Comment

CLIVE GRANGER*

If causality can be equated with some measurable quantity, then statisticians should be able to devise tests for causation. I believe this to be an important topic but it is one that statisticians have been rather unhelpful about, even negligent, in the past. Only a few statisticians have attempted to discuss this difficult field, and thus I welcome Holland's article as a useful further contribution.

To appreciate his work it is important to consider first a variety of causal-type questions. If one looks at units of a population and asks why one unit has a different value of some variable than another, this is a cross-sectional question. An example is when one tries to explain why one household spends more on electricity, say, than does another household. Many causality questions, however, are of quite a different nature, such as asking why electricity demand has fallen this year or why crime rates have increased. Here, one is asking causal questions about data that are a group of time series. In cross-sectional causation one is asking why a particular unit is placed in a certain part of the distribution for the variable of interest. In temporal causality one is asking why parameters of that distribution have changed through time. The two types are very different in nature and probably require different definitions and methods of analysis.

Holland's article deals just with cross-sectional causal questions, as is clear from the discussion of association in Section 2 and the causal model in Section 3. The use of experiments to illuminate statistical questions have a venerable past and an experimental viewpoint seems to be an entirely sensible one for consideration of cross-sectional causation. I found Holland's discussion in Section 3 very helpful and largely convincing for the particular class of questions being asked. Of course, the experiment actually has to take place for the analysis; hypothetical experiments will not be relevant. According to this article, it is also required that the treatment variable actually can be controlled for all units of the population. It follows that one cannot tackle questions such as whether race or sex affects income or crime rates. Thus many causal questions cannot be tackled within this framework, such as most of those arising in history, economics, sociology, meteorology, oceanology, political science, anthropology, or law. This is, of course, a serious limitation. Examples of topical importance are the questions of whether pornography causes changes in rape rates and whether the death sentence causes decreases in murder rates.

There are some very important advantages of trying to analyze causality by experimentation. One can hold constant, or at least potentially control, many other variables that otherwise could be disturbing so that it is not necessary to condition on these variables during the analysis. Further, one does know which is the treatment variable and which is the experiment's outcome variable. This is not always true in nonexperimental situations—for example, does crime cause poverty or does poverty cause crime? It might also be noted that the value of the treatment variable—the cause—is determined before the experiment starts, and thus before the output variable is observed, and this will be known in an experiment. There are also difficulties with experiments, however. Human subjects may behave differently in experimental situations than in the real world, making findings not easily transferable and so of limited value. Further, some "irrelevant" variables may be controlled and disturb the actual causal relationship. For example, when studying the effects of a price raise on consumption, if the hours worked by consumers and hence their incomes are kept constant, the wrong causal implication may be reached. One also cannot ask questions about two-way causation, such as poverty causing crime and crime causing poverty.

I am rather surprised that Holland concentrates his attention on the differences between the two means of Y_t and Y_c, whereas other differences in the distributions of these two variables, such as variances, could be very important to someone reaching a decision on the basis of the experiment. After all, in decision making under uncertainty, risk is as important as expected return.

My own particular interest is in temporal causality. I think that necessary conditions for a cause are that it occurs before the effect and contains unique information about it. From these ideas, it follows that knowing the cause helps forecast many aspects of the effect, and tests can be based on this simple idea. I do not see that the experimental context contradicts these ideas. I have also tried to emphasize that the purpose of causal analysis, including statistical analysis, is to try to change people's "degrees of belief," which might be conveniently summarized as a probability that a suggested causal relationship is true. These beliefs are required for decision making by economic agents. These views are expanded in two papers: Granger (1980, 1985).

The question obviously arises whether or not the experimental framework used here is also relevant for testing temporal causality. We may think of two types of experimental units—those with memory and those without. Examples of units without memory would be physical objects and possibly land or lower animals, the classical units used in experimental work. Certainly human subjects will have memory, as will many animals. If a unit has memory, it will be very difficult to devise a time sequence of experiments obeying Holland's requirements to test some theory about the effects of a price change or income level on consumption, say, because the idea of potentially being

* Clive Granger is Professor, Department of Economics, University of California, San Diego, CA 92093.

