specified models labeling α as a constant means that α really is constant. Therefore that variable is not a candidate for intervention, and is not affected by interventions. Reminding analysts that if their models are misspecified their diagnoses of causation are likely to be wrong is hardly necessary. We see that invariance disappears as a feature of causal attributions that requires extended discussion.

6. In Wold (1954) (see also Wermuth, 1992) it is argued that multiequation models that are fully recursive inherit many of the properties of single-equation models. The fact that y_1, but not y_2, IN-causes y_3 in the model (4)-(6) suggests that this result does not carry over directly to causation.

7. Here and throughout we take the definition of genericity to exclude special cases (but not to rule out the presence of zeros in the G matrix, as in (8)-(9)). For the simplest example of what is ruled out, note that the restriction $\gamma_{21} = \gamma_{22} = \alpha_{21} = 0$ satisfies the condition (10), but we do not have $y_1 \Rightarrow y_2$. This is so because under the restriction just stated y_2 becomes a rescaled version of the external variable x_3, implying that interventions on y_1 do not cause changes in y_2. The stated restriction is nongeneric within the space of parameter values satisfying (10), so we ignore such cases.
In economic models, particularly those modeling agents assumed to have rational expectations, one

 DOI: 10.5709/ce.1897-9254.276

finds relations of the form $y_1 = y_2 - \delta y_3$, *where* δ *is the regression coefficient of* y_2 *on* y_3. *Here* y_1, *the regression residual, and* y_3, *the regression explanatory variable, are uncorrelated by construction. Such relations violate the genericity requirement since* δ *is a function of the other parameters in the model. One way to handle this problem is to relabel* δ *as an internal variable and include the function determining its value as an equation in the model. This, however, results in a nonlinear model.*

8. *To make the same point in vector-matrix notation, note that the restricted reduced form can be written as*

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & 0 \\ \alpha_{21}\beta_{11} & \alpha_{21}\beta_{12} & \beta_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \qquad (17)$$

This can be shown to coincide with the structural form (12)-(13), written in vector-matrix form as

$$\begin{bmatrix} 1 & 0 \\ -\alpha_{21} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & 0 \\ 0 & 0 & \beta_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad (18)$$

by inverting the matrix and multiplying.

9. *As noted in LeRoy (2016), the above representation of IN-causation in terms of graphs in which all paths from the external set of the cause variable to the effect variable pass through the cause variable is described in Woodward (2007).*

 DOI: 10.5709/ce.1897-9254.276