Journal of Economic Dynamics and Control 2 (1980) 329–352. © North-Holland

TESTING FOR CAUSALITY

A Personal Viewpoint

C.W.J. GRANGER

University of California, San Diego Campus, La Jolla, CA 92093, USA

Received February 1980, final version received June 1980

A general definition of causality is introduced and then specialized to become operational. By considering simple examples a number of advantages, and also difficulties, with the definition are discussed. Tests based on the definitions are then considered and the use of post-sample data emphasized, rather than relying on the same data to fit a model and use it to test causality. It is suggested that a Bayesian viewpoint should be taken in interpreting the results of these tests. Finally, the results of a study relating advertising and consumption are briefly presented.

1. The problem and a definition

Most statisticians meet the concept of causality early in their careers as, when discussing the interpretation of a correlation coefficient or a regression, most textbooks warn that an observed relationship does not allow one to say anything about causation between the variables. Of course this warning has much to recommend it, but consider the following special situation: suppose that X and Y are the only two random variables in the universe and that a strong correlation is observed between them. Further suppose that God, or an acceptable substitute, tells one that X does not cause Y, leaving open the possibility of Y causing X. In the circumstances, the strong observed correlation might lead to acceptance of the proposition that Y does cause X. This possibility occurs because of the extra structure imposed on the situation by the knowledge that X does not cause Y. As will be seen, the way structure is imposed will be important in definitions of causality.

The textbooks, having given a cautionary warning about causality, virtually never then go on with a positive statement of the form 'the procedure to test for causality is...', although a few do say that causality can be detected from a properly conducted experiment. The obvious reason for the lack of such positive statements is that there is no generally accepted procedure for testing for causality, partially because of a lack of a definition of this concept that is universally liked.

Attitudes towards causality differ widely, from the defeatist one that it is impossible to define causality, let alone test for it, to the populist viewpoint that everyone has their own personal definition and so it is unlikely that a generally acceptable definition exists. It is clearly a topic in which individual tastes predominate, and it would be improper to try to force research workers to accept a definition with which they feel uneasy. My own experience is that, unlike art, causality is a concept whose definition people know what they do not like but few know what they do like. It might therefore be helpful to present a definition that some of us appear to think has some acceptable features so that it can be publicly debated and compared with alternative definitions.

For ease of exposition, a universe is considered in which all variables are measured just at prespecified time points at constant intervals $t=1,2,\ldots$ When at time n, let all the knowledge in the universe available at that time be denoted Ω_n and denote by $\Omega_n - Y_n$ this information except the values taken by a variable Y_t up to time n, where $Y_n \in \Omega_n$. Ω_n includes no variates measured at time points $t>n$, although it may well contain expectations or forecasts of such values. However, these expectations will simply be functions of Ω_n. Ω_n will certainly be multivariate and Y_n could be, and both will be stochastic variables. To provide structure to the situation, the following axioms will be assumed to hold:

Axiom A. The past and present may cause the future, but the future cannot cause the past.

Axiom B. Ω_n contains no redundant information, so that if some variable Z_n is functionally related to one or more other variables, in a deterministic fashion, then Z_n should be excluded from Ω_n.

Thus, for example, if temperature is measured hourly at some location both in degrees Fahrenheit and degrees Centigrade, there is no point in including both of these variables in the universal information set.

Suppose that we are interested in the proposition that the variable Y causes the variable X. At time n, the value X_{n+1} will be, in general, a random variable and so can be characterized by probability statements of the form $\text{Prob}(X_{n+1} \in A)$ for a set A. This suggests the following:

General Definition. Y_n is said to cause X_{n+1} if

$$\text{Prob}(X_{n+1} \in A \mid \Omega_n) \neq \text{Prob}(X_{n+1} \in A \mid \Omega_n - Y_n) \quad \text{for some } A.$$

For causation to occur, the variable Y_n needs to have some unique information about what value X_{n+1} will take in the immediate future.

The ultimate objective is to produce an operational definition, which this is certainly not, by adding sufficient limitations. This process will be discussed in section 3, and the definition will also be defended there. In the following section some more general background material will be introduced which will, hopefully, make the defence a little easier.

2. A variety of viewpoints on causality

The obvious place to look for definitions of causality and discussions of the concept is the writings of philosophers on the topic, of which there have been plenty from Aristotle onwards. A useful discussion of parts of this literature can be found in Bunch (1963). I think that it is fair to say that the philosophers have not reached a consensus of opinion on the topic, have not found a definition that a majority can accept and, in particular, have not produced much that is useful to practising scientists. Most of the examples traditionally used by philosophers come from classical physics or chemistry, such as asking what causes the flame when a match is struck, or noting that applying heat to a metal rod causes it to become longer. Much of the literature attempts to discuss unique causes in deterministic situations, so that if A occurs then B must occur. Although most writers seem to agree with Axiom A, that causes must precede effects, even this is not universally accepted. Quite a few philosophers, at least in the past, seem to believe that causes and effects should be contiguous both in time and space, which undoubtedly reflects the pre-occupation with classical physics. Social scientists would surely want to consider the possibility that an event occurring in one part of the world could cause an event elsewhere at a later time. The philosophers are not constrained to look for operational definitions and can end up with asking questions of the ilk: 'If two people at separate pianos each strike the same key at the same time and I hear a note, which person caused the note that I hear?' The answer to such questions is, of course: 'Who cares?' For an interesting discussion of the lack of usefulness of the philosophers' contribution by a pair of lawyers, another group which clearly requires an operation definition of causation; see Hart and Honoré (1959). They take the viewpoint that 'the cause is a difference to the normal course which accounts for the difference in the outcome'. They also point out that legally this difference can be not doing something, 'as the driver did not put on the brakes, the train crashed'. One interesting aspect of the philosophers' contribution is that they often try to discuss what the term causality means in 'common usage', although they make no attempt to use common usage terms in their discussion. Rather than trying to decide what the public thinks they mean by such a difficult concept as causality, it may be preferable to try to influence common usage towards a sounder definition.

The philosophers and others have provided a variety of definitions, but no attempt to review them will be made here, as most are of little relevance to statisticians. Once a definition has been presented, it is very easy for someone to say 'but that is not what I mean by causation'. Such a remark has to be taken as a vote against the particular definition, but it is entirely destructive rather than constructive. To be constructive, the critic needs to continue and provide an alternative definition. What is surely required is a menu of definitions that can be discussed and criticized but at least defended by someone. Only by providing such a menu can a debate be undertaken which, hopefully, will result in one, or a few, definitions that can receive widespread support. I believe that definitions should be allowed to evolve due to debate rather than be judged solely on a truth or not scale. It is possible, as has been suggested, that everyone has their own definition so that no convergence will occur, but this outcome does seem to be unlikely.

Before proceeding further, it is worthwhile asking if there is any need, or demand, for a testable definition of causality. It is worth noting that the Social Science Citation Index lists over 1000 papers with words such as causal, causation or causality in their titles, and in a recent five-year period the Science Citation Index lists over 3000 such articles. Papers mentioning such words in the body of the paper, not in the title, are vastly more numerous. There does therefore seem to be a need for a widely accepted definition. Statisticians already have methods for measuring relationships between variables, but causal relations may be thought of as being in some sense deeper than the ordinarily observed kind. Consider the following three time series:

X_t = number of patients entering a maternity hospital in day t,
Y_t = number of patients leaving the same hospital in day t,
Z_t = ice cream sales in the same city in day t.

It seems very likely that the series X_t is useful in forecasting Y_t, and it is also possible that Z_t may appear to be useful in forecasting Y_t, as both variables contain seasonal components. However, most people would surely expect that if a more careful analysis was conducted, using perhaps longer data series, larger information sets including more explanatory variables or more sophisticated techniques, then the observed (forecasting) relationship between X_t and Y_t is likely to continue to be found, whereas that between Z_t and Y_t may well disappear. The deeper relationship is a candidate for the title causal. There thus appears to be both a need, and a demand, for techniques to investigate causality. The possible uses of a causal relationship, if found, will be discussed below.

It has been suggested that although such deeper relations need to be named, that name should not involve words like 'cause' or 'causality', as these words are too emotion-laden, involve too much preconception and have too long a history. Alternative phrases such as 'due to', 'temporally interrelated', 'temporally prior' and 'feedback free' have been proposed, for example. To my mind, this suggestion reflects a basic misunderstanding about language and its use. Most of the components of a language are just a notation, with generally agreed meanings. If I use words such as 'apple' or 'fear', I will not need to first define them, as it is understood that most people mean approximately the same thing by them. Occasionally, with unusual or technical words, such as 'therm' or temperature', I might need to add a definition. If I start a piece of written work, or a lecture, by carefully defining something, then I can use this as a notation throughout, such as distribution, mean, or variance. If my definition is quite different from general usage, then I may be unpopular but will not be logically incorrect, as, for example, if I write cos$x$ for what is usually denoted by x^3. As causation has no generally accepted definition, this criticism cannot apply. Provided I define what I personally mean by causation, I can use the term. I could, if I so wish, replace the word cause throughout my lecture by some other words, such as 'oshkosh' or 'snerd', but what would be gained? It is like saying that whenever I use x, you would prefer me to use z. If others wanted to refer to my definition, they can just call it 'Granger causality' to distinguish it from alternative definitions. There already exist many papers in economics which do just that, some of which are referenced later, and no misunderstanding occurs. If it is later observed that which is called 'Granger causality' is identical to the definition introduced by some earlier writer, then the name should be altered. In fact, I would be very surprised if the definition to be discussed in the next section has not been suggested many times in the past. Part of the definition was certainly proposed by Norbert Wiener (1958). It would not be a telling argument to appeal to 'common usage' in connection with the words cause or causality, as statisticians continually use words in ways different from common usage, examples being mean, variance, moments, probable, significant, normal, regression and distribution.

These remarks made so far in this section are designed to defuse certain criticisms that can be made of what is to follow. My experience suggests that I will be unsuccessful in this aim.

When discussing deterministic causation, philosophers distinguish two cases:

(a) Necessity — if A occurs, then B must occur.
(b) Sufficiency — if I observe B did occur, this means that A must have occurred.

For example, if one has a metal rod, then event A might be that one heats the rod and event B is that the rod expands. Although causality is defined for pairs of sequences, or functions, obeying axiom A in parts of mathematical science, any statistician or any worker dealing with data generated by an animal body, a person's behavior, part of an economy or an atmosphere, for example, will not be happy with these deterministic definitions. Rather than saying 'If A occurs, then B must occur', they would probably be happier with statements such as 'If A occurs, then the probability of B occurring increases (or changes)'. For example, if a person smokes, he does not necessarily get cancer, but he does increase the probability of cancer. If a person goes sailing, he does not necessarily get wet, but he does increase the probability of getting wet. It is therefore important for a useful definition to deal with stochastic events or processes. It is interesting to note that the advent of quantum physics had a big impact on the philosophical writings about causality, which had relied heavily on classical physics for examples. Bertrand Russell, in particular, dramatically changed his views of causality at that time.

There have, of course, been several attempts to introduce probabilistic theories of causality. A particularly convincing attempt, well worth reading, is that by Suppes (1970). One of his definitions is:

An event $B_{t'}$ (occurring at time t') is a prima facie cause of the event A_t if and only if (i) $t'<t$, (ii) $\text{Prob}(B_{t'})>0$, and (iii) $\text{Prob}(A_t \mid B_{t'})>P(A_t)$.

One might observe a large African population, for example, and find that the probability of not getting cholera is 0.91 but that of those inoculated against the disease, the probability of not getting cholera is 0.98. If A_t is not getting cholera and B_t is inoculation, then the evidence suggests that 'inoculation is a Suppes prima facie cause of not getting cholera'. Note that, by replacing A and B with their complements, the same evidence is also likely to lead to the conclusion 'not having inoculation is a Suppes prima facie cause of getting cholera'. There is obvious arbitrariness in practice in defining an event. If the inequality in (iii) is reversed, Suppes talks of negative causation. Nevertheless, for probabilistic events, rather than variables or processes, the discussion by Suppes is very useful and is certainly potentially applicable to a series of properly conducted random experiments.

Good (1961, 1962) has a somewhat similar definition, although he effectively hides it amongst 24 assumptions and 17 theorems combined with very little interpretation. If E and F are two events with F occurring before E, then he says that there is a tendency for F to cause E, given some state of the universe, if $\text{Prob}(E \mid F)>\text{Prob}(E \mid \text{not } F)$. It would be a lengthy task to critically discuss and compare such definitions, and so I will not attempt it at this time.

At the very start of this paper, the case where random variables X and Y are correlated, but God tells you that causation in one direction is impossible, was briefly discussed. Virtually all definitions of causality require some imposed structure, such as that provided here by God. In many definitions, Axiom A provides this structure, but not all definitions follow this route. The causality concepts discussed by Simon, Wold and Blalock [see Blalock (1964)] and others do not require Axiom A but do presume special knowledge about the structure of relations between two or more variables. Given this structure, the possibility of causal relationships can then be discussed, usually in terms of the vanishing or not of correlation or partial correlation coefficients. Because these definitions require a number of assumptions about structure to be true, they will be called conditional causation definitions. If the assumptions are correct, or can be accepted as being correct, these definitions may have some value. However, if the assumptions are somewhat doubtful, these definitions do not prove to be useful. Sims (1977) has discussed the Simon and Wold approach and found it not operational in practice. Certainly there has been little use made of these definitions in recent years, at least in economics. The 'path analysis' of Sewell Wright (1964) has similarities with the Wold and Simon approach, but he does state that he would prefer to use his analysis together with Axiom A, which would bring it nearer to the definition discussed in the next section. The full question of priority in these matters is a complex one and, I think, need not detain us here.

The question of whether any real statement can be made about causality based just on statistical data is clearly an important one. Naturally, as a statistician, I think that proper statements can be made, if they are carefully phrased. The link between smoking and cancer provides an example. So far the only convincing link has been a statistical one, but it is now generally accepted. The real question for most people is not 'Does smoking cause cancer?' but rather 'How does smoking cause cancer?' Before the accumulation of statistical evidence, people could be thought of as having subjective, personal probabilities that the statement 'Smoking causes cancer, in a statistical sense' is true. Since the evidence has been presented, for most people these subjective probabilities have greatly increased and may well be near one. The weight-of-evidence is certainly in favor of this causality. Smoking is certainly a prima facie cause of cancer and is probably more than that, in the opinion of the majority. A decision, such as for an individual to stop smoking or for a government to ban it, could be a wrong one, but statisticians are used to making decisions under uncertainty and realize that when properly based on the statistical evidence can be wrong but are usually correct.

There is one problem with the statistical approach which was pointed out by the philosopher Hume as applying to any testing procedure. It is always possible that the evidence from the past may be irrelevant, as causation can change from the past to the future. It is therefore necessary to introduce:

Axiom C. All causal relationships remain constant in direction throughout time.

The strength, and lags, of these relationships may change, but causal laws are not allowed to change from positive strength to zero, or go from zero to positive strength, through time. This axiom is, of course, central to the applicability of all scientific laws and so is generally accepted, even though it is not necessarily true.

3. An operational definition

The general definition introduced above is not operational, in that it cannot be used with actual data. To become operational, a number of constraints need to be introduced. To do this, it is convenient to first re-state the general definition. Suppose that one is interested in the possibility that a vector series $\boldsymbol{Y}_t$ causes another vector $\boldsymbol{X}_t$. Let J_n be an information set available at time n, consisting of terms of the vector series $\boldsymbol{Z}_t$, i.e.,

$$J_n : \boldsymbol{Z}_{n-j}, \qquad j \geqq 0.$$

J_n is said to be a proper information set with respect to $\boldsymbol{X}_t$ if $\boldsymbol{X}_t$ is included within $\boldsymbol{Z}_t$. Further, suppose that $\boldsymbol{Z}_t$ does not include any components of $\boldsymbol{Y}_t$, so that the intersection of $\boldsymbol{Z}_t$ and $\boldsymbol{Y}_t$ is zero. Further, define

$$J'_n : \boldsymbol{Z}_{n-j}, \boldsymbol{Y}_{n-j}, \qquad j \geqq 0,$$

so that J'_n is the information set J_n plus the values in past and present $\boldsymbol{Y}_t$.

Denote by $F(\boldsymbol{X}_{n+1} \mid J_n)$ the conditional distribution function of $\boldsymbol{X}_{n+1}$ given J_n, so that this distribution has mean $E[\boldsymbol{X}_{n+1} \mid J_n]$. The notation using other information sets is obvious. These expressions are used in the following definitions:

Definition 1. $\boldsymbol{Y}_n$ does *not cause* $\boldsymbol{X}_{n+1}$ with respect to J'_n if

$$F(\boldsymbol{X}_{n+1} \mid J_n) = F(\boldsymbol{X}_{n+1} \mid J'_n),$$

so that the extra information in J'_n has not affected the conditional distribution. A necessary condition is that

$$E[\boldsymbol{X}_{n+1} \mid J_n] = E[\boldsymbol{X}_{n+1} \mid J'_n].$$

Definition 2. If $J'_n \equiv \Omega_n$, the universal information set, and if

$$F(\boldsymbol{X}_{n+1} \mid \Omega_n) \neq F(\boldsymbol{X}_{n+1} \mid \Omega_n - \boldsymbol{Y}_n),$$

then $\boldsymbol{Y}_n$ is said to *cause* $\boldsymbol{X}_{n+1}$.

Definition 3. If

$$F(X_{n+1} \mid J'_n) \neq F(X_{n+1} \mid J'_n),$$

then Y_n is said to be a *prima facie cause* of X_{n+1} with respect to the information set J'_n.

Definition 4. Y_n is said *not to cause* X_{n+1} *in mean* with respect to J'_n if

$$\delta_{n+1}(J'_n) = E[X_{n+1} \mid J'_n] - E[X_{n+1} \mid J_n]$$

is identically zero.

Definition 5. If $\delta_{n+1}(\Omega_n)$ is not zero, then Y_n is said to *cause* X_{n+1} *in mean.*

Definition 6. If $\delta_{n+1}(J'_n)$ is not identically zero, then Y_n is said to be a *prima facie cause in mean* of X_{n+1} with respect to J'_n.

Definition 2 is equivalent to the general definition introduced in the first section, which was discussed in Granger and Newbold (1977). If a less general information set than the universal set is available, J'_n, then a prima facie cause can occur, as in Definitions 1 and 3. These definitions can be strengthened by adding phrases such as 'almost surely', or 'except on sets of measure zero' at appropriate points, but as these will not help towards the eventual aim of an operational definition capable of being tested, such niceties are ignored.

If, rather than discussing the whole distribution of X_{n+1}, one is content with just point forecasts using a least squares criterion, then the final three definitions become relevant. To ask for causality in mean is much less stringent than asking for full causality, but does provide a definition much nearer to being operational. If one wishes to use some criterion other than least squares, this can be done, but point forecasts will be made much more difficult to obtain. Definition 6 can be rephrased: Let $\sigma^2(X \mid J_n)$ be the variance of the one-step forecast error of X_{n+1} given J_n, and similarly for $\sigma^2(X \mid J_n, Y) \equiv \sigma^2(X \mid J'_n)$, then Y is a prima facie cause of X, with respect to J', if $\sigma^2(X \mid J_n, Y) < \sigma^2(X \mid J_n)$. Thus knowledge of Y_n increases one's ability to forecast X_{n+1}, in a least squares sense. This corresponds to a definition hinted at by Wiener (1958), introduced specifically in Granger (1964) and Granger and Hatanaka (1964), re-introduced in Granger (1969), amplified and applied by Sims (1972, 1977) and then used by numerous authors since, including Black (1978), Williams, Goodhart and Gowland (1976), Skoog (1976), Sargent (1976), Mehra (1977), Gordon (1977), Feige and Pearce (1976a, b), Ciccolo (1978), and Caines, Sethi and Brotherton (1977).

However, it should be said that some of the recent writers on this topic, because they have not looked at the original papers, have evolved somewhat unclear and incorrect forms of this definition. It is rather like the party game where a phrase or rumor is whispered around the room, ending up quite differently from how it started.

In this newer formulation, Axiom B becomes:

Axiom B′. $F[Y_n|J_n]$ is not a singular distribution, so that is not that of a variable taking only a constant value. This implies that Y_n is not deterministically related to the contents of J_n.

If purely time-series techniques are used to generate one-step forecasts, these forecasts will usually be linear functions of the information set because of the present state of the art, although some progress in the use of certain non-linear models is occurring; see for instance, Granger and Andersen (1978) and Swamy and Tinsley (1980). However, if forecasts are made from reduced form equations derived from a, possible non-linear, structural econometric model, then the contents of the information set may be utilized non-linearly. The definitions discussed here do not require that only linear models are used, although most of the actual applications so far and much of the theoretical discussions have concentrated on the linear case. If the available information in J_n is used only linearly, then it may be possible to observe that Y_n is a linear prima facie cause in mean of X_{n+1} with respect to J'_n and with the available modelling and forecasting techniques this provides the operational definition that is being sought. For the remainder of this paper the phrase 'linear prima facie cause' will be replaced simply by 'cause' for convenience, unless a more general case is being considered. The definition as given relates a pair of vectors, Y_n and X_{n+1}, but the usual case will be concerned with just a pair of individual series, Y_n and X_{n+1}. Further, to actually model data it will usually be necessary to either assume that the series are stationary or belong to some simple class of models with time-varying parameters. Again, this is not strictly necessary for the definition but is required for practical implementation.

There are a number of important implications of the definition of cause here developed. If, for example, it is found that Y_n causes X_{n+1} with respect to some information set, then this implies no restrictions on whether or not X_n causes Y_{n+1}; this second causation may occur but need not. If both causations occur, one may say that there is feedback between the two series X_t and Y_t. A simple example is

$$X_t=\varepsilon_t+\eta_{t-1}, \qquad Y_t=\eta_t+\varepsilon_{t-1},$$

where ε_t, η_t are a pair of independent white noise series. Further, if there are three series X_t, Y_t and Z_t and it is observed that X causes Y and Y causes Z, then it is not necessarily true that X causes Z, although it can occur.

Example 1. $\quad X_t = \varepsilon_t, \quad Y_t = \varepsilon_{t-1} + \eta_t, \quad Z_t = \eta_{t-1},$

where again ε_t, η_t are independent white noises. There are four information sets that need to be considered: $J_n(X, Y)$ — consisting of past and present X_{n-j}, Y_{n-j} $(j \geqq 0)$, and similarly, $J_n(X,Z)$, $J_n(Y,Z)$, and $J_n(X,Y,Z)$ — consisting of past and present X_{n-j}, Y_{n-j}, Z_{n-j} $(j \geqq 0)$. Then clearly X causes Y with respect to either $J_n(X,Y)$ or $J_n(X,Y,Z)$, Y causes Z with respect to $J_n(Y,Z)$ and to $J_n(X,Y,Z)$, but X does not cause Z with respect to $J_n(X,Z)$ but it does cause Z with respect to $J_n(X,Y,Z)$. This last result occurs because Z_{n+1} is completely predetermined from Y_{n-j}, X_{n-j} $(j \geqq 0)$ but not from just Y_{n-j} $(j \geqq 0)$. The importance of stating the information set being utilized is well illustrated by this example. A further example shows a different situation:

Example 2. $\quad X_t = \varepsilon_t + \omega_t, \quad Y_t = \varepsilon_{t-1}, \quad Z_t = \varepsilon_{t-2} + \eta_t,$

where ε_t, η_t and ω_t are three independent white noises. Here X causes Z in $J_n(X,Z)$ but not in $J_n(X,Y,Z)$.

One thing that is immediately clear from the definition is that if Y_n causes X_{n+1}, then $Y'_n = a(B)Y_n$ causes $X'_{n+1} = b(B)X_{n+1}$ if $a(B)$ and $b(B)$ are each one-sided filters of the form $a(B) = \sum_{j=0}^{\ell} a_j B^j$. However, if two-sided filters are used, as occurs for example in some seasonal adjustment procedures, then causality can obviously be lost because Axiom A is disrupted.

The use of proper information sets, that is, sets including the past and present values of the series to be forecast X_t, does have the following important implication: it is impossible to find a cause for a series that is self-deterministic, that is, a series that can be forecast without error from its own past. The basic idea of the causal definition being discussed is that knowledge of the causal variable helps forecast the variable being caused. If a variable is perfectly forecastable from its own past, clearly no other variable can improve matters.

Example 3. $\quad X_t = \alpha + bt + ct^2 \quad \text{and} \quad Y_t = dX_{t+1}.$

Then the following three equations generate X_t exactly, without error, i.e.:

$$X_t = \alpha + bt + ct^2, \quad X_t = d^{-1} Y_{t-1}, \quad X_t = 2X_{t-1} - X_{t-2} + 2c,$$

so that at first sight X_t is 'caused' by time, or by Y_{t-1}, or by its own past. If all three equations fit equally well, that is perfectly, it is clear that no kind of data analysis can distinguish between them. It is therefore obvious that in this circumstance a statistical test for causality is impossible, unless some extra structure is imposed on the situation. It may be noted that causality tests can be made with variables that contain deterministic components, as proved formally by Hosoya (1977), but with this definition one cannot say that the deterministic component of one variable causes the deterministic component of another variable.

4. Some difficulties

Virtually any sophisticated statistical procedure has some problems associated with it, and there is every reason that this will be true also with any operational definition of causality. These difficulties can either be intrinsic to the definition itself or be associated with its practical implementation.

Some of the difficulties will arise because of data inadequacies. One obvious problem arises when the data is gathered insufficiently frequently. Suppose that a change in wood prices causes a change in furniture prices one week later, but prices are only recorded monthly; then the true causal relationship will appear to be instantaneous. It is perhaps worth defining 'prima facie apparent instantaneous causality in mean', henceforth instantaneous causality, between X_{n+1} and Y_{n+1} with respect to J'_n if

$$E[X_{n+1} \mid J_n, Y_{n+1}] \neq E[X_{n+1} \mid J_n].$$

Although the phrase 'instantaneous causality' is somewhat useful on occasions, the concept is a weak one, partly because Axiom A is not being applied and because, at least in the linear case, it is not possible to differentiate between instantaneous causation of X by Y, of Y by X or of feedback between X and Y, as simple examples show. If extra structure is imposed, it may be possible to distinguish between these possibilities, as will be discussed below. If one totally accepts Axiom A, then instantaneous causality will either occur because of the data collection problem just mentioned or because both series have a common cause which is not included in the information set J'_n being used.

The problem of missing variables, and consequential mis-interpretation of one's results, is a familiar one in those parts of statistics which consider relationships between variables. A simple example of apparent causation due to a common cause is:

Example 4. $\quad Z_t = \eta_t, \quad X_t = \eta_{t-1} + \delta_t, \quad Y_t = \eta_{t-2} + \varepsilon_t,$

where ε_t, η_t and δ_t are independent white noises. Here Z_t is causing both X_t and Y_t with respect to information sets $J_n(X, Z)$, $J_n(X, Z)$ and $J_n(X, Y, Z)$, but X_t is causing Y_t in $J_n(X, Y)$ but not $J_n(X, Y, Z)$.

This apparent causation of Y by X in $J_n(X, Y)$ may be thought of as spurious because it vanishes when the information set is expanded, something one would not expect with a true cause. Sims (1977) has studied the system

$$Y_t = c(B)Z_t + \varepsilon_t, \qquad X_t = d(B)Z_t + \eta_t,$$

and found that it is unlikely to give rise to a spurious one-way causation between X and Y based on $J_n(X, Y)$, although presumably a feedback relationship between X and Y is more likely to be found.

An important case where missing variables can lead to misleading interpretations is when one variable is measured with an error having time structure. The following example illustrates the difficulty:

Example 5. $\quad X_t = \eta_t, \quad Y_t = \delta_t, \quad Z_t = X_t + \varepsilon_t + \beta\varepsilon_{t-1},$

where η_t and δ_t are white noises with correlation $(\eta_t, \delta_s) = 0$, $t \neq s$, but this correlation equals λ when $t = s$ and ε_t is a white noise independent of η_t and δ_t. Z_t may be thought of as X_t with an MA(1) measurement error. There is no causation between X_t and Y_t apart from instantaneous causation. As Z_t is the sum of a white noise and an MA(1) term, it will be MA(1), so that there exists a constant θ with $|\theta| < 1$ and a white noise series e_t so that

$$Z_t = (1 + \theta B)e_t.$$

It follows that

$$e_t = (1 + \theta B)^{-1}\eta_t + (1 + \theta B)^{-1}(\varepsilon_t + \beta\varepsilon_{t-1}).$$

The one-step forecast of Z_{n+1} using Z_{n-j} $(j \geqq 0)$ is just θe_n with error e_{n+1}, but this error is a function of η_{n-j} $(j \geqq 0)$ which is correlated with δ_{n-j} $(j \geqq 0)$ which is equal to Y_{n-j}. It thus follows that the Y_{n-j} will help forecast Z_{n+1}, so that apparently Y_n causes Z_{n+1} with respect to $J_n(Y, Z)$, but this would not be the case if X_{n-j} were observable, so that $J_n(X, Y, Z)$ could be considered. This result is at first sight quite worrying, as in many disciplines, such as economics, variables are almost inevitably observed with error, so that x_t — the missing variable — will always be missing. However, as the results of Sims (1977) and Newbold (1978) show, by no means does the addition of measurement error to variables necessarily produce spurious causations, as the error has to have particular time-series structure compared to the original series. Nevertheless, the possibility of mis-leading results occurring from a common type of situation has to be kept in mind when interpreting results.

Another situation which needs care in interpretation is when the time that a variable is recorded is different from the time at which the event occurred that led to the variable's value. For example, March unemployment figures in New York City and New York State may not become known to the public until April 1 and April 15, respectively. The values must be associated with March, not the time of their release, otherwise spurious causation may well occur. A further example of this problem is the relationship between lightning and thunder. As the lightning is usually observed before the thunder, because light travels faster than sound, it might seem that lightning causes thunder. However, both are manifestations of what is essentially the same event, and if the observations are placed at the time of the original electrical discharge, the spurious causation disappears. If one is being pedantic, the light-producing part of the discharge does occur before the sound-producing part, but both lightning and thunder do have a common cause.

A further interpretation problem can arise because of Axiom B. Suppose one has three variables which are related through some linear identity, such as

$$\text{Work force} = \text{Unemployed} + \text{Employed}.$$

It is clear that all three variables cannot be in the information set to be used, but it is not necessarily obvious which one should be excluded. If, for example, total consumption is caused by size of work force, but this latter variable is excluded, one may expect to find that numbers of both unemployed and employed appear to cause consumption. Once one is aware of such interpretational difficulties, it is not difficult to invent strategies for analyzing them, such as excluding different variables and repeating the analysis, or by testing equality of certain coefficients in the model, for example.

One apparently serious problem of interpretation, which is suggested by the thunder and lightning example, arises from the idea of a leading indicator. Suppose that X causes both Y and Z, but that the causal lag is shown from X to Y, then from X to Z. If now X is not observed, Y will appear to cause Z. Example 4 shows such a situation. The search for such leading indicators occurs in various fields. In economics, for example, the Bureau of the Census publishes a list of such indicators, plus an index, which are supposed to help indicate when the economy is about to experience a down-turn or an up-turn. A number of possible leading indicators for earthquakes are also being considered, an example being unusual animal behavior. If leading indicators are included in an information set, tests may well indicate prima facie causality. In most cases this will be just another example of the missing variable problem. Sometimes the missing variable will be available and, when added to the information set, the leading indicator will no longer appear to cause. In other cases the missing variable is not observable and, when this occurs, it will not always be obvious whether a variable is a cause or merely a leading indicator. This relates to the question of how to interpret the outcomes of the causality tests, which are discussed in the next section, given that the information sets used in practice will always be limited in extent. One way of viewing the test results is as an informal Bayesian. A person may start with a prior belief, or probability, that X causes Y, say, and then use the test results to alter this probability, if the test is viewed as being relevant. If the causality definition being invoked by the test is not liked, then the probability need not change. Once more, the purely personal aspect of the attitude towards causality is seen to be of real importance. For some cases, such as whether changes in animal behavior causes earthquakes, the prior probability will start at zero and will remain there despite any test results, so these changes will still be relevant leading indicators, but will not be thought of as causes. Surely no one believes that a lot of cows jumping up and down will trigger an earthquake. On the other hand, one may have a prior probability that smoking is a cause of heart failure, and, after an appropriate and well-conducted test, this probability may be changed, depending on the results of the test.

One very important aspect of forming these prior probabilities will be the availability, or otherwise, of some convincing theory for the causal situation being considered. If there exists, for example, some piece of economic theory that one strongly believes to be true, this could certainly influence, or even fully determine, one's belief about a causal relationship. In what is at present the only carefully thought out critical discussion of what has been called Granger causality, Zellner (1978) strongly proposes that causality should only be considered in the context of some accepted theory. He states: 'Perhaps a more satisfactory position would be to define causality, as Feige and others have, in terms of predictability according to well thought out economic laws.' In terms of the General Definition, or Definition 2, this approach adds nothing, as these definitions assume total knowledge of all relevant distribution functions. Any true or correct law or theory is simply an observed constraint that is found to apply to these distribution functions, and so knowledge of it does not add to the assumed available information set. However, when trying to make the definition more operative, the use of some theory may well be helpful, depending on the quality of the theory and on its nature. For example, if the theory simply tells one that certain variables need to be included in the information set to be utilized, that is extremely useful. It is also helpful to be told that certain variables can be safely omitted from the information set, as this will greatly simplify the data analysis. However, these would be rather vague theories and, I suspect, are not what Zellner had in mind. If the theory is much more specific, such as 'interest changes do not cause changes in production' or 'money supply changes do cause changes in prices', then, if true, analysis of other possible causes would again be simplified. But such theories may be precisely what the causality test is designed to verify. To impose any theory, one has conditional causality testing and, like all such situations, if the theory is correct it is helpful to use it, if it is incorrect one may well be worse off by its use. This is certainly true of any causality test that is conditional on the truth of some very specific theory. Whereas in many fields there may be theories, specific or not, that are generally accepted as being true, such theories are much more difficult to find in economics. It is interesting to note that Zellner in his paper never gives a single example of what he would consider to be a 'well thought out economic theory' nor even of a specific theory or law that is generally accepted by the majority of economists. Again, one returns to the personal belief aspect of causality testing; an individual may strongly believe some theory and is happy to test causality conditional on this theory, whereas someone else would not want to do that. One obvious place where a good theory would be particularly useful would be where extra structure is required to resolve causal directions in what appears to be instantaneous causality/feedback. For example, if sufficient structure can be put on a model to ensure identifiability — in the econometrician's sense of having a unique model — then a conditional causal test can be constructed. This is very much in the spirit of the Simon and Wold approach to causality, which is very well summarized in Zellner's article. However, it must be emphasized that only conditional causality can result, and this is potentially very much weaker than the unconditional causality definition discussed earlier.

The definitions of causation introduced in the previous section admittedly have a number of arbitrary aspects, some of which are potentially removable, others perhaps not. The data is assumed to be measurable on a cardinal scale, whereas actual data often occurs on different scales. If the data is intrinsically ordinal, I consider that it may be difficult to use these definitions, because of the lack of suitable distribution functions. However, it may be possible to build and evaluate forecasting models for such data, and so one aspect of the definitions will go through. With attribute data, without any natural order to the categories, the general definition remains unstable, but clearly the 'causality in mean' definitions are not relevant. This type of data is much nearer to the situation of one event causing another that was discussed by Suppes (1970) and Good (1961/62) and may often occur as the outcome of designed experiments. To be relevant to statisticians, a sequence of experiments will be required, as there seems to be no possibility of investigating causal relationships between unique events using statistical procedures.

A further arbitrary feature of the definitions is the use of one-step forecasts rather than h-step for any h. It is usually by no means clear what is the natural length of the step, and the pragmatic procedure is to use just the data period of the publicly available data, which can lead to the apparent instantaneous causation problem mentioned above. In the bivariate information set case, where one asks if Y causes X with respect to $J_n(X, Y)$, Pierce (1975) has shown that if Y causes X using an h-step forecasting criterion, with $h>1$, then it will necessarily be found that Y causes X with a one-step criterion. However, this does not seem to be true in the multivariate case:

Example 6. $\quad X_t=\varepsilon_t, \quad Y_t=\varepsilon_{t-2}+\eta_t, \quad Z_t=\varepsilon_{t-1}+\theta_t,$

where ε_t, η_t, θ_t are independent zero-mean, white noise series.

Here Z_n causes X_{n+1} with respect to $J_n(X,Z)$ and $J_n(X,Y,Z)$, Y_n causes X_{n+2} with respect to $J_n(X,Y)$ and $J_n(X,Y,Z)$, Y_n (or Y_{n-1}) causes X_{n+1} with respect to $J_n(X,Y)$ but not with respect to $J_n(X,Y,Z)$. Although some justification can be made that one-step forecasts are the most natural to consider, it will remain an arbitrary aspect of the definitions.

It is, on occasion, possible to distinguish between different types of causes by considering alternative information sets. For example, one might call Y a primary cause of X, if tests show this to be so for $J_n(X,Y)$, $J_n(X,Y,Z)$ and for all other information sets containing X and Y and any other series. A secondary cause might be one such that X causes Z in $J_n(X,Y,Z)$ but not in $J_n(X,Z)$, as illustrated in example one above. This example shows that X can cause Z, according to the definition, even though X and Z are statistically independent, provided that X can add further information to the primary cause, which in Example 1 is Y. The existence of such secondary causes may be upsetting to some readers, and so it might be relevant to alter the basic definition to deal only with primary causes. However, I personally would not, at this time, wish to emphasize such a change.

Most of the problems and difficulties discussed in this section relate not to the basic definition but with making it operational, in my opinion. Some are inherent to any statistical study using an incomplete or finite data set. Many of the difficulties become considerably reduced in inportance once care is taken with interpretation of test results.

In the following section a brief discussion of actual test procedures is presented, and in the final section some further important interpretational questions are considered, such as the relevance of control variables and the meaning of exogeneity.

5. Test procedures

There has been a lot of thought given in recent years to the question of how the above definitions can be actually tested, although the major attention has been given to the case of whether X causes Y with respect to $J_n(X,Y)$, that is, just the two-variable case. Although most empirical studies have considered this case, it is probably not a particularly important one in economics, as it is easy to suggest relevant missing variables. It is clear that more attention is needed on how to utilize bigger information sets. As the two-variable case has been well summarized recently by Pierce and Haugh (1977), only a few of the more important aspects will be discussed here. To give some structure to the discussion, consider the pair of zero-mean, jointly stationary series x_t, y_t, which are purely non-deterministic.

The moving-average, or Wold, representation can be denoted, following Pierce and Haugh, by

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{bmatrix} \psi_{11}(B) & \psi_{12}(B) \\ \psi_{21}(B) & \psi_{22}(B) \end{bmatrix} \begin{pmatrix} a_t \\ b_t \end{pmatrix}, \tag{1}$$

where each $\psi_{ij}(B)$ is a power-series, possibly infinite in length, in the backward operator B and $(a_t, b_t)'$ is a two-element white noise vector, with zero correlation between a_t and b_s, except possibly when $t=s$. Assuming that the moving average matrix operator is invertible, the corresponding autoregressive model can be denoted by

$$\begin{bmatrix} A(B) & H(B) \\ C(B) & D(B) \end{bmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} a_t \\ b_t \end{pmatrix}. \tag{2}$$

Rather than considering models for the actual series, one can equally well consider relationships between prewhitened series. If the filters $F(B)x_t = u_t$ and $G(B)y_t = v_t$ produce a pair of series u_t and v_t that are individually white noises, then moving average and autoregressive models will exist of the form

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} = \begin{bmatrix} \theta_{11}(B) & \theta_{12}(B) \\ \theta_{21}(B) & \theta_{22}(B) \end{bmatrix} \begin{pmatrix} a_t \\ b_t \end{pmatrix}, \tag{3}$$

and

$$\begin{bmatrix} \alpha(B) & \beta(B) \\ \gamma(B) & \delta(B) \end{bmatrix} \begin{pmatrix} u_t \\ v_t \end{pmatrix} = \begin{pmatrix} a_t \\ b_t \end{pmatrix}. \tag{4}$$

There are obviously relationships between the various operators, as described by Pierce and Haugh.

Denote the correlation between u_{t-k} and v_t by $\rho_{uv}(k)$ and consider the regression

$$v_t = \sum_{j=-\infty}^{\infty} \omega_j u_{t-j} + f_t, \tag{5}$$

where $\rho_{uv}(k) = (\sigma_u/\sigma_v)\omega_k$. Similarly, one can consider the regression

$$y_t = V(B)x_t + h_t. \tag{6}$$

Here $V(B)=(F(B)/G(B))\omega(B)$ and f_t, h_t are residuals which are uncorrelated with u_{t-j}, x_{t-j}, respectively, but are not necessarily white noises. Using this notation, Pierce and Haugh (1977) prove the following two theorems, amongst others:

Theorem 1. Instantaneous (prima facie) causality (in mean) exists if and only if the following equivalent conditions hold:

(i) *at least one of* $\operatorname{cov}(a_t, b_t), \gamma(0), \beta(0)$ *in* (4) *are non-zero, or*

(ii) *at least one of* $\operatorname{cov}(a_t, b_t), H(0), C(0)$ *in* (2) *are non-zero.*

In their 1977 paper, Pierce and Haugh had further conditions, such as $\rho_{uv}(0) \neq 0$ or $\omega_0 \neq 0$, but Price (1979) and Pierce and Haugh (1979) show that these conditions are not necessarily correct when there is feedback between x and y.

Theorem 2. y *is not a (prima facie) cause (in mean) of* x *if and only if the following equivalent conditions hold:*

(1) $\psi_{12}(B)$ *[equivalently* $\theta_{12}(B)$*] can be chosen zero.*
(2) $\theta_{12}(B)$ *is either* 0 *or a constant.*
(3) $\psi_{12}(B)$ *is either* 0 *or proportional to* $\psi_{11}(B)$.
(4) $V_j=0$ $(j<0)$ *in* (6).
(5) $\beta(B)$ *is either* 0 *or a constant.*
(6) $H(B)$ *is either* 0 *or proportional to* $A(B)$.
(7) $\rho_{uv}(k)=0$, *or equivalently* $\omega_k=0$ $(k<0)$.

If any of these conditions do not hold, then y will be a prima facie cause of x in mean with respect to $J_n(x, y)$. (1) and (4) were pointed out by Sims (1972), the first part of (6) was mentioned in Granger (1969), and that of (7) was emphasized in Granger and Newbold (1977). Multivariate generalizations of these conditions, concerning the possibility that the vector $\boldsymbol{y}$ may cause the vector $\boldsymbol{x}$, have been discussed by Caines and Chan (1975) and elsewhere. Because of this variety of equivalent conditions, there are clearly numerous statistical tests that can be devised based on these conditions. The performance of these tests needs further investigation, either using statistical theory or Monte Carlo study, especially as some are suspected to be occasionally biased or to be lacking in power.

My own experience has largely been with the autoregressive form (2), first fitting the bivariate model with $H(B)$ constrained to be zero and then re-fitting without this constraint, to see if a significant decrease in the variance of the residual for the x_t equation can be achieved. This experience, using both simulated and actual data as, for example, in Chiang (1978), suggests that misleading results do not occur but that the power is not particularly satisfactory. However these tests are not of considerable importance for two basic reasons: (i) they deal only with the bivariate case, whereas the more important applications are likely to involve more variables; and (ii) they are not properly based on the definitions presented above. This latter point arises because these definitions are explicitly based on the extra forecasting ability schieved from one information set over another, whereas the equivalent conditions given in Theorem 2, for example, make no mention of forecasts. This makes no difference for populations, as the definition of non-causation in mean and the conditions in Theorem 2 are then equivalent. However, if only a finite sample is available, as will always occur in practice, the equivalence disappears. Suppose that a sample is used to model the relationship between x_t and y_t in the autoregressive form (2) and the estimate of $H(B)$ is found to be significantly different from zero. Then the result is essentially saying that if this fact were known at the start of the sample, it could have been used to improve forecasts of x_t. This is quite different from actually producing improved forecasts. It is generally accepted that to find a model that apparently fits better than another is much easier than to find one that forecasts better. Thus tests based on the 'equivalent conditions' in Theorem 2 are just tests of goodness of fit, whereas the original definition requires evidence of improved forecasts. To satisfy this requirement, alternative models, based on different information, can be identified and estimated using the first part of the sample and then their respective forecasting abilities compared on the later part of the sample. The best way to actually test for differences in 'post-sample' forecasting ability and the optimum way to divide the sample into a modelling part and a forecast evaluation part need further investigation, but at least a test that is in sympathy with the basis of the definition would result.

An application of these ideas, in a two-variable case, is provided by Ashley, Granger and Schmalensee (1979), who consider possible causal relationships between aggregate advertising expenditures and consumption spending. They use a five-step procedure:

(i) Using a block of data, which is called the sample, each series is prewhitened by building ARIMA models, to get u_t, v_t as above.

(ii) The cross-correlations $\rho_{uv}(k)$ are examined to see if there is evidence of possible causal relationships.

(iii) For each indicated possible causal relationship, a model is built on these residuals u_t, v_t. If a one-way cause is suggested, the transfer function methods of Box and Jenkins (1970) may be utilized, but if a two-way causality appears to be present, the method for modelling this situation suggested in Granger and Newbold (1977) can be used.

(iv) The models in stages (i) and (iii) are then put together to suggest a model for the original data, in differenced form where necessary. This model is estimated, insignificant terms dropped and a final model achieved.
(v) The forecasting ability, in terms of mean-squared one-step forecast error, of the bivariate model and the single series ARIMA model, are then compared using post-sample data. If the bivariate model forecasts significantly better, then evidence of causation is found.

These stages are somewhat biased against finding causation, as, if in stage (ii) no evidence of causes is found, then no bivariate models will be constructed. The separation of the modelling period and the evaluation period does prevent evidence for spurious causation occurring because of data mining. However, a weakness is that if an important structural change occurs between the sample and the post-sample, the test will lose power. The relevance of Axiom C is evident. Ashley, Granger and Schmalensee, using quarterly data, find evidence that consumption causes advertising, but that advertising does not cause consumption except instantaneously. These results agree with parts of the advertising literature that find advertising expenditure is determined by management from previous sales figures and that advertising has little or no long-memory ability. On the other hand, these results might well be the opposite of the pre-conceptions of many economists, which illustrates both the relevance of performing a test and also of not relying on some partly formed theory.

6. Discussion and conclusions

The definition of causation proposed and defended above essentially says that X_{n+1} will consist of a part that can be explained by some proper information set, excluding Y_{n-j} $(j \geqq 0)$, plus an unexplained part. If the Y_{n-j} can be used to partly forecast the unexplained part of X_{n+1}, then Y is said to be a prima facie cause of X. It is clear that in practice the quality of the answer one gets from a test is related to the sophistication of the analysis used in deciding what is explained and by what. The definition also relies very heavily on Axiom A, that the future cannot cause the past, as using the 'arrow of time' imposes the structure necessary for the definition to hold. It also means that the definition does emphasize forecasting. If one does not accept Axiom A, the rest of the work connected with the definition becomes irrelevant. It is important to realize that the truth of Axiom A cannot be tested using the methods discussed in this paper. I should point out that the work by physicists on 'time-reversibility' does not seem to contradict Axiom A, as a careful reading of the review article by Overseth (1967) will show. Because of the way the definition is framed, and the tests based on it are organized, it is only appropriate for use with sequences of data. It cannot say anything about unique events or contribute to topics such as whether there exists an ultimate or first cause. Such topics have to remain the province of philosophers and theologians. In interpreting the test results it has been suggested above that one thinks in terms of changing personal beliefs about whether Y causes X. There is nothing essentially new in this suggestion, as it is certainly what occurs in practice. The definition and tests based on it provide a way to organize the available data in such a way that some workers will feel is appropriate for them to need to possibly change their prior probabilities. I leave to others the discussion of the effect of this procedure, and of the whole causation testing methodology on scientific methodology.

Some of the economists writing about what is called Granger causality have related this concept to the more familiar one of exogeneity; see, for example, Sims (1977) and Geweke (1978). When econometric models are constructed it is usual to divide variables into exogenous ($\boldsymbol{Z}$) and endogenous ($\boldsymbol{Y}$), and it is assumed that components of $\boldsymbol{Z}$ may cause components of $\boldsymbol{Y}$ but not vice versa. There is thus assumed to be a one-way causal relationship from $\boldsymbol{Z}$ to $\boldsymbol{Y}$. For estimation and econometric identification purposes, it is important that this classification be correct as questions of efficiency, model uniqueness and model specification are concerned. Tests for exogeneity are with respect not only to the information set used but also to the division of variables picked. One may find, for instance, that $\boldsymbol{Z}$ minus W is exogenous to $\boldsymbol{Y}$ plus W, for some variable W. It is also possible that missing variables can disrupt the exogenous interpretation, as when $\boldsymbol{Z}$ is exogenous to $\boldsymbol{Y}$ but not to some extended $\boldsymbol{Y}$. The possibility of 'instantaneous causality' obviously greatly complicates the problem of how to test for exogeneity. Some of these problems have been discussed elsewhere [Granger (1980)] and so will not be followed up here. Some variables are such that prior beliefs will be strong that they are exogenous: an example is that weather is probably exogenous to the economy. However, other variables have often been considered to be exogenous yet need to be tested, the best examples being the control variables. One can argue that a government controlled interest rate is in fact partly determined by previous movements elsewhere in the economy, and so is not strictly exogenous. The true exogenous part of such a variable is that which cannot be forecast from other variables and its own past, and it follows that it is only this part that has any policy impact. The theory of rational expectations, currently attracting a lot of attention in economics, is relevant here but its discussion is not really appropriate.