Journal of the American Statistical Association
December 1986, Vol. 81, No. 396, Theory and Methods

able to place every unit under every value of the controlled variable at every moment of time becomes less plausible. We are back to not being able to relate race or sex to income. Because of the memory, it seems that all such experiments become strictly impossible, as what happened in the past will potentially affect the outcome of the present experiment. It seems to me that most of the solutions to what Holland calls the "Fundamental Problem of Causal Inference" will no longer work in this case, including the "statistical solution," without conditioning on the past. I am thus unclear that the experimental model is even theoretically helpful for temporal causality in the behavioral sciences. If one does condition on the past, the statistical solution may be relevant, but the basis for the inference will then be quite different from that proposed here.

ADDITIONAL REFERENCE

Granger, C. W. J. (in press), "Causality Testing in a Decision Science," to appear in proceedings of the "Conference on Probability and Causality," held at the University of California, Irvine, July 1985.

Rejoinder

PAUL W. HOLLAND

I thank all of the discussants for their very thoughtful comments. Not surprisingly, I agree more with the views expressed by Cox and Rubin than with those of Glymour and Granger, but each discussant makes important points that expand and illuminate issues that arise in the article. Space does not permit a response to every point mentioned, and the more critical comments of Glymour and Granger tend to be balanced by the comments of Cox and Rubin. Hence I will restrict my rejoinder to those issues that I feel need emphasis or to which I feel I can add a useful point of view.

In reflecting upon the discussants' remarks I realized that nowhere in the article, or elsewhere, is there a purely mathematical description of Rubin's model. Such a formulation ought to help separate the *model* itself from its *applications*. For this reason I will begin my rejoinder with a brief, mathematical statement of Rubin's model and its interpretation in terms of my article. Then I will address some of the issues raised by each discussant.

1. A MATHEMATICAL STATEMENT OF RUBIN'S MODEL

In its simplest form, stripped of all of the interpretative language, Rubin's model is a quadruple, $R = (U, K, Y, S)$, in which U and K are sets, Y is a real-valued function defined on $U \times K$, and S is a mapping from U to K. In the language of the article the meaning of the components of R is as follows. U is the population of units, and K is a set of labels or descriptions of the various causes or treatments under consideration. For any $u \in U$ and $k \in K$, $Y(u, k)$ is the value of the response that would be measured on u if u were exposed to cause k. The value of $S(u)$ is the cause or treatment to which u is actually exposed prior to the measurement of the response. In the article I used the equivalent subscript notation, that is, $Y_k(u) = Y(u, k)$, and I let $K = \{t, c\}$. Of course, in general K could contain more than two elements.

In real applications of Rubin's model other measurements besides the response Y need to be represented. I think that all measurements should be regarded as functions defined on $U \times K$, just as Y is. If X is such a function, then $X(u, k)$ is the value of the X measurement that would be made on u if u were exposed to cause $k \in K$. One special type of measurement needs mention here. If the value of $X(u, k)$ does not depend on which cause k to which u is exposed I shall call X an *attribute* of u; that is, $X(u, k) = X(u)$ for all $u \in U$ and $k \in K$. Important examples of attributes are (a) pre-exposure variables (Sec. 3) and (b) post-exposure variables that cannot be affected by k. Among the measurements that are *not* attributes I include other response variables besides Y and "post-treatment concomitant variables" (Rosenbaum 1984b).

The purpose of Rubin's model is to provide a language for discussing causation, and this language takes *units*, *causes*, and *responses* as primitive notions that are not defined further. These three elements, however, are not arbitrary and must satisfy the basic property that Y is defined on all of $U \times K$. The *effect* of cause t relative to c is then defined in terms of these primitive notions, that is, as $Y(u, t) - Y(u, c)$, and the observed response on each unit is also defined in terms of the elements of R, that is, $Y_S(u) = Y(u, S(u))$.

By taking units, causes, and responses as the primitives of his theory and defining effects and observed data in terms of them, Rubin's model breaks with an ancient philosophical tradition that takes "events" or "phenomena" as primitives and attempts to define what is meant by one event being *the* cause of another.

An *application* of Rubin's model requires an identification of the elements of R with features of a real-world problem. What are the units, the causes, the responses? How are units actually exposed to the action of the causes? Is Y defined on all of $U \times K$? If the identification of the elements of the real-world application with those of Rubin's model leads to a faithful representation of the real-

Journal of the American Statistical Association
December 1986, Vol. 81, No. 396, Theory and Methods