The effect of the presence of control variables on causal relationships was considered by Sims (1977). It is certainly possible that the actions of a controller can lead to what appears to be a causal relationship between two variables. Equally, it is possible that two variables that would be causally related if no controls were used, would seem to be unrelated in the presence of a control. It is also worth pointing out that controlability is a much deeper property than causality, in my opinion, although some writers have confused the two concepts. If Y causes X, it does not necessarily mean that Y can be used to control X. An example is if one observes that the editorial recommendations of the New York Times about which candidates to support causes some voters to change their votes. However, if one started controlling these editorials, and this became known, the previously observed causality may well disappear. The reason is clearly that the structure has been altered by changing a previously uncontrolled variable to one that is controlled. If causation is found between a controlled variable and something else, this could be useful in deciding how to control, provided movements are kept near those observed in the past. It seems quite possible that some variables used in the past by governments to control may be so ineffectual that causation will not be found, so testing is worthwhile. The relationship between control, causation and the recent rational expectations literature is potentially an interesting one, but is too large a topic to be considered here.

There is clearly much more discussion required of this and other definitions and more experience required with the various methods of testing that have been suggested. It is my personal belief that the topic is of sufficient importance, and of interest, to justify further work in this field.

References

Ashley, R., C.W.J. Granger and R. Schmalensee, 1979, Advertising and aggregate consumption: An analysis of causality, Research report (Department of Economics, University of California, San Diego, CA).

Black, H., 1978, Inflation and the issue of unidirectional causality, Journal of Money, Credit and Banking X, Feb., 99–101.

Blalock, H.M. Jr., 1964, Causal inferences in non-experimental research (University of North Carolina Press, Chapel Hill, NC).

Box, G.E.P. and G.M. Jenkins, 1970, Time series analysis, forecasting and control (Holden Day, San Francisco, CA).

Bunch, M., 1963, Causality (Meridian Books, Cleveland, OH).

Caines, P.E. and C.W. Chan, 1975, Feedback between stationary stochastic processes, IEEE Transactions on Automatic Control AC 20, 498–508.

Caines, P.E., S.P. Sethi and T. Brotherton, 1977, Impulse response identification and causality detection for the Lydia–Pinkham data, Annals of Economic and Social Measurement 6, Spring, 147–164.

Chiang, C., 1978, An investigation of the relationship between price series, Ph.D. thesis (University of California, San Diego, CA).

Ciccolo, J.H. Jr., 1978, Money equity values and income: Tests for exogeneity, Journal of Money, Credit and Banking X, Feb., 45–64.

Feige, F.L. and D.K. Pearce, 1976a, Inflation and incomes policy: An application of time series models, Journal of Monetary Economics, Supplementary Series 2, 273–302.

Feige, E.L. and D.K. Pearce, 1976b, Economically rational expectations: Are innovations in the rate of inflation independent of innovations in measures of monetary and fiscal policy?, Journal of Political Economy 84, 499–522.

Geweke, J., 1978, Testing the exogeneity specification in the complete dynamic simultaneous equation model, Journal of Econometrics 7, no. 2, 163–186.

Good, I.J., 1961/62, A causal calculus, I/II, British Journal of Philosophical Society 11, 305/12, 43–51.

Gordon, R.J., 1977, World inflation and monetary accommodation in eight countries, Brookings Papers on Economic Activity, Part 3, 409–478.

Granger, C.W.J., 1963, Economic processes involving feedback, Information and Control 6, 28–48.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37, 424–438.

Granger, C.W.J., 1980, Generating mechanisms, models and causality, Paper presented to World Econometrics Congress, Aix-en-Provence, Sept. 1980.

Granger, C.W.J. and A. Andersen, 1978, An introduction to bilinear time series models (Vandenhoeck and Ruprecht, Göttingen).

Granger, C.W.J. and M. Hatanaka, 1964, Spectral analysis of economic time series (Princeton University Press, Princeton, NJ).

Granger, C.W.J. and P. Newbold, 1977, Forecasting economic time series (Academic Press, New York).

Hart, H.L.A. and A.M. Honoré, 1959, Causation in the law (Oxford University Press, Oxford).

Hosoya, Y., 1977, On the Granger condition for non-causality, Econometrica 45, no. 7, 1735–1736.

Mehra, Y.P., 1977, Money wages, prices and causality, Journal of Political Economy 85, Dec., 1227–1244.

Newbold, P., 1978, Feedback induced by measurement errors, International Economic Review 19, 787–791.

Overseth, O.E., 1967, Experiments in time reversal, Scientific American 221, Oct., 88–101.

Pierce, D.A., 1975, Forecasting in dynamic models with stochastic regression, Journal of Econometrics 3, 349–374.

Pierce, D.A. and L.D. Haugh, 1977, The assessment and detection of causality in temporal systems, Journal of Econometrics 5, 265–293.

Pierce, D.A. and L.D. Haugh, 1979, Comment on price, Journal of Econometrics 10, 257–260.

Price, J.M., 1979, A characterization of instantaneous causality: A correction, Journal of Econometrics 10, 253–256.

Sargent, T.F., 1976, A classical macroeconometric model for the U.S., Journal of Political Economy 84, 207–238.

Sims, C.A., 1972, Money, income and causality, American Economic Review 62, 540–552.

Sims, C.A., 1977, Exogeneity and causal ordering in macroeconomic models, in: New methods in business cycle research, Proceedings of a conference (Federal Reserve Bank, Minneapolis, MN).

Skoog, G.R., 1976, Causality characterizations: Bivariate, trivariate and multivariate propositions, Staff Report no. 14 (Federal Reserve Bank, Minneapolis, MN).

Suppes, P., 1979, A probabilistic theory of causality (North-Holland, Amsterdam).

Wiener, N., 1958, The theory of prediction, in: E.F. Bec, ed., Modern mathematics for engineers, Series 1, Ch. 8.

Williams, D., C.A.E. Goodhart and D.H. Gowland, 1976, Money, income and causality: The U.K. experience, American Economic Review 66, 417–423.

Wright, S., 1964, The interpretation of multivariate systems, in: O. Kempthorne et al., Statistics and mathematics in biology, Ch. 2 (Hafner, New York).

Zellner, A., 1978, Causality and econometrics, in: Proceedings of conference held at University of Rochester, NY, forthcoming